

ACQUISITION OF STRUCTURE AND INTERPRETATION:
CASES FROM MANDARIN BARE AND NON-BARE NOUN PHRASES

By

Hsiang-Hua Chang

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Linguistics

2011

ABSTRACT

ACQUISITION OF STRUCTURE AND INTERPRETATION: CASES FROM MANDARIN BARE AND NON-BARE NOUN PHRASES

By

Hsiang-Hua Chang

Children's production of bare nominals is universal. When acquiring languages disallowing bare nominals, children will develop from the bare to the non-bare stage. However, Mandarin nominals may appear bare or non-bare in various positions with all kinds of interpretations. This dissertation conducts two acquisition studies to examine the production and interpretation of bare and non-bare nominals in Mandarin.

The production study examined the spontaneous speech data produced by two Mandarin-speaking children: a girl (2;0-2;6) and a boy (2;10-3;3). Distributional analyses and variation analyses using GoldVarb were conducted to compare children's production of bare and non-bare nominals, taking into consideration various syntactic and semantic aspects. The results show that four variables – MLU, Interpretation, Verb Type, and Aspect Marker – have significant effects on the variation of bare and non-bare nominals. Within the nominal phrase, (1) adjective is the first element children add to a bare noun root (age 2;0). (2) Possessives nominals emerge as early as age 2;1. (3) The order of appearance frequency: possessives, classifiers, numerals, and demonstratives. (4) Non-bare nominals most frequently associate with the object position and the existential interpretation.

The comprehension study tested the interpretation of bare and demonstrative nominals among 110 Mandarin-speaking children and adults. The study finds that nominal types and age have significant effects on the interpretation. The findings show

that (1) children, like adults, have both generic and existential definite readings for bare nominals. (2) They distinguish between bare and demonstrative nominals by assigning more generic interpretations to bare nominals. (3) They prefer generic readings for bare nominals, as opposed to existential definite readings. Demonstrative nominals also receive a considerable amount of generic interpretations. (4) Two variables significantly affect the interpretation of nominals: pragmatics and non-linguistic properties of the predicates.

This study discovers a non-target generic interpretation of Mandarin demonstrative nominals assigned by both children and adults. The individual analysis finds a significantly positive correlation between the generic reading for demonstratives and that for bare nominals, which suggests that the generic reading of demonstratives may result from individual preference. The non-target generic interpretation may also be accounted for semantically and syntactically: (1) Mandarin-speaking children treat demonstratives as a less-specified determiner like the Spanish definite; (2) Mandarin-speaking children project demonstrative nominals as NP, not DP.

The production study discovers variables that significantly associate with the emergence of the non-bare nominals and the order of emergence of various nominal-internal elements. The comprehension study reveals empirical data of the interpretation of bare and demonstrative nominals. This dissertation contributes to the understanding of the acquisition of Mandarin noun phrases and sheds light on further intralinguistic and crosslinguistic research.

Copyright by
HSIANG-HUA CHANG
2011

ACKNOWLEDGMENTS

Walking through the darkest valley, I am grateful to finally see the light at the end of the tunnel. My very first class in linguistics, also my very first class in the United States, took place in a Monday morning. That was Introduction to Syntax, taught by my dissertation advisor, Cristina Schmitt. How fitting it is to start and end my journey at Michigan State with her. I am especially thankful for her support when I gave up the research that I had been working for about two years and did not know what to do next. It would be impossible to write this dissertation without her help, guidance, and patience. My deepest gratitude also goes to my committee members, Barbara Abbott, Yen-Hewi Lin, Alan Munn, and Ana Teresa Pérez-Leroux. I enjoyed all of their classes and am grateful for their insightful comments and thought provoking discussions. Barbara Abbott is one of the best teachers that I have had in my life. I wish someday I can teach and write as well as she. I am also indebted to Cristina Schmitt and Yen-Hwei Lin for their generous help at every stage of my job search.

I would also like to thank Alan Beretta, whose class I enjoyed very much. There are other wonderful teachers in the department that I did not have an opportunity to take their classes. I thank Grover Hudson, Mutsuko Endo-Hudson, and John Hale for good chats on the hallway and in the elevator. I give thanks to Susanne Evens Wagner for introducing me to a great book “Analyzing Sociolinguistics Variation” by Sali Tagliamonte, from which I learned how to do language variation and use GoldVarb from scratch. The study regarding free speech data in this dissertation cannot be done without such a good book. Thanks also go to Xiaoshi Li, who came to our department when I just

started to learn how to use GoldVarb. She provided good and timely suggestions for my GoldVarb analysis. I owe my deepest gratitude to Haiyong Liu for giving me a precious opportunity to teach linguistics courses at Wayne State University. The valuable experience helped build up my confidence at the later stage of my job search. I am also grateful for his genuine interest in my dissertation project. Thanks also go to Miao-Ling Hsieh for kindly sharing her monograph about Mandarin noun phrases and good advice. I wish I could have spent more time reading her book. I will definitely do so in the near future.

I would also like to thank all of my experiment participants and parents. It was very kind of Yi-ching Su and I-ru Su to allow me to run experiments in their classes. I am especially grateful for Yi-ching's generous encouragement and friendship, not to mention her expertise on language acquisition. You-Shan Chung gave me a listening ear and understanding heart on the cloudy days of becoming language acquisitionists, and I am truly grateful for that.

I am very thankful to the good staff in the department, as well as for the teaching assistantship and research assistantship I was offered. I thank my "short-term" officemates, Dennis Preston and Tara Sanchez, for good laughs. I was fortunate to have shared my office with these two wonderful linguists. I would also like to thank my colleagues: Irina Agafonova, Boonjeera Chiravate, Bo-Young Kwon, Chun (Jimmy) Huang, Yu-fen Liou, Ai Matsui, Sachie Miyazaki, Nao Nakano, Diane Ogiela, Li-Jen Shih, Walter Sistrunk, Mingyu Sun, Masaaki Takashima, Malinee Vongsakulkasem, and many more for good sharing – especially, Chunhua Ma, Karen Miller, Ming Xiang, and Xiaofei Zhang for the help and heart-felt encouragement.

It is an honor for me to work with Linda Cooke Johnson. I was fortunate to be her research assistant for her book project about Chinese history. Working with her was truly a pleasant experience. I thank her for helping me with the job search, allowing me to sit-in one of her classes, and generously sharing with me all of her presentation files and the precious pictures and images that she had been collecting for many years.

Landing my current job at Oakland University would not have been possible if not for her help, because it requires teaching a Chinese civilization class. I would also like to thank Hsiao-ping Wang and Mei-hsiao Tang for helping me with the job search. I give thanks to my previous and current "bosses" at Oakland University, Chris Clason, Stephen Filler, Seigo Nakao, and Ron Rapin, for their encouragement and understanding.

I am also indebted to my friends and family members. I thank Fei-Ya Chen, Ting Hong, Hsiu-Shuo Hu, Shaohua Mu, Ying Qin, Anne Tseng, Lydia Yao, my mother, and many other friends for the faithful prayers. In addition, my two best friends for almost thirty years, Hui-Ling Su and Sz-Ying Lu, have been offering their unfailing friendship and listening ears throughout this journey. I am thankful to my relatives in Taiwan, especially my father, mother, brother, sister, brother-in-law, and sister-in-law for their moral and financial support. I give special thanks to my sister-in-law, who helped recruit children subjects. People can always realize who the true friends are in times of challenge and difficulty. My sincere regards and blessings go to those who supported me in any respect throughout this tough journey.

My biggest debt is to my husband, Cheng-Chieh Hsu, who has been walking side by side with me every step of the way in this journey. I could not have done it without him. He shared with me all those long and sleepless nights, the arduous weekends, the

lost health, and the stress and despair resulted from such a huge project. Although with a very demanding job and sometimes having to work over eighty hours a week, he still managed to help me in fixing computer problems, taking part in my experiments, proof-reading my manuscripts, and most important of all, listening to all of my whining. He has been my best friend, playing multiple roles as my research assistant, computer technician, writing consultant, and psychiatric counselor. He has been remarkably supportive and unselfish. I know he is far happier than I am now that the journey is about to end. Without him in my life, a doctoral degree is not only impossible but also unthinkable. All glory to God for He is merciful and able.

TABLE OF CONTENTS

LIST OF TABLES.....	XII
LIST OF FIGURES.....	XIV
ABBREVIATIONS.....	XVI
CHAPTER ONE	
INTRODUCTION.....	1
CHAPTER TWO	
MANDARIN NOUN PHRASES.....	6
2.1 STRUCTURE OF MANDARIN BARE AND NON-BARE NOMINALS	6
2.2 INTERPRETATION OF MANDARIN BARE NOMINALS	9
2.2.1 Chierchia (1998)	9
2.2.2 Cheng and Sybesma (1999)	14
2.3 SUMMARY AND DISCUSSION	17
CHAPTER THREE	
ACQUISITION ASSUMPTIONS	21
3.1 DEFAULTS AND THE SUBSET PRINCIPLE.....	22
3.2 CHERCHIA, GUASTI, AND GUALMINI (1999)	24
3.3 ROEPER (2006).....	27
3.4 SUMMARY	31
CHAPTER FOUR	
PRODUCTION OF BARE AND NON-BARE NOUN PHRASES	34
4.1 INTRODUCTION.....	34
4.2 LINGUISTIC BACKGROUND.....	35
4.3 HYPOTHESES AND PREDICTIONS FOR EMERGENCE AND INTERPRETATION OF NON- BARE NOMINALS	43
4.3.1 The Element/Structural Projection that First Appears above NP	44
4.3.2 The Argument Position that Non-Bare Nominals First Appear.....	47
4.3.3 The Interpretation that Non-Bare Nominals Most Frequently Associated with	49
4.3.4 Summary	50
4.4 METHODOLOGY.....	51
4.4.1 Data and Subjects.....	51
4.4.2 The GoldVarb Analysis	52

4.4.3 Defining Variables	53
4.4.3.1 Developmental and semantic variables.....	58
4.4.3.2 Variables of the nominal internal structures	60
4.4.3.3 Variables of the nominal external structures	61
4.5 RESULTS	71
4.5.1 Distributional Analyses.....	72
4.5.1.1 General distribution and developmental variables.....	72
4.5.1.2 Variables of the nominal domain.....	74
4.5.1.3 Cross-tabulation between nominal internal variables and age.....	77
4.5.1.4 Cross-tabulation among nominal internal variables	82
4.5.1.5 Summary and Discussion.....	87
4.5.2 Multivariate Analyses	89
4.5.2.1 General result.....	89
4.5.2.2 Effect of the variables in the verbal domain	92
4.5.2.3 Effect of the variables in the clausal domain.....	99
4.5.2.4 Effect of the semantic variable	101
4.5.2.5 Summary	102
4.6 GENERAL DISCUSSION AND CONCLUDING REMARKS.....	102

CHAPTER FIVE

COMPREHENSION OF MANDARIN BARE AND DEMONSTRATIVE NOUN

PHRASES	106
5.1 INTRODUCTION.....	107
5.2 LINGUISTIC BACKGROUND.....	109
5.2.1 Linguistic Characteristics of Generic Expressions	109
5.2.2 Generic Noun Phrases.....	112
5.3 ACQUISITION BACKGROUND	118
5.3.1 Acquisition of Generics and Definites	119
5.3.2 Gelman and Raman (2003)	123
5.3.3 Pérez-Leroux, Munn, Schmitt, and DeIrish (2004)	124
5.4 RESEARCH QUESTIONS AND HYPOTHESES	127
5.5 A COMPREHENSION EXPERIMENT	133
5.5.1 Stimuli and Procedure.....	134
5.5.2 Subjects	137
5.5.3 Results.....	138
5.5.4 Individual Analyses and Discussion	144
5.5.5 Summary	151
5.6 GENERAL DISCUSSION AND CONCLUDING REMARKS.....	151
5.6.1 Interpretation of Bare and Demonstrative Noun Phrases	153
5.6.2 Variables Affecting the Interpretation	154
5.6.3 Non-target Interpretation of Demonstrative Noun Phrases	155

CHAPTER SIX

CONCLUSION

6.1 GOAL AND HYPOTHESES OF THE DISSERTATION	159
---	-----

6.2 SUMMARY AND CONCLUDING REMARKS.....	160
APPENDICES	165
APPENDIX I DATA TRANSCRIPTION AND CALCULATION OF MEAN LENGTH OF UTTERANCES.....	166
APPENDIX II DATA EXTRACTING AND CRITERIA FOR ELIMINATING UNWANTED SENTENCES.....	170
APPENDIX III TARGET QUESTION SETS FOR THE COMPREHENSION EXPERIMENT ...	173
APPENDIX IV TARGET QUESTIONS FOR THE KNOWLEDGE TEST.....	178
REFERENCES	179

LIST OF TABLES

Table 2.1	Mandarin nominals	18
Table 4.1	Syntactic positions, verb forms, and nominal types	42
Table 4.2	Interpretation and nominal types	43
Table 4.3	Information of Nana's data	51
Table 4.4	Information of Didi's data	51
Table 4.5	Variables in analyzing Mandarin production data	54
Table 4.6	Nana's tokens of coded nominals across age	71
Table 4.7	Didi's tokens of coded nominals across age	71
Table 4.8	Overall distribution of bare and non-bare nominals	72
Table 4.9	Distribution of adjective-like factors in bare and non-bare nominals	74
Table 4.10	Distribution of nominal internal factors in non-bare nominals	75
Table 4.11	The use of adjective-like modifiers in all nominals across age	78
Table 4.12	The use of modification above NP in non-bare nominals across age	78
Table 4.13	The use of classifiers in non-bare nominals across age	78
Table 4.14	The use of determiners in non-bare nominals across age	80
Table 4.15	Cross-tabulation of adjective-like modifiers and classifiers in non-bare nominals	83
Table 4.16	Cross-tabulation of modification above NP and classifiers in non-bare nominals	83
Table 4.17	Cross-tabulation of determiners and classifiers in non-bare nominals	84
Table 4.18	Cross-tabulation of determiners and adjective-like modifiers in non-bare nominals	86
Table 4.19	Cross-tabulation of determiners and modification above NP in non-bare nominals	86

Table 4.20	Multivariate analysis of the contribution of each factor to the probability of non-bare nominal occurrence	90
Table 4.21	Cross-tabulation of post-copular position and interpretations	92
Table 4.22	Cross-tabulation of activity verbs and interpretations	95
Table 4.23	Cross-tabulation of activity verbs and syntactic positions	95
Table 4.24	Cross-tabulation of object position and aspect	96
Table 5.1	Interpretation of subject nominals with characterizing statements in English, Spanish, and Mandarin	116
Table 5.2	Subjects of the Experiment in Version A and B	137
Table 5.3	Subjects of the Experiment	138
Table 5.4	Percentage of generic responses: presentation order	139
Table 5.5	Percentage of generic responses: property of predicates	142
Table 5.6	Listing of individual responses of adults	146
Table 5.7	Listing of individual responses of three-year-olds	147
Table 5.8	Listing of individual responses of four-year-olds	147
Table 5.9	Listing of individual responses of five-year-olds	148

LIST OF FIGURES

Figure 4.1	MLU of Nana and Didi.....	52
Figure 4.2	Distribution of non-bare nominals across age	73
Figure 4.3	Percentage of each classifier type in non-bare nominals.....	80
Figure 4.4	Percentage of each determiner type in non-bare nominals.....	82
Figure 5.1	English experiment example	124
Figure 5.2	The counterbalancing design in version A and B of the comprehension experiment.....	135
Figure 5.3	Mandarin experiment example	136
Figure 5.5	Mean proportion of generic responses	138
Figure 5.6	Percentage of generic responses for bare nominals: presentation order.....	140
Figure 5.7	Percentage of generic responses for demonstrative nominals: presentation order	141
Figure 5.8	Percentage of generic responses for bare nominals: property of predicates	142
Figure 5.9	Percentage of generic responses for demonstrative nominals: property of predicates	143
Figure 5.10	Individual adult's generic reading	148
Figure 5.11	Individual child's generic reading	149
Figure 5.12	Percentage of people who prefer generic readings.....	150
Figure AIII.1	Question Set 1	173
Figure AIII.2	Question Set 2.....	174
Figure AIII.3	Question Set 3.....	174
Figure AIII.4	Question Set 4.....	175
Figure AIII.5	Question Set 5.....	175

Figure AIII.6 Question Set 6.....	176
Figure AIII.7 Question Set 7.....	176
Figure AIII.8 Question Set 8.....	177

ABBREVIATIONS

A	affirmation question particle <i>a</i>
arg	argument
BA	direct object marker <i>ba</i>
CL	classifier
CIP	Classifier Phrase
D	determiner
DE	pre-nominal modification marker or postverbal resultative marker <i>de</i>
Dem	demonstrative
DOU	universal quantification adverb <i>dou</i>
DP	Determiner Phrase
DUR	durative marker <i>-zhe</i>
EXP	experiential marker <i>-guo</i>
ILP	Individual-Level Predicate
MA	yes/no question particle <i>ma</i>
MLU	Mean Length of Utterances
N	noun
NP	Noun Phrase
Num	numeral
NumP	Numeral Phrase
PERF	perfective marker <i>-le</i> or <i>you</i>
PL	plural
pred	predicate
PROG	progressive marker <i>zai-</i>
Q	quantifier
QP	Quantifier Phrase
RC	relative clause
SFP	sentence final particle
Sing.	singular
SLP	Stage-Level Predicate
SVO	Subject Verb Object
UG	Universal Grammar
VP	Verb Phrase

CHAPTER ONE

INTRODUCTION

The behavior of bare and non-bare nominals within and across languages have motivated many syntactic, semantic, and acquisition studies (Cheng and Sybesma 1999, Chierchia 1998, Chierchia et al. 1999, Longobardi 1994, to name a few). Since Cheng and Sybesma's classic study (1999) of Mandarin bare and non-bare nominals, the interest about the internal structure of Mandarin nominals has increased (Hsieh 2008, Sio 2006, Yang 2001); however, the studies focusing on the development of Mandarin bare and non-bare nominals in child language remain very few.¹ This dissertation fills part of this gap by exploring children's production and interpretation of bare and non-bare nominals in Mandarin.

Bare nominals without any inflection are forbidden in argument position in some languages. In Mandarin, a language without articles and without obligatory plural morphology,² bare nominals are not only acceptable but also able to convey various interpretations (generic, definite, or indefinite) and appear in various syntactic positions (argument or non-argument). Examples (1), (2), and (3) show that Mandarin bare nominals can function as arguments – subject (1), object (2), or object of a preposition (3)

¹ The current research investigates the behavior of Mandarin Chinese, which is the official language spoken in China, Taiwan, Hong Kong, and Singapore. Other studies cited in this research may refer to it as either Chinese or Mandarin. For example, in Chierchia (1998), only 'Chinese' is used, not 'Mandarin.' For consistency, only 'Mandarin' will be used in this research.

² The Mandarin plural suffix *-men* has a very restricted distribution. It is only obligatory when marking the plurality of pronouns, such as *wo-men* 'we' and *ta-men* 'they.' *Men* can also appear with human terms to form a definite description as in *haizi-men* 'the children.' Except for the above situations, *-men* cannot be used with other common nouns (Li and Thompson 1981).

– and can be interpreted as generic, definite, indefinite, singular or plural. Example (4) illustrates that Mandarin bare nominals may also function as non-arguments, such as predicates.

(1) Subject:

houzi	chi	xiangjiao	
monkey	eat	banana	
'Monkeys eat bananas.'			(generic)
'The monkey eats/ate bananas.'			(singular definite)
'The monkeys eat/ate bananas.'			(plural definite)

(2) Object:

wo	yao	qu	mai	pingguo	
I	want	go	buy	apple	
'I am going to buy the apple/an apple.'					(singular definite/indefinite)
'I am going to buy the apples/apples.'					(plural definite/indefinite)

(3) Object of preposition:

zhe-zhong	yifu	shi	zuo	gei	daren	chuan-de
this-kind	clothes	is	make	for	adult	wear-DE
'This kind of clothes is made for adults.'						

(4) Predicate:

Mimi	shi	mao	;	Wangwang	shi	gou
Mimi	be	cat		Wangwang	be	dog
'Mimi is a/the cat; Wangwang is a/the dog.'						

As bare nominals, Mandarin non-bare nominals can also appear in different syntactic positions and have various semantic functions. Except the generic reading of example (1) that requires a bare subject, sentences (1) to (4) may also be expressed by using various types of non-bare nominals, such as singular/plural demonstrative nominals, numeral nominals or classifier nominals, as illustrated in (5) to (8).

(5) Subject:

a. **na-zhi** **houzi** chi xiangjiao
that-CL monkey eat banana
'That monkey eats/ate bananas.' (singular definite)

b. **naxie** **houzi** chi xiangjiao
those monkey eat banana
'Those monkeys eat/ate bananas.' (plural definite)

(6) Object:

a. wo yao qu mai **na-ke/yi-ke** **pingguo**
I want go buy that-CL/one-CL apple
'I am going to buy that apple/an apple.' (singular definite/indefinite)

b. wo yao qu mai **naxie/yixie** **pingguo**
I want go buy those/some apple
'I am going to buy those apples/some apples.' (plural definite/indefinite)

(7) Object of preposition:

zhe-zhong yifu shi zuo gei **suoyoude** **daren** chuan-de
this-kind clothes is make for all adult wear-DE
'This kind of clothes is made for all the adults.'

(8) Predicate:

a. Mimi shi **zhi** **mao**; Wangwang shi **zhi** **gou**
Mimi be CL cat Wangwang be CL dog
'Mimi is a cat; Wangwang is a dog.'

b. Mimi shi **na-zhi** **mao**; Wangwang shi **na-zhi** **gou**
Mimi be that-CL cat Wangwang be that-CL dog
'Mimi is that cat; Wangwang is that dog.'

From the acquisition standpoint, the production of bare nominals is found to be a universal phenomenon in early child speech regardless of their acceptability in the adult target language. Determiners tend to be omitted in early child speech, much like other functional categories (Brown 1973, Radford 1990, Rizzi 1993/4, 1994, Wexler 1994, Chierchia, Guasti, and Gualmini 1999, and Guasti 2000, among others). However, the occurrence of bare nominals in child Mandarin cannot be attributed to determiner

omission because bare nominals are legitimate in Mandarin, with a variety of interpretations in various positions. Besides, Mandarin, as is well known, does not have a definite or an indefinite article. The language-specific challenges that Mandarin poses to language acquisition motivate the research questions for the current study.

Challenge 1. Production

Since bare nominals are able to convey almost all kinds of interpretations in an appropriate context, what are the contexts that trigger the emergence of non-bare nominals?

Challenge 2. Interpretation

- a. How do Mandarin-speaking children interpret bare nominals given that bare nominals can have various interpretations?
- b. In contrast, what is the interpretation that children assign to non-bare nominals, such as demonstrative nominals?
- c. Do children have the same interpretation as adults do?

The goal of this dissertation is to fill in some voids in the study of the acquisition of noun phrases in Mandarin by conducting two acquisition case studies.

This dissertation is organized as follows. Chapter 2 presents the assumptions about the Mandarin nominal structure adopted in this dissertation. It also discusses the possible accounts of the interpretation of Mandarin bare nominals (non-bare nominals will also be discussed wherever applicable). Chapter 3 briefly presents the assumptions about the acquisition of nominals adopted in the dissertation. Chapter 4 presents a study of the longitudinal spontaneous speech data of two Mandarin-speaking children. I present findings from a variation analysis using GoldVarb (Robinson, Lawrence, and Tagliamonte 2001) to compare children's production of bare and non-bare nominals, taking into consideration various syntactic and semantic aspects. Chapter 5 presents a comprehension study, testing Mandarin-speaking children's and adults' interpretation of bare and demonstrative nominals. This study is controlled in such a way that only the

generic and existential definite readings of the bare nominals are possible. Mandarin lacks definite articles. In order to compare the existential definite reading of bare nominals, demonstrative nominals – a type of definite expressions – are used in the experiment. This chapter discusses how children's interpretation of bare and demonstrative nominals deviates from that of adults. Chapter 6 concludes with a summary of findings and suggestions for future research.

CHAPTER TWO

MANDARIN NOUN PHRASES

2.1 Structure of Mandarin Bare and Non-Bare Nominals

The syntactic representation of Mandarin nominals has been subject to much debate since the work by Abney (1987). This dissertation adopts the well-accepted view that DPs denote an entity or a generalized quantifier and can serve as an argument, while NPs denote properties (Huang et al. 2009). In line with Huang et al. (2009) and Longobardi (1994), I assume that object-denoting expressions (kind or individual) are DPs. Accordingly, it is assumed in this study that in Mandarin nominals in argument positions are DPs, while nominals in non-argument positions may have less structure than DPs. Similar to Longobardi's analysis of European languages, an empty D exists in those Mandarin argument nominals that do not have an overt determiner, such as the object in (1a). Non-argumental bare nominals are NPs, such as the nominals in the *as* construction, as illustrated in (1b).

- (1) a. Mao chi-diao-le [DP *e* [NP binggan]]
 cat eat-off-PERF cookie
 'The cat ate up the cookie.'
- b. Lili dang [NP laohu], Mingming dang [NP shizi]
 Lili act as tiger, Mingming act as lion
 (Two children are acting.) 'Lili acts as a tiger and Mingming acts as a lion.'

The structural assumption for non-bare nominals in this study is based on two proposals: Huang et al. (2009) and Cinque (2005). Huang et al. (2009) propose that Mandarin demonstratives are D heads, and numerals, classifiers, and nouns project NumP, ClP, and NP, respectively, as presented in (2).

- (2) [DP na [NumP san [CLP ben [NP shu]]]]
 that three CL book
 'those three books'

This is in accord with Cinque (2005) who finds that the structure illustrated in (3) is the universal structure for DPs, which is essentially identical to what is presented in (2), with the addition of A (adjective) in (3). This study adopts Cinque's proposal that A and NP form a constituent and further assumes that A is an adjunct to NP.

- (3) [Dem...[Num...[Cl...[A...NP]]]] (adopted from Cinque 2005:(11))

Based on the proposals that D, Num, CL, and N are heads in Mandarin and A adjuncts to NP, a DP such as *na san-ben jiu shu* 'those three old books' is assumed to have the following structure.

- (4) [DP na [NumP san [CLP ben [NP jiu [NP shu]]]]
 that three CL old book
 'those three old books'

Obviously, the Mandarin nominal structure can be much more complicated than the structure in (4). Other elements that can appear in noun phrases include quantifiers and relative clauses. Hsieh (2008), in line with Vangsnes (2001), argues that there is a higher projection above DP in Mandarin. For instance, example (5) illustrates that a universal quantifier can be added to a Mandarin DP.

- (5) [QP suoyou [DP na [NumP san [CLP ben [NP jiu [NP shu]]]]]
 all that three CL old book
 'all those three old books'

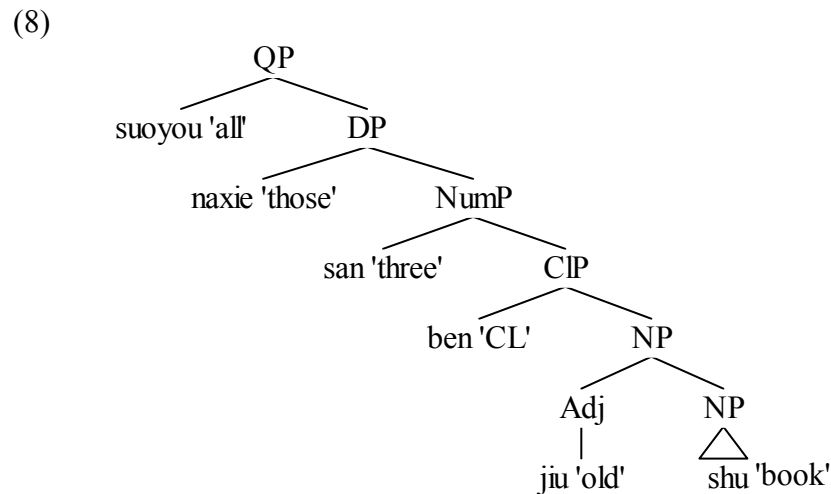
However, the universal quantifier can also follow the demonstrative, as shown in (6). The difference between *suoyou na* 'all those' and *na suoyou* 'those all' in Mandarin is subtle and very hard to distinguish.

- (6) [DP na [QP suoyou [Numsan [CLP ben [NP jiu [NP shu]]]]]]
 that all three CL old book
 'all those three old books'

Cinque (2005:318) also notes that universal quantifiers may follow demonstratives, as is found in Korean. Since there is not a consensus of the QP structure in Mandarin, for the purpose of the current study, which focuses on the contrast between bare nominals and non-bare nominals (mainly DPs), the universal nominal structure proposed by Cinque is adopted, as shown in (7) (adopted from Cinque 2005:(11)).

- (7) [Quniversal [Dem [RC [Num [Cl [A NP]]]]]]

Following the structure in (7), a possible representation for the complex Mandarin nominal in (5) is illustrated below.³



³ In Cinque (2005), it is not clear whether RC is a head or an adjunct. According to the bracketing shown in (7), RC seems to be a head for him. But since adjectives are adjuncts, relative clauses are probably adjuncts for him, too. Note that Cinque also mentions that RCs can appear lower, such as below the numerals (p.328). Example (8) simply shows one possible construction of Mandarin complex nominals. There are only 10 tokens of RCs among the 1969 nominals analyzed in the production study of this dissertation. Therefore, I will assume Cinque's nominal structure and leave the detailed structure of mandarin RCs for further inquiry.

For this dissertation, the relative order of Q and D, either Q+D in (5) or D+Q in (6), is less important. What is more important is the relative order of the elements – D Num CL Adj N – inside the noun phrase, and their possible interpretations.

In brief, the current study assumes that Mandarin referential arguments are DPs in the adult language. Thus, for Mandarin bare nominals, those in argument positions are DPs, and those in non-argument positions have less structure than DPs. For Mandarin non-bare nominals, the basic structure is assumed to be [DP [NumP [CIP [A NP]]]] with adjectives being adjuncts of NPs.

2.2 Interpretation of Mandarin Bare Nominals

Bare nominals have been the subject of many syntactic and semantic studies. In this section, I will focus on the interpretation of Mandarin bare nominal arguments and leave the interpretation of non-arguments to be discussed wherever necessary.

There are different proposals that may account for the various interpretations of Mandarin bare nominals. This section summarizes the studies of Chierchia (1998) and Cheng and Sybesma (1999) and discusses how to obtain the generic, definite, and indefinite interpretations of Mandarin bare arguments from different approaches.

2.2.1 Chierchia (1998)

Since Carlson (1977), many researchers have followed his idea and taken kind-referring as the default meaning of bare nominals in English and some other languages (Krifka 1995, Chierchia 1998, and Yang 2001; but see Chierchia's critique (p.358) on

Krifka). This subsection summarizes part of the study in Chierchia (1998) and focuses on how the generic and definite interpretations emerge for Mandarin bare nominals.

Chierchia's typological research (1998) of bare argument nominals proposes that languages vary in what their nominals denote. He proposes that the denotation of nominals is set by a semantic parameter, the Nominal Mapping Parameter. In his view, nominals have two features: $[\pm \text{arg(ument)}]$ and $[\pm \text{pred(icate)}]$. $[+\text{Arg}]$ nominals can map into arguments directly and get kinds as their value. $[+\text{Pred}]$ can map into predicates directly and get property as their value. According to Chierchia, a typology of the possible nominal settings can be shown as follows.

- | | |
|--|---|
| (9) a. $[+\text{arg}, -\text{pred}]$ e.g. Mandarin | <ul style="list-style-type: none"> - generalized bare arguments - the extension of all nouns is mass - no plural morphology - generalized classifier system |
| b. $[-\text{arg}, +\text{pred}]$ e.g. French,
Spanish | <ul style="list-style-type: none"> - no bare nouns in argument position - mass/count distinction - plural morphology |
| c. $[+\text{arg}, +\text{pred}]$ e.g. English | <ul style="list-style-type: none"> - bare mass nouns can be arguments;
singular count nouns cannot - mass/count distinction
($[+\text{arg}]$ N: mass, $[+\text{pred}]$ N: count) - plurals can be arguments - plural morphology |
| d. $[-\text{arg}, -\text{pred}]$ non-existent | |

According to Chierchia, Mandarin is in the $[+\text{arg}, -\text{pred}]$ category and Mandarin bare nouns denote kinds (p.353) and thus are allowed to occur freely as arguments (p.353, 401). English is a $[+\text{arg}, +\text{pred}]$ language. Bare mass nouns and count plurals can be arguments and kind-denoting in English. However, this does not mean bare nominals always have the same interpretation in Mandarin and English. Mandarin bare nominal arguments can be interpreted as generic or definite in subject position, as shown in (10),

(and as generic, definite, or indefinite in object position). English bare subjects are also ambiguous and can receive a generic or an indefinite interpretation, as shown in (11).

- (10) **diaonao** neng kongzhi feiji
 computer can control plane
 'Computers can control planes.' (generic)
 'The computer/s can control planes.' (definite)

- (11) **Computers** route modern planes. (generic or indefinite)

Although the generic reading of the bare nominal subjects in (10) and (11) can be accounted for by assuming that bare nominals are kind-referring by default, an explanation is needed for the differences in interpretation between English and Mandarin.

Assuming Partee's (1987) type-shifting devices, Chierchia proposes that the Blocking Principle and type-shifting operations can account for the bare nominal's capability of obtaining various readings. His approach can also account for the Mandarin data. The Blocking Principle states that if in a language there is a determiner whose meaning is equivalent to a particular type-shifting operation, then the use of that covert type-shifting operation as an automatic type-changing functor is blocked (p.360). In languages like Mandarin, which do not have a definite or an indefinite article to block automatic type-shifting, bare nominals are able to obtain either definite or indefinite readings. Chierchia assumes that for a language that does not have definite articles the non-overt ι operator is universally available. The ι operator is a type-shifter and has the semantics of a definite article. For a language that has definite articles, such as English, ι is not available based on the Blocking Principle (p.360). Since Mandarin does not have definite articles, it will automatically resort to ι for definite interpretation of bare nominals. Therefore, in order to generate the intended definite interpretation of the

subject in (10) (repeated here as (12a)), which is determined by the context, the correct structure should be (12b).

- (12) a. **diannao** neng kongzhi feiji
computer can control plane
(definite) (generic)
- 'The computers can control planes.' definite
*'Some computers can control planes.' indefinite
- b. [_{IP} t_i **diannao** [_{IP} feiji]_i [_{ASPP} Gn [_{MODP} neng [_{VP} kongzhi t_j]]]]
(def.) computer plane can control
- 'The computers can control planes.'

In (12), the object *feiji* 'plane' receives the generic reading by being raised outside of the VP (a scope-shifting operation available at LF), while the subject *diannao* 'computer' obtains the definite reading by resorting to *ι* for the definite interpretation.

As for the indefinite reading of Mandarin bare nominals, we have to first consider how the indefinite reading is obtained in English. To derive the indefinite reading for bare nominals, such as the English subject in (11), Chierchia adopts the scope-shifting operation at LF and Diesing's (1992) generalization that bare nominals are interpreted as weak existential (indefinite) within the VP and as generic outside the VP when there is a Gn operator in the clause (pp. 366-368), as shown in (13).

- (13)(adopted from Chierchia 1998:(39))
- a. **Computers** route modern planes.
 (indefinite) (generic)
- Intended meaning: 'Some computers route modern planes.'
- b. [_{NP} modern planes]_i [_{ASpp} Gn [_{VP} **computers** route t_i]]

The subject in (13), *computers*, is interpreted as indefinite after being reconstructed in a VP internal position, and the object, *modern planes*, obtains the generic reading after

being raised outside the VP. Although as shown in (12), Mandarin bare nominal subjects do not have indefinite interpretation, how the indefinite reading is derived in English can be applied in Mandarin to derive the indefinite interpretation of a bare nominal, as demonstrated in (14).

- (14) a. *xuesheng* *yao* *mai* **shu**
 student want buy book
 (definite) (indefinite)

'The student wants to buy (some) books.'

- b. [_{IP} ɿ xuesheng [_{VP} yao mai **shu**]]
(def.) student want buy book

'The student wants to buy (some) books.'

A bare nominal argument in Mandarin, such as *shu* 'book' in (14), can be interpreted as indefinite when it appears within the VP. As for the lack of indefinite readings in subject position in Mandarin,⁴ one could argue that they cannot reconstruct to a position inside the VP.

In short, Chierchia's approach can account for the various interpretations of Mandarin bare nominal arguments. He argues that Mandarin bare nouns (NP) are arguments and kind-referring by default. The existential readings (definite and indefinite) can be explained by the Blocking Principle, the type-shifting operation, and the scope-shifting operation at LF.

However, there seems to be a problem to adopt Huang et al.'s (2009) syntactic assumption and Chierchia's proposal at the same time. Huang et al. assume that Mandarin nouns are predicates (NP) and Mandarin arguments will project as DPs. For

⁴ Chapter 4 will discuss the distribution of various types of Mandarin nominals. For the prohibition of indefinite nominals in Mandarin subject position, see Tsai (2001).

Chierchia, Mandarin nouns (NP) are inherently argumental and there is no need to project D for arguments. The study which will be discussed immediately (Cheng and Sybesma 1999) also argues that Mandarin bare nouns are predicates. I will try to resolve this issue at the end of the chapter, Section 2.3.

2.2.2 Cheng and Sybesma (1999)

Another approach that accounts for the different interpretations of Mandarin nominals is Cheng and Sybesma's study (1999). Cheng and Sybesma propose that the various interpretations of nominals in Mandarin and Cantonese result from different syntactic structures. In this section, I summarize their proposal and focus on how different interpretations of Mandarin nominals are derived within their proposal.

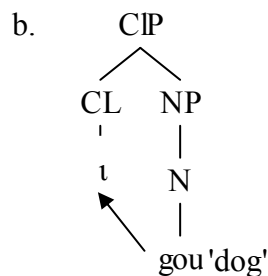
Cheng and Sybesma assume that it is a property of Universal Grammar that some expressions describe, whereas other expressions perform the deictic function of linking the description to a particular object or event in the real world. In the verbal domain, VP describes and T(ense) refers. In the nominal domain, NP describes and D refers. For articulated languages, the articles carry out the deictic function. In Mandarin, where there are no articles, Cheng and Sybesma argue that the classifier takes up some of the functions of D (p.518).

What motivates Cheng and Sybesma (1999: 520, 535) to propose that the classifier is similar to the determiner is as follows. Both CL and D have a singularizing function, and they both are type-shifters, changing predicates to arguments. Besides, similar to the classifier system, the determiner system in some languages is also involved in classification. Determiners are often encoded with gender marking (e.g. in German and Spanish.). Gender marking is a way of classifying nouns into masculine, feminine

and neuter. In other words, the classifiers in Mandarin and gender marking in articulated languages both classify nouns, although the criteria used for the classification are different.

Based on the argument that Mandarin CL are like D in articulated languages, they propose that bare nouns in Mandarin are never bare – there is always a Classifier Phrase in the structure for bare nouns. For Mandarin bare nouns to have definite reading, the CL head is filled with an ι operator, followed by the N-to-CL movement. Because the ι operator changes the nominal from a predicate to an individual, the nominal cannot stay in an NP any more, and thus must raise from N to CL (p.522), as illustrated in (15).⁵

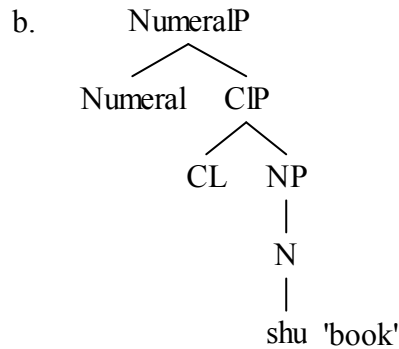
- (15) a. **gou** yao guo malu (= Cheng and Sybesma (2a))
 dog want cross road
 ‘The dog wants to cross the road.’



For indefinite noun phrases in Mandarin, Cheng and Sybesma propose that there is a Numeral Phrase on top of the Classifier Phrase. The classifier and the numeral can be overt or covert. Since Numeral Phrases are inherently indefinite, the noun phrase will have an indefinite interpretation. The structure of an indefinite bare nominal is illustrated in (16).

⁵ Cheng and Sybesma do not discuss predicative bare Mandarin nominals. However, it seems they assume predicative bare nominals are NPs according to: (i) their proposal that definite bare nominals are CIPs and indefinite bare nominals are NumPs; (ii) their statement that nominals that denote individuals, instead of predicates, cannot stay in NP and must raise to CL. In other words, predicative bare nominals are NPs and argument bare nominals are at least CIPs.

- (16) a. Hufei mai **shu** qu le (= Cheng and Sybesma (1a))
 Hufei buy book go SFP
 'Hufei went to buy a book/books.'



As for the generic reading, for Cheng and Sybesma, bare nominals with definite or generic interpretation have the same structure (CLP) and must undergo the N-to-CL movement. The parallel between definiteness and generics is reasonable in that the definite operator/article selects the maximal set, while the generic reading can also be argued to select the maximal set. They state that the ability of Mandarin bare nominals to be interpreted as generic or definite is similar to that of English definite singular nominals, which can also receive either the generic or the definite reading, as shown in (17).

- (17) a. **The tiger** lives in the jungle. (generic or definite)
 b. **laohu** zhu zai conglin li (generic or definite)
 tiger live in jungle in
 'Tigers live in jungles.'
 'The tiger lives in jungles/the jungle.'

As for how the kind interpretation arises for Mandarin, they adopt Chierchia (1998) and assume the application of the \cap 'down' operator, which is applied in [+pred] nominals, such as English count plurals, to shift property to kind. In English, \cap dogs will refer to dog-kind. They claim that since Mandarin nouns are neutralized in number (bare nominals can have plural or singular readings), the 'down' function can always apply (p. 533). However, by applying the \cap 'down' function on Mandarin nominals, Cheng and

Sybesma's proposal is clearly distinct from Chierchia's view for Mandarin in that they assume Mandarin nominals are predicates, while Chierchia assumes that they are arguments.

In Cheng and Sybesma's framework, the definite and indefinite readings of Mandarin bare nominals can be accounted for by different syntactic structures – definite nominals are CIP and indefinite nominals are NumP (CL and Numeral can be overt or covert). Mandarin kind-referring bare nominals also have the CIP structure, same as the bare nominals with definite reading. The kind interpretation is due to the fact that Mandarin nouns are predicates and neutralized in number, which enable the 'down' operator to be applied freely.

2.3 Summary and Discussion

It is assumed in this study that in Mandarin the demonstrative is one of the possible heads of DP, numerals are heads of NumP, classifiers are heads of CIP, and nouns are heads of NP. Adjectives adjunct to NP and Adj-NP will form a constituent, a larger NP. Mandarin referential nominal arguments are DPs, while nominals in non-argument positions may have less structure than DPs. The production study of this dissertation will investigate Mandarin nominals in free speech data to examine how the structure of nominals is developed in child language.

Although syntactically, following Huang et al. (2009) I assume that Mandarin object-denoting nominals are DPs, their study is unable to provide the semantic insight needed for this chapter. Two other theories, Chierchia (1998) and Cheng and Sybesma (1999), were discussed. My intention is to illustrate that, based on these two approaches,

it is possible to account for the various interpretations of Mandarin bare nominals without completely endorsing either of the approaches. The interpretation study of this dissertation will inspect adults' and children's interpretations of nominals to examine what the preferred interpretation is and why it is preferred.

A closer look reveals that these three studies make quite different assumptions about Mandarin. For Huang et al. (in line with Longobardi 1994), Mandarin nominals are predicates and arguments are DPs. Chierchia argues that Mandarin nominals (NP) are arguments and kind-referring by default. In Cheng and Sybesma, Mandarin nominals are predicates and Mandarin CL, like English D, makes referring expressions out of predicates. Therefore, the CL head is always needed for Mandarin arguments, even for bare nominal arguments. For them, kind-referring bare nominals are CIP. These hypotheses about Mandarin nominals are summarized below.

Table 2.1 Mandarin nominals

	Huang et al.	Chierchia	Cheng and Sybesma
Noun	predicate	argument	predicate
Kind-referring Bare Argument	DP	NP	CIP

There is a problem to adopt Huang et al.'s (2009) syntactic assumption and Chierchia's (1998) proposal at the same time. For Huang et al. (following Longobardi 1994), Mandarin referential arguments are DPs. For Chierchia, Mandarin nouns (NP) are arguments and it is not necessary to project D.

Longobardi in his recent study (2008) argues that the principle that arguments must be DPs does hold in Germanic and Romance languages, but certain other languages, such as Japanese, 'are likely to exhibit nominal arguments without D, as expected in

Chierchia's (1998) framework' (p.189). Therefore, he proposes a new analysis of the D head and argues that this analysis will prevail cross-linguistically, as stated below.

(18) a. D is the Person head.

b. Individuals are denoted through the Person feature.

He argues that D is the Person head, which consists of the person features, such as gender and number. Individuals are denoted through the Person feature in D; thus referential arguments are DPs. For languages with grammaticalized person features, such as Romance and Germanic languages, (18b) applies in narrow syntax and arguments are associated with D overtly or covertly. He notes that languages without grammaticalized person features, such as Japanese, 'may indeed have bare (i.e. without D) NPs in argument position' (p.207), as suggested in Chierchia (1998). He argues that these languages 'may associate person to the relevant expressions freely (or at most under pragmatic constraints)' and (18b) 'applies in conceptual or pragmatic representations' (p.207). In a word, the statements in (18) hold for both types of languages. Mandarin is a language without grammaticalized person features. Based on Longobardi's new framework, although syntactically Mandarin allows bare NPs in argument position, Mandarin arguments are actually DPs conceptually or pragmatically.

Following Huang et al. (2008) and Longobardi (1994), I assume that nouns are predicates across languages and thus Mandarin referential arguments are DPs. To adopt Chierchia (1998) and Longobardi (2008), I assume that Mandarin allows bare NPs as arguments with the type-shifting operation, which creates expressions denoting kinds, definites, etc., located in D (the Person Head). Nevertheless, whether bare nominal arguments are NP or DP will not affect the discussion of the data in the study. In the production study, given that the data were produced by young children, only the phrasal

projections with a head filled by a lexical word will be discussed. More linguistic background concerning Mandarin nominal structure will be discussed in the chapter of the production study. In the interpretation study, how Chierchia accounts for the generic interpretation in different languages will be discussed. His proposal that Mandarin bare nominals are kind-referring by default will play an important role in making research predictions and in understanding the data.

CHAPTER THREE

ACQUISITION ASSUMPTIONS

Generative theories (Chomsky 1965, 1981, 1986, 1995) assume a common system innate to humans: Universal Grammar (UG), which consists of universal principles and parameters of possible variation in languages. This innate linguistic system equips children with the capability of acquiring any language in their environment. The presence of such a system can account for children's ability to determine the entire rule system of a language and to generate grammatical sentences that they have never heard before, with the limited input they encounter (the poverty of the stimulus) and the lack of negative evidence in their environment.^{6,7} Researchers argue that these facts are only explainable by UG, the predisposition of humans to acquire language.

The current study adopts these assumptions and investigates the production and interpretation of bare and non-bare nominals in child Mandarin. The main research questions focus on (i) how do Mandarin-speaking children proceed from bare to non-bare nominals and what are the contexts that trigger the emergence of non-bare nominals? (ii) how do they interpret nominals at different ages (in different developmental phases)?

Since child language acquisition studies focusing on the contrast between bare and non-bare nominals are few, this chapter will discuss two previous works that will

⁶ See Laurence and Margolis (2001) and Pullum and Scholz (2002) for a detailed review of the poverty of the stimulus argument.

⁷ Negative evidence is the information about what is not grammatical in the target language. Negative evidence is not available because caregivers usually correct children for truthfulness, not grammaticality (Cairns 1996). For the unavailability of negative evidence, see Marcus (1993).

inform the hypotheses in the current study: Chierchia, Guasti, and Gualmini (1999) and Roeper (2006).

3.1 Defaults and The Subset Principle

The acquisition studies that will be discussed in this chapter, Chierchia et al. (1999) and Roeper (2006), both assume the existence of defaults in the acquisition of nominals. The theory of Chierchia et al. is built on the assumption of defaults and the Subset principle. This section will briefly discuss the idea of defaults and the Subset Principle.

To determine what the initial state is and how children move from the initial state to the target grammar are two of the main goals of child language acquisition research. First, what is the initial state? Chomsky (1988:61) claims the initial state of the language faculty before any experience is made up by UG. This initial state includes universal principles and parameters of possible variation. All languages share certain universal principles, e.g. all have nouns and verbs. Languages differ in the settings of some parameters. For example, English is an articulated-language, while Mandarin is article-less. A question immediately arises: how are parameters set? It is possible that parameters may have a default setting which is subject to possible resetting. According to Fodor (1998), defaults enable the learners to parse the input and cope with ambiguity that natural language displays. An important task of language acquisition is to reset the settings of some parameters to conform to the target grammar.

Second, how do children move from the initial state to the target grammar? Any plausible theory to answer this question must deal with the learnability problem mentioned above – the unavailability of negative evidence. It is a challenge to formulate

an account of how children proceed from one stage to another with the assumption that only positive evidence is available to them. In view of that, a language learning mechanism that avoids negative evidence is proposed, i.e. the Subset Principle. The Subset Principle is briefly discussed as follows.

Assuming the existence of UG and the default settings, the following scenario is presented in order to set the stage for the Subset Principle. Suppose that Language 0 is a subset of Language 1. Setting the value of a parameter at 0 will yield Language 0. The same goes with Language 1 – setting the value of the same parameter at 1 will yield Language 1. If children start to acquire a language by setting the parameter value at 1 (the superset language), given the absence of negative evidence, it is then not possible for them to falsify the parameter of value 1 and reset the parameter value to 0 (the subset language). To solve this problem, the Subset Principle (Berwick 1985, Wexler and Manzini 1987 among others) suggests that language learners should start out with settings that yield the smallest (subset) language. The settings in a subset language are more falsifiable; therefore, when encountering constructions that are incompatible with such settings, language learners are able to reconsider the hypothesis based on the positive evidence. For that reason, the default setting of nominals should be most falsifiable, which readily allows positive evidence to trigger the change to another setting.

With the assumption of defaults and the Subset Principle in mind, we turn to the two studies that will inform the hypotheses in the current research.

3.2 Chierchia, Guasti, and Gualmini (1999)

The production of bare noun roots is found to be a universal phenomenon in early child speech regardless of their acceptability in the adult target language. A possible account for this is that bare nominals functioning as arguments are the default setting of UG when children begin acquiring the nominal system of their target language (Chierchia 1998, Chierchia et al. 1999). That is why young children produce bare nominals even in the positions that disallow bare nominals in their target languages. This section briefly summarizes the child language study of Chierchia et al. (1999).

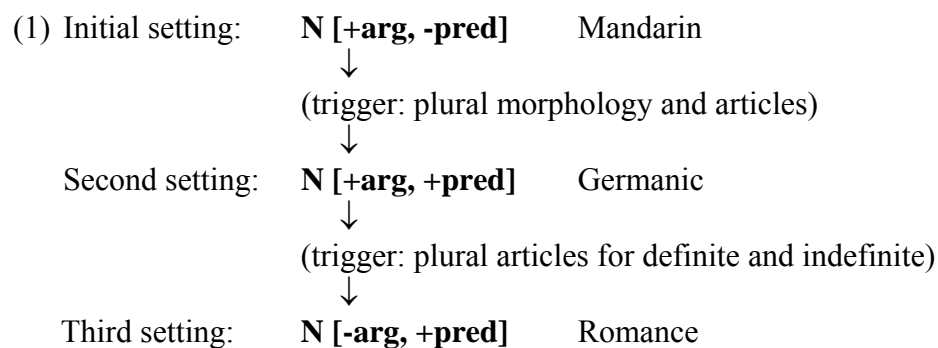
They depart from Chierchia's (1998) Nominal Mapping Parameter, which categorizes languages according to the semantics of their nominals. According to him, there are three different types of nominals: the Mandarin type [+arg, -pred], the English type [+arg, +pred], and the Romance type [-arg, +pred], as presented in Chapter 2. In line with the Subset Principle, Chierchia et al. (1999:28-29) suggest that children start to learn the nominal system in their language by using the Mandarin setting of nominals, [+arg, -pred], allowing their nominal arguments, count or mass, to appear totally bare (mass-like) because this setting is most falsifiable. They argue that the Mandarin type of nominals is the default setting for nominals in the UG. Accordingly, regardless of the rules set forth in their target language grammar, children use bare nominals in all syntactic positions – as permitted by the Mandarin grammar. This is as if they use their language under the Mandarin grammar.

For that matter, English-speaking children start with the Mandarin setting when acquiring nouns – they start to acquire nominals by treating every noun as mass, which is able to appear as a bare nominal argument. According to Chierchia et al., when children discover obligatory articles and plural morphology for count nouns in argument position,

they learn that English bare arguments may be mass nouns or count plurals, but not count singulars, i.e. bare arguments are allowed but restricted by lexical properties. Thus, they will be triggered to reset the default Mandarin setting to the English setting [+arg, +pred], allowing two types of bare arguments (mass and count plurals).

Similarly, Romance language-speaking children start out by allowing every noun to appear in its bare form without any articles (the Mandarin setting). Take French-speaking children as an example, they start to acquire nominals by treating every noun as Mandarin common nouns – being able to appear as bare nominal arguments. Then they learn the existence of articles and plural morphology and reset to the English setting. The English setting allows French bare nouns (mass or count plurals) to appear as arguments without projecting D. This setting cannot be supported by positive evidence from the target language, French, in which D should always appear. Eventually, they acquire the full French article system and discover that D must always project, resulting in the French [-arg, +pred] setting.

The discussion above shows the acquisition path of English and French nominals based on the assumption that the Mandarin setting is the default setting of nominals in UG. The three settings of nominal parameter mappings of Chierchia et al. are summarized below (adopted from Chierchia et al. 1999: (26)).



Chierchia et al. argue that the child speech data they examined support this proposal. They examined the pattern of determiner omission in four languages – French, Italian, English, and Swedish – by analyzing the spontaneous speech data of 16 children (mostly two-year-olds). They find three phases of determiner development: total omission of determiners, optional use of determiners, and convergence to the target language. Their main finding is that beyond the phase of total omission of determiners, all children optionally omit articles for a varied period of time. French and Italian children cease to omit articles at an earlier MLU point than that of English and Swedish children. In the French and Italian data, the optional determiner phase is shorter and ends abruptly. In the English and Swedish data, this phase lasts much longer and converges toward the target gradually.

The initial setting, the Mandarin-type setting shown in (1), can account for the first phase, in which bare nouns appear without any determiners. The second setting, the English-type setting, enables children to choose whether a noun starts out as an argument or predicate. Depending on their choice, a determiner may be used or not. Therefore, in this phase, nouns appear optionally with a determiner – children produce both '*I want truck*' and '*I want the truck*' in this phase. Because learners of a Germanic language know that nouns can be either arguments or predicates but are not sure yet which noun is which, they stay in the phase of optional determiners longer than Romance language learners. For Romance language learners, as soon as they discover that determiners are generally obligatory for nominals, they end the second phase and quickly move on to their target setting. The proposal summarized in (1) can account for the three phases found in their data.

In short, according to Chierchia et al. (1999), Mandarin nominals are of type [+arg, -pred], which is the default setting for nominals in UG. All children start with this setting when acquiring nominals; therefore, they are found, cross-linguistically, producing bare nominals, even in languages that prohibit bare nominals. Assuming Chierchia et al., it is expected that in the production study of this dissertation, bare nominals will be found to have a freer distribution than non-bare nominals since the Mandarin setting of nominals is the default setting in UG. In the comprehension study, it is expected that children will prefer that bare nominals have a generic reading instead of other readings – even though the definite reading of the bare nominals is acceptable in some given context – because Mandarin nominals refer to kinds by default based on Chierchia (1998) (in line with Carlson (1977)).

3.3 Roeper (2006)

Roeper (2006, cf. 1999) also assumes the crucial role of defaults in language acquisition. Roeper argues that there is clear and strong evidence for the following generalizations ((2) is adopted from Roeper's (59a,b,c)).

- (2) a. Children move from less specific to more specific.
- b. Such path proceeds from N to NP and then to DP.
- c. Defaults play a prominent role in the grammar aspect of language acquisition.

More details about these generalizations are discussed immediately. He argues that the acquisition path starts with Minimal Default Grammar, which is the Initial State projection from UG and contains nodes that are universal in a hierarchical relation, such as the basic form of CP, IP, VP, and NP. According to him, there are three defaults of

nominal projections (p.38), as shown in (3): (i) nonspecific bare Ns (e.g. *cookie*), (ii) relational NPs (NP has a Spec position and the Spec has an anaphoric link allowing the N to have a predicative reading, e.g. *home*), and (iii) particular types of DPs, namely, proper names (e.g. *Johnny*) and deictic pronouns (e.g. *that*). Roeper calls the second category ‘predicate NPs.’ This is somewhat idiosyncratic and different from what is called ‘predicate’ in the current study. To avoid confusion, ‘relational noun’ is used here.

(3) N = Kind	(want <i>cookie</i>)
NP= Relational noun	(Joshua <i>home</i>) ⁸
DP = Proper name/Deictic pronoun	(<i>Johnny, that</i>)

Roeper's proposal, which assumes that the default interpretation of N is kind, echoes with the theory of kind reference (Carlson 1977 and Chierchia 1998 among others). He also notes that children begin with these defaults, all of which do not have an article. This is consistent with the well known observation (Chierchia et al. 1999 among others) that cross-linguistically children start their nominal acquisition journey by producing bare nominals without articles.

⁸ Fillmore (1992) argues with corpus evidence that the English word *home* has special grammatical and semantic properties. His discussion about *home* undoubtedly covers more properties than what is concerned by Roeper. Roeper's proposal focuses on children's distinction between N and NP, which is supported by corpus data and an experimental study by Blumenfeld (1999), as exemplified below. (Examples (i) and (ii) are Roeper's (41) and used in Blumenfeld's experiment.)

- (i) He gave her home-made cookies.
(*home* as a N and can be interpreted as anybody's home)
- (ii) He gave her cookies made at home.
(*home* as a controlled NP with an anaphoric reading)

With regard to the properties of *home* that Fillmore discusses, he also notes the existence of the anaphoric reading in the meaning of *home*, which is consistent with Roeper's proposal.

Roeper (2006:40) also argues that the hierarchical order of the nodes within a nominal phrase is fixed (hypothetically) in UG, but not all grammars instantiate every node in the hierarchy – language learners have to decide which nodes in the hierarchy should be instantiated by their target language. Based on various sources of evidence, he attests that in a tree structure of nominals, children will move from less specific (e.g. NP) to more specific (e.g. DP) and roughly from lower to higher nodes (p.46).

In line with Roeper, one strategy to build up a nominal structure is to instantiate a head in the hierarchy of functional categories provided by UG and add elements under it. This may account for the observation that Mandarin-speaking children overuse the general classifier *ge* with nouns that require specific classifiers as if they are using it as a syntactic place holder of the Classifier head (Hu 1993).⁹ The following example illustrates that after the child successfully instantiated the CL head for Mandarin nominals, she still did not use the semantically correct classifier – even though the adult demonstrated using the correct classifier in the context.

(4) Adult: you ji-**zhi** xiongxiong?
 exist how:many-CL bear
 'How many bears are there?'

Nana (2;2): you san-**ge** xiongxiong.
 exist three-CL bear
 'There are three bears.'

In addition to the generalizations summarized in (2), another crucial argument of Roeper's study is that 'the concept of compositionality is natural and should play a central role' (= Roeper's (59d)). Since every node in the structure requires semantic information to justify its projection, he adopts a Fregean approach to compositionality (as is

⁹ Although Roeper does not explicitly discuss the position of classifiers in his study, classifier phrase is noteworthy when studying the acquisition of Mandarin nominals.

explicated in Heim and Kratzer (1998)), 'where higher nodes are strictly composed of the contents of lower ones' (p.39). He finds that one of the observations that can help explain how compositionality works in child language is that new morphemes at first take no function. For instance, English *-ed* is first added to verbs without clear past reference. Additionally, 20% of the first use of the definite article by a child acquiring English do not have a clear antecedent (Matthewson and Schaeffer 2000, cited in Roeper 2006) (Adult data are not available). This suggests that *the* may initially function as an expletive without any meaning (2006:44-45). It seems that children may project a structure without actually knowing what it means. Accordingly, for the comprehension study of this dissertation, it is expected that Mandarin-speaking children will first use demonstratives as expletives, and demonstrative+noun sequences will allow generic readings, given that N refers to kind by default, as proposed by Roeper's Minimal Default Grammar. At the same time, demonstrative must also allow regular definite (existential) readings since D is a default projection and can also be referential by default.

Another observation as it relates to compositionality in language acquisition is that in a child's initial analysis, s/he first identifies a Head and adds something to it. This phenomenon can be called 'merge to head.' An example of 'merge to head' is to add a modifier to the head, as illustrated in (5).

(5) Merge to Head: X + Head

doll + house = doll house
 modifier + head

Roeper points out that although example (5) seems common sense, there are other accounts that may be used to oppose this compositional analysis. However, evidence supporting these accounts are rarely seen in child language. The first account proposes

that children may treat any sequence as a single word without analysis. Accordingly, many erroneous compounds should arise. For example, when a parent hands a child a spoon and says 'eat rice,' the child may understand *eatrice* as 'spoon' and says 'where my eatrice? ' Few such errors are found. Another account is that children may pursue conjunction for word combinations, and then words like 'Mommydaddy,' 'spoonfork,' or 'hatcoat' will be expected. Yet, these cases are very rare in children's utterances. Besides, this account will predict that children treat 'dollhouse' as referring to two different objects – doll and house – which contradicts the reality. Therefore, Roeper argues that children's analysis identifies a Head and adds something to it (Roeper 2006:41-42).

In sum, the acquisition path of the nominal structure in Roeper's proposal goes from non-specific to specific nodes (roughly from bottom to top) in a series of compositional steps to build the full determiner phrase of the target language. For Roeper, the existence of the Minimal Default Grammar and the application of compositionality are essential in children's acquisition of nominal structure. The notion of Default Grammar is assumed in the current study, while the use of compositionality will not be discussed.

3.4 Summary

This chapter gives a brief overview of the acquisition assumptions adopted in the current study – especially the notions of UG and defaults. Two previous acquisition studies have been discussed in this chapter, Chierchia et al. (1999) and Roeper (2006), to provide a background of the current study. Chierchia et al. and Roeper are introduced to argue for kinds as the default interpretation of NPs and Ns respectively. According to

Chierchia, Mandarin bare arguments are NPs. Roeper's Minimal Default Grammar argues that default bare Ns are kinds, default NP are Ns with a spec, and default DPs are proper names and deictic pronouns. Therefore, there is a difference between these two studies.

As discussed in Chapter 2, I assume that Mandarin allows arguments to be bare NPs and Chierchia's type-shifting operation (e.g. for creating kind or definite referential nominals) is located in D (the Person Head in Longobardi (2008)). Although Roeper assumes the referential arguments are by default DPs, he uses these syntactic category terms in a more semantic way. For example, article-less nominals with predicative, generic, or existential readings are NPs, while articulated nominals referring to discourse familiar or culturally familiar objects or unique objects are DPs, as illustrated below ((6) is adopted from Roeper 2006: (29)).

- | | |
|---|---------------------------------------|
| (6) a. Bare Noun, kind | (want <i>cookie</i>) |
| b. NP, predicate with anaphoric control | (go <i>home</i>) |
| c. NP, generic | (I like <i>cats</i>) |
| d. NP, existential | (<i>cats</i> are in the yard) |
| e. DP, discourse familiar object | (I have a hat. <i>The hat</i> green.) |
| f. DP, unique object focal stress | (did he get <i>THE bike</i>) |
| g. DP, culturally familiar object | (where is <i>the sun</i>) |

Based on the Minimal Default Grammar and (6), NPs are Ns with a spec, which have an anaphoric link allowing the nominal to have a predicate reading (*home*) or generic reading (*cats*), etc. In other words, semantically, Ns are kind-referring by default and NPs can have a generic reading. This does not counter the assumption that nominal arguments are DPs. For example, *cats* in (6c) has a generic interpretation, and as an referential argument it will project as DP (either syntactically or pragmatically, Longobardi 2008).

The current study adopts the following assumptions. Semantically, assuming Chierchia et al., regardless of their target languages, all children begin learning the system of nominals by treating them as Mandarin nouns – which are kind-referring by default. Accordingly, it is more economical for Mandarin-speaking children to interpret nominals as generic than as other readings. Syntactically, assuming Roeper, when children enter the nominal learning process, they already possess the knowledge of Minimal Default Grammar: N, NP, and DP. Following Roeper, it is hypothesized that in the acquisition of nominals, children move from less specific to more specific and the acquisition path mirrors the tree from N to NP and then from NP to DP.

The current study uses empirical and experimental data from Mandarin-speaking children to investigate the production and comprehension of Mandarin noun phrases. More details about the acquisition assumptions will be discussed in the relevant chapters.

CHAPTER FOUR

PRODUCTION OF BARE AND NON-BARE NOUN PHRASES

4.1 Introduction

In some languages, only certain types of nominals (mass nouns in English, for example) can appear completely bare, and only in certain syntactic positions (as objects but not subjects in Spanish, for example) (Alexiadou et al. 2007, Munn and Schmitt 1999). However, in Mandarin, bare nominals are not only widely acceptable, but also able to convey various interpretations (generic, definite, and indefinite), and they can appear in various syntactic positions (subject, object, and predicate), as discussed in Chapter 2.

In language acquisition, the production of bare nominals is found to be a universal phenomenon in early child speech, regardless of the acceptability of bare nominals in the adult target language. As a result, we hear young children say *I want cookie* (Naomi 1;11.11, Sachs corpora), instead of *I want a cookie* or *I want cookies*. It is not surprising that articles or inflections tend to be omitted in child language, just like other functional categories (Brown 1973, Radford 1990, Rizzi 1993/4, 1994, Wexler 1994, Guasti 2000, and Chierchia, Guasti, and Gualmini 1999, to name a few).

When acquiring languages that generally do not allow bare nominals, children will proceed from the bare nominal stage to the non-bare nominal stage producing grammatical nominal forms with articles or inflections. However, the occurrences of bare nominals in child Mandarin do not necessarily result from determiner omission, as found in some other child languages, because bare nominals are grammatically legitimate

in Mandarin. Therefore, the first challenge that Mandarin poses to the acquisition of nominals is the following.

Challenge 1. Since bare nominals can convey almost all kinds of interpretation in an appropriate context, what are the contexts that trigger the emergence of non-bare nominals? In other words, what are the factors that determine whether or not to use non-bare nominals?

This research tries to answer these questions by studying two longitudinal corpora of child Mandarin. This chapter is divided as follows: 4.2 presents some syntactic and semantic characteristics of Mandarin bare and non-bare nominals. 4.3 discusses the research hypotheses and predictions. 4.4 presents the methodology: how data were collected, the software used to analyze data, and how research variables were defined and coded. 4.5 presents the study results and discussion, and 4.6 the general discussion and conclusion.

4.2 Linguistic Background

To answer how children acquire bare and non-bare nominals in Mandarin, we have to first understand some linguistic properties of Mandarin nominals. In this study, all nominals are divided into two groups: bare and non-bare.

Bare nominals refer to bare noun roots, such as (1a) *shu* ‘book,’ and modified bare nominals, such as adjective-modified nominals as in (1b) *hao shu* ‘good book.’

Non-bare nominals refer to any nominal that includes a bare nominal and something else,

as exemplified in (1c) to (1g).¹⁰ I also assume the following nominal structures for Mandarin.¹¹

- (1) a. [NP shu]
 ‘book’
- b. [NP hao [NP shu]]
 ‘good book’
- c. [DP ta de [NP shu]]
 ‘his book’
- d. [ClP ben [NP shu]]
 ‘CL book’
- e. [NumP san [ClP ben [NP shu]]]
 ‘three CL book’
- f. [DP zhe [ClP ben [NP shu]]]
 ‘this CL book’
- g. [QP mei [NumP san [ClP ben [NP shu]]]]
 ‘every three CL book’

¹⁰ Mandarin is famous for having substantial amount of compounds and there is great deal of disagreement over the definition of compounds in Mandarin (Li and Thompson 1981:45). In the current study, since the subjects are young (two- and three-year-olds), many of their compounds are arguably fixed terms and will be transcribed as one morpheme (see Appendix I for transcription criteria). Therefore, the words that are arguably compounds in adult Mandarin are treated as bare nominals if they are transcribed as one morpheme in the corpora. For example, *dongwu yuan* (animal garden) ‘zoo’ and *yangyu pian* ‘potato chip’ produced by Nana (2;3) are counted as single words and as bare nominals.

¹¹ More complicated structures may be needed when these nominals appearing in different syntactic positions in sentences. Although I assume Longobardi (1994, 2001) and Huang et al. (2009) that Mandarin arguments are DPs, for the data from child language, I will only postulate a structure when there is overt evidence for it, at least with respect to the description of the data.

Examples in (1) show that nouns, possessives, classifiers, numerals, demonstratives (there are no articles in Mandarin), and quantifiers will head their own maximal projections, while adjectives (such as in (1b)) will be adjuncts to NP, which is consistent with Cinque (2005). Although bare noun root (NP) and adjective+noun (NP) are different in their surface forms, they share some common features. For example, both can be kind-referring (e.g. *laohu* ‘tigers,’ and *white laohu* ‘white tigers’). This study categorized nominals to two groups: bare and non-bare, with adjective+noun sequences as bare.

This section will present some linguistic properties of Mandarin nominals to demonstrate the distinct distribution of bare and non-bare nominals. This is to justify the methodology used in this study. In the methodology section, I will present more linguistic details about nominals when defining all possible variables that may influence the appearance of bare and non-bare nominals.

First, all syntactic positions allow bare nominals, but different syntactic positions select different types of non-bare nominals. For example, there is a subject-object asymmetry for the type of nominals allowed in either position. The subject position allows bare nouns, modifier+nouns, demonstrative+classifier+nouns, and quantifier+nouns, but it disallows classifier+nouns, and numeral+classifier+nouns, as illustrated in (2a). On the other hand, the object position allows all of them, as shown in (2b).¹²

¹² In this study of child speech production, a structure will be projected when there is overt evidence for it. Therefore, the abbreviations listed below should be understood literally.

(i) Bare NP: bare nouns

Modified NP: modifier+nouns

- (2) a. **Mao/Bosi mao/*zhi mao/*san-zhi mao/zhe-zhi mao/quanbude mao dou**
 Cat/Persian Cat/CL cat/three-CL cat/this-CL cat/some cat/all cat DOU

chi-le yu
 eat-PERF fish

‘The cat(s)/The Persian cat(s)/*A cat /*Three cats/This cat/All the cats ate fish.’

- b. Zhe-zhi mao chi-le **yu/guantou yu/tiao yu/san-tiao yu/zhe-tiao yu/**
 This-CL cat eat-PERF fish/can fish/CL fish/three-CL fish/this-CL fish

quanbude yu
 all fish

‘This cat ate fish/canned fish/a fish/three fish/this fish/all the fish.’

Among non-bare nominals, demonstrative+classifier+nouns and quantifier+nouns may show up in either the subject or object position, but classifier+nouns with or without a numeral can only be found in the object position.

Second, besides positions, other elements in the sentence also contribute to the distribution of bare and non-bare nominals. Take nominals in the object position as an example. There are restrictions on the verb forms and the type of objects they can co-occur with. The following examples show the first aspect marker acquired by Mandarin-speaking children, the perfective marker *le* (Chang 2002, Erbaugh 1992). A verb appearing with the perfective aspect *-le* can co-occur with both bare and non-bare nominals (bare NP, modified NP, CIP, NumP, DP, and QP) in its object position, but a bare verb without any aspect marker can only allow bare nominals (bare NPs and modified NPs) as its object, as shown below.

CIP: classifier+nouns	NumP: numeral+classifier+nouns
DP: demonstrative+classifier+nouns	QP: quantifier+(classifier)+nouns

For ease of discussion, the abbreviated form will be used in this chapter and the full term will be spelled-out whenever necessary.

(3) Context outside the nominal domain: existence of the perfective marker *le*

- a. Zhe-zhi mao chi-le **yu/guantou yu/tiao yu/san-tiao yu/zhe-tiao yu/**
This-CL cateat-PERF fish/can fish/CL fish/three-CL fish/this-CL fish

quanbude yu
all fish

‘This cat ate (the) fish/canned fish/a fish/three fish/this fish/all the fish.’

- b. zhe-zhi mao chi **yu/guantou yu/*tiao yu/*san-tiao yu/*zhe-tiao yu/**
This-CL cat eat fish/canned fish/CL fish/three-CL fish/this-CL fish

***quanbude yu**
all fish

‘This cat eats fish/canned fish/a fish/three fish/this fish/all the fish.’

Accordingly, a bare verb can only take bare nominals as its complement,¹³ while verbs with an aspect marker can take all kinds of bare and non-bare nominals. When non-bare

¹³ Even if the word *zuotian* ‘yesterday’ is added to force the past reading of (3b), the types of nominals accepted in this sentence do not change, as shown in (i).

- (i) zhe-zhi mao zuotian chi yu/guantou yu/*tiao yu/*san-tiao yu/*zhe-tiao yu/*quanbude yu
this-CL cat yesterday eat fish/canned fish/CL fish/three-CL fish/this-CL fish/all fish
‘This cat ate fish/canned fish/a fish/three fish/this fish/all the fish yesterday.’

Only when making a list will all the types of nominals be able to occur, as shown in (ii), which lists the different behavior of ‘this cat’ and ‘that cat.’

- (ii) zhe-zhi mao chi yu/guantou yu/tiao yu/san-tiao yu/zhe-tiao yu/quanbude yu,
this-CL cat eat fish/canned fish/CL fish/three-CL fish/this-CL fish/all fish

na-zhi mao he tang/guantou tang/wan tang/san-wan tang/na-wan tang/quanbude tang
that-CL cat drink soup/canned soup/CL soup/three-CL soup/that-CL soup/all soup

‘This cat eats fish/canned fish/a fish/three fish/this fish/all the fish.
That cat drinks soup/canned soup/a bowl of soup/three bowls of soup/that bowl of soup/all the soup.’

- b. **Zhangsan de mao** chi yu (*generic, existential)
 Zhangsan DE cat eat fish
 'Zhangsan's cat eats fish.'
- c. **Zhe-zhi mao** chi yu (*generic, existential)
 This-CL cat eat fish
 'This cat eat fish.'
- d. **Quanbude mao** dou chi yu (universal generic, existential)
 all cat DOU eat fish
 'All the cats eat fish.'
- e. You **yi-zhi mao** bu chi yu (*generic, existential)
 Have one-CL cat not eat fish
 'There is a cat that does not eat fish.'
- f. (Ni kan!) **Yang/yi-zhi yang**
 (predicative/property)¹⁶
 (You look) sheep/one-CL sheep - here and now -
 '(Look!) Sheep!/ A sheep'
- g. Zhe shi **gou/ye gou/-zhi gou/yi-zhi gou** (predicative/property)
 this be dog/wild dog/-CL gou/one-CL dog - predicate position -
 'This is dog/wild dog/a dog/a dog.'

In sum, there are at least three characteristics of Mandarin nominal grammar that distinguish bare nominals from non-bare nominals – syntactic positions, aspect markers, and interpretations. As for syntactic positions, all syntactic positions allow bare nouns, while different syntactic positions select different types of non-bare nominals. Subject positions are restricted to bare nominals and nominals with demonstratives or quantifiers,¹⁷ while object positions allow bare and various types of non-bare nominals.

¹⁶ In the current study, some types of nominals appearing in this type of utterances are coded as predicative uses, not as an object of *kan* 'look' based on two observations: there is a pause after *kan* in the data and according to the context, the nominals are used for naming.

¹⁷ Num-CL-Ns cannot appear in the subject or topic positions in Mandarin. However, there is a special type of Num-CL-N phrases that can appear in subject positions, as shown below.

As for verb forms, non-bare nominals in the object position must co-occur with an aspect marker. In other words, aspect markers prompt the use of non-bare nominals in object position. Table 4.1 summarizes the properties of nominals regarding syntactic positions and aspect markers.

Table 4.1 Syntactic positions, verb forms, and nominal types

		Bare NP		Non-Bare NP			
		N	Modifier-N	CL-N	Num-CL-N	Dem-CL-N	Q-N
Subject		✓	✓	✗	✗	✓	✓
Object	Bare verb	✓	✓	✗	✗	✗	✗
	Verb + aspect	✓	✓	✓	✓	✓	✓

When it comes to interpretations, different nominal interpretations require different nominal structures, as summarized in Table 4.2. Bare nominals can have generic, existential, or predicative interpretation, while non-bare nominals cannot have

- (i) liang-ge gongren bu gou
two-CL worker not enough
'Two workers is not enough.'
- (ii) liang-ge gongren ban-bu-dong zhe-zhang zuozi
two-CL worker move-not-move this-CL desk
'Two workers won't be able to move this desk.'

As discussed in Li (1998), these are non-referential 'quantity number expressions' – not referential individual-denoting – and should be treated as exception to the summary in Table 4.1 (also see Huang et al. 2009:289). A piece of evidence that illustrates that these quantity expressions are different from referential individual-denoting nominals is that they do not bind a pronoun, as shown below (for other evidence please see Li 1998). Example (iiia) shows that Num-CL-Ns are not allowed in the subject position and do not bind a pronoun. Example (iiib) shows that *you* 'there is/are' must be present to make the sentence grammatical.

- (iii) a. *San-ge ren_i chibuwan ni gei tamen_i de wu-wan fan
three-CL men eat-not-finish you give them DE five-bowl rice
'Three men cannot finish five bowls of rice you gave to them.'
- b. Yousan-ge ren_i chibuwan ni gei tamen_i de wu-wan fan
have three-CL men eat-not-finish you give them DE five-bowlrice
'There are three men that are unable to finish the five bowls of rice you gave to them.'

generic interpretation, except the nominals with a universal quantifier, which have a universal generic reading.

Table 4.2 Interpretation and nominal types

	Bare NP		Non-Bare NP			
	N	Modifier-N	CI-N	Num-CL-N	Dem-CL-N	Q-N
Generic	✓	✓	✗	✗	✗	✗/✓
Existential	✓	✓	✓	✓	✓	✓
Predicative	✓	✓	✓	✓	✓	✓

In sum, bare nominals can appear in different syntactic positions and convey generic, existential, or predicative interpretations, while non-bare nominals are restricted to certain syntactic positions with certain interpretations.

4.3 Hypotheses and Predictions For Emergence and Interpretation of Non-bare Nominals

In adult Mandarin, the distribution of different nominal types in different syntactic positions with different semantic interpretations is not completely without restrictions. The restrictions, as discussed in the previous section, include (i) syntactic positions (e.g. subject position requires certain nominal types), (ii) context where nominals appear (e.g. bare verbs without aspect markers only allow certain nominal types in object position), and (iii) the interpretation which the nominal conveys (e.g. generic interpretation generally requires bare nominals). Accordingly, the hypothesis is stated below.

(5) Hypothesis:

If children's speech production reflects the target language, according to the distinct property of Mandarin bare and non-bare nominals, it is expected that when acquiring nominals, Mandarin-speaking children will systematically produce nominals with particular structures, in particular constructions, and with particular interpretations.

This leads us back to the main question in this study: what factors will trigger or restrict the use of non-bare nominals in child Mandarin? More questions need to be

answered regarding the particular context in which children use non-bare nominals: with what element (nominal structure), in what position, and with which interpretation, do non-bare nominals first appear in child Mandarin? In this section, predictions are made for these questions.

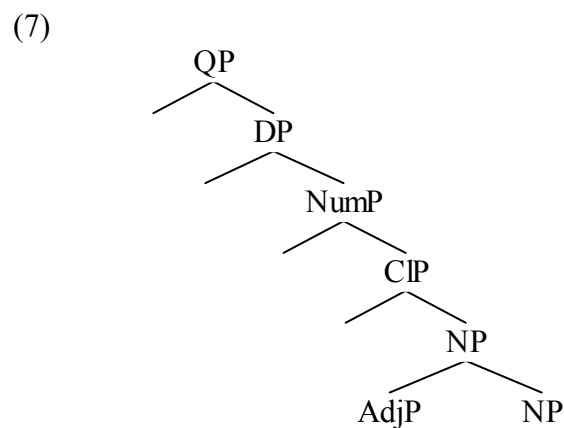
4.3.1 The Element/Structural Projection that First Appears above NP

In child Mandarin, what is the element within the nominal structure that non-bare nominals first appear with? To answer this question, we have to look for the most likely first element to be added to bare nominals to form non-bare nominals in the syntactic structure and in acquisition theory.

In this study, as mentioned in Chapter Two, I follow Cinque's (2005:328) nominal structure, which includes all major projections in nominals, as shown in (6).

(6) [Quniversal [Dem [RC [Num [Cl [A NP]]]]]]

In the spirit of Cinque, I assume the following structure of Mandarin nominals to be examined in the current study. I assume that adjectives will project as NP adjuncts and thus the nominals with adjectives will still project as NP.



Roeper (2006) proposes that children move from nonspecific to specific nodes in a series of compositional steps to build the nominal structure (DP) of their language.¹⁸ He argues that this acquisition path starts with the Minimal Default Grammar, which is the Initial State projection from UG and contains nodes that are universal in a hierarchical relation, such as the basic form of CP, IP, VP, and NP. According to him, there are three defaults of nominal projections (p.38), as shown in (8): (i) nonspecific bare Ns (e.g. *cookie*), (ii) relational NPs (NP has a Spec position and the Spec has an anaphoric link allowing the N to have a predicative reading, e.g. *home*), and (iii) particular types of DPs, namely, proper names (e.g. *Johnny*) and deictic pronouns (e.g. *that*).

- | | |
|----------------------------------|---------------------------------|
| (8) N = Kind | (want <i>cookie</i>) |
| NP= Relational noun | (Joshua <i>home</i>) |
| DP = Proper name/Deictic pronoun | (<i>Johnny</i> , <i>that</i>) |

Given the scope of the current study, I focus on the comparison between bare and non-bare nominals. Since proper names and pronouns (including deictic pronouns) usually appear bare without the bare and non-bare variation, they are excluded in the current study. According to Roeper, the hierarchical order of the nodes within a nominal phrase is ‘(by hypothesis) fixed in UG but not every one is present in every grammar, and so the child needs evidence to determine which ones are present in his grammar’ (p.38). I adopt Roeper’s argument (2006:46) that in a tree structure of nominals, ‘the child moves roughly from lower to higher nodes.’ The generalizations in (9) summarize part of his proposal that is relevant to the current study ((9) is adopted from Roeper's (59a,b,c)).

- (9) a. Children move from less specific to more specific.
- b. Such path proceeds from N to NP and then to DP.

¹⁸ Roeper (2006) has been discussed in Chapter 3. The summary here is to set the theoretical foundation of the research predictions in this chapter.

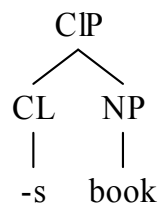
- c. Defaults play a prominent role in the grammar aspect of language acquisition.

In this study, I assume Roeper's acquisition hypothesis and adopt Cinque's (2005) nominal structure as the target structure to be acquired by children. Following Roeper's hypothesis, I predict that the initial elements that children add to a Mandarin bare nominal will appear as low as possible above NP. Since adjectives do not add projections (adjective+NP remains to be NP), adjectives will be the first elements added to bare noun roots to appear within the nominal phrase. Examples of a two-year-old below support the prediction that children use adjectives at a very early age.

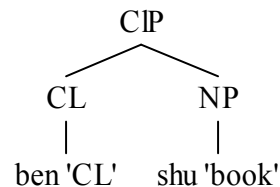
- (10) a. **xiao niao** si-diao (Nana 2;0)
 small bird die-out
 b. **da pengyou** (Nana 2;1)
 big friend

Although Roeper does not explicitly discuss the position of classifiers in his study, classifier phrase is noteworthy when studying the acquisition of Mandarin nominals. Following Borer's framework (2005:93-94), Chang (2007) argues that Mandarin classifiers and English plural *-s* both head the Classifier Phrase and project a phrase that immediately dominates the NP, as shown below.

- (11) a. English CIP



- b. Mandarin CL-N: CIP



Assuming with Borer (2005) and Chang (2007) that classifiers head the first phrasal projection above noun phrases and children acquire nominals roughly from bottom up in the nominal tree (Roeper 2006), I predict that Mandarin classifiers will be the lowest

head found in child Mandarin that is between NP and DP (both are the default nominal projections according to Roeper).

4.3.2 The Argument Position that Non-Bare Nominals First Appear

The research about argument positions in Mandarin acquisition is mostly about the issue of null subjects and null objects in Mandarin (Wang et al. 1992). There is no known research directly studying the internal structure of nominals in particular syntactic positions in Mandarin acquisition.

Mandarin-speaking children follow the SVO order at a very early age. According to Erbaugh (1992:402-405), they enter the strict SVO order stage right after the first word stage (also see Lee 1996). Even when the MLU is as short as between 1.8 and 2.5 and children produce mostly SV or VO, and very few SVO, their utterances already seem to present distinctions between subjects and objects. They also map the appropriate thematic roles to the subject and object positions.

The MLU range for the current study is 1.73 to 3.57, which makes a good period to examine the emergence of non-bare nominals, given that at least three morphemes will be required for SV or VO sequences to include a non-bare nominal. Recall the subject-object asymmetry on nominal forms discussed in (2), repeated below.

- (12) a. **Mao/Bosi mao/*zhi mao/*san-zhi mao/zhe-zhi mao/quanbude mao dou**
Cat/Persian Cat/CL cat/three-CL cat/this-CL cat/some cat/all cat DOU
- chi-le yu
eat-PERF fish
- ‘The cat(s)/The Persian cat(s)/*A cat /*Three cats/This cat/Many cats/All the cats ate fish.’

- b. Zhe-zhi mao chi-le **yu/guantou yu/tiao yu/san-tiao yu/zhe-tiao yu/**
 This-CL cat eat-PERF fish/can fish/CL fish/three-CL fish/this-CL fish

quanbude yu
 all fish

‘This cat ate fish/canned fish/a fish/three fish/this fish/all the fish.’

Subject positions allow bare NPs, modified NPs, DPs, and QPs, but disallow CIPs and NumPs. However, object positions allow all these types of nominals. That is, in the input, children encounter more types of non-bare nominals appearing as objects, as opposed to subjects.

In addition to argument positions, different types of nominals can be found in the predicate position as well. In the predicate position, as presented in (13), different verbs allow different nominal types. The post-copular position (13a) allows bare nouns and classifier+nouns only, and the predicate position after the verb *xiang* ‘look like’ allow all types of nominals, as illustrated in (13b).

- (13)a. Zhangsan shi **laoshi/ge-laoshi/?yi-ge laoshi/*zhe-ge laoshi**
 Zhangsan be teacher/CL-teacher/one-CL teacher/this-CL teacher
 ‘Zhangsan is Teacher/a teacher.’

- b. Xiaoli xiang **gongzhu/ge-gongzhu/yi-ge gongzhu/zhe-ge gongzhu**
 Xiaoli look:like princess/CL-princess/one-CL princess/this-CL princess
 ‘Xiaoli is princess-like.’
 ‘Xiaoli looks like Princess/a princess/one (certain) princess/this princess.’

Generally speaking, predicate positions have more restrictions on the nominal types than object positions do. Therefore, it is possible that non-bare nominals will emerge earlier in object positions than in the predicate positions.

Based on the argument of Erbaugh (1992) that children know the distinction between subject and object and considering that objects allow all types of non-bare nominals but subject and predicate positions have more restrictions on the form of non-

bare nominals, it is predicted that the argument position for non-bare nominals to first appear in a sentence is the object position.

4.3.3 The Interpretation that Non-Bare Nominals Most Frequently Associated with

The studies about the acquisition of nominal semantics in Mandarin mostly focus on the interpretation of pronouns and reflexives (Chien and Wexler 1987a, 1987b, 1990), and quantifier scope (Chien and Wexler 1989, Lee 1986ab, Su 2001). Some focus on genericity (Chang 2008, Gelman and Tardif 1998). There is no known research studying the acquisition of nominal interpretations in Mandarin in general, which is one of the goals of the current study.

Mandarin non-bare nominals can have either existential or predicative interpretation, but no generic interpretation in general. As discussed before, the production of predicative nominals is limited in either the naming context with the characteristics of 'here and now' or in the predicate position. Besides, all types of bare and non-bare nominals can all associate with existential interpretation, while predicative nominals only appear as bare nominals, CLP, or *one*-CL-nouns – other types of non-bare nominals have existential interpretation in general – as exemplified below.

- (14) a. (Nikan!) **Yang/yi-zhi yang** (predicative/property)
 (You look) sheep/one-CL sheep - here and now -
 ‘(Look!) Sheep!/ A sheep’
- b. Zhe shi **gou/ye gou/-zhi gou/yi-zhi gou** (predicative/property)
 this be dog/wild dog/-CL gou/one-CL dog - predicate position -
 ‘This is dog/wild dog/a dog/a dog.’
- c. Wangwang shi **ta de gou/na-zhi gou** (existential)
 Wangwang be he DE dog/that-CL dog
 ‘This is dog/wild dog/a dog/a dog.’

Based on these two observations – (i) non-bare nominals can be interpreted as existential or predicative; (ii) existential interpretation may be conveyed in more types of nominal forms than predicative interpretation – it is predicted that non-bare nominals will be most frequently associated the existential interpretation.

4.3.4 Summary

The research questions and predictions for this study are summarized in the following.

(15) Questions and predictions:

- a. What are the contexts that will trigger or associate with the use (or non-use) of a non-bare nominal?

Factors such as the use of aspect markers will prompt the use of non-bare nominals. Non-bare nominals mostly associated with the existential interpretation. Generic interpretation disallows non-bare nominals.

- b. With which element/structural projection do non-bare nominals first appear beyond the bare nominal stage?

Adjectives, which occupy the adjunct position of NP in the syntactic tree, will be the first elements that children add to a Mandarin bare noun root within the NP structure.

Classifiers will be the first X^0 element appearing beyond the NP projection and below DP.

- c. In which syntactic position do non-bare nominals first appear?

The first syntactic position for non-bare nominals to emerge is the object position, which has less restrictions on non-bare nominal types than the subject and predicate positions do.

- d. With which interpretation do non-bare nominals most frequently associated?

The interpretation that non-bare nominals most frequently associated with will be the existential interpretation.

4.4 Methodology

4.4.1 Data and Subjects

The current study analyzes the longitudinal spontaneous speech data from Mandarin-speaking children from Chang corpora (2002), which were transcribed in the format of Codes for the Human Analysis of Transcripts (CHAT) (MacWhinney 2000). The subjects were recorded in the environmental settings, such as homes or neighborhood playgrounds, where children feel most comfortable to talk freely (Demuth 1996). Natural speech data are argued to be least likely influenced by inappropriate experimental task demands or the required repeated measures of the similar test which affect children's performance. Consequently, natural speech data are particularly advantageous when investigating the early stages of acquisition in young children (Stromswold 1996).

The study analyzes the data of a two-year-old girl, Nana (born in 1998), and a three-year-old boy, Didi (born in 1997). Nana and Didi's data were selected because their mean length of utterance (MLU) falls in the so-called morphosyntactically interesting period – between 1.75 and 4 morphemes (Demuth 1996:4), which is suitable for this developmental study of nominal structure. The data include a two-hour recording in each month for half a year per child, except the first file of Nana, which is a one and a half hour recording. The basic information of the data is provided in Table 4.3 and 4.4.

Table 4.3 Information of Nana's data

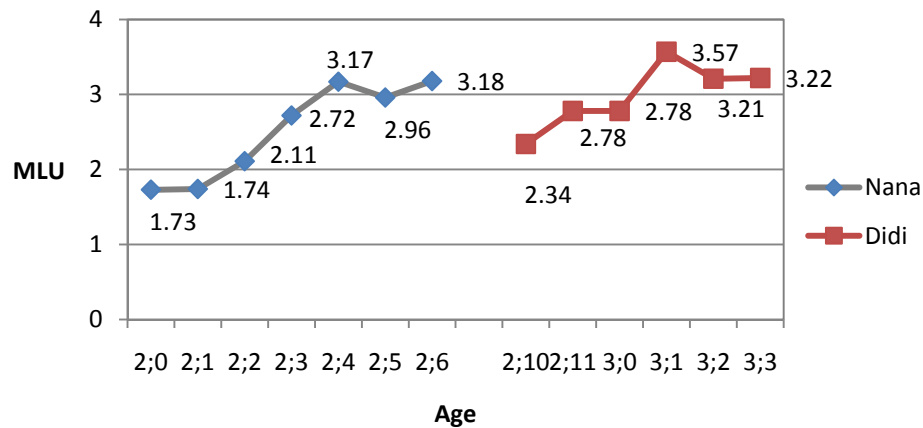
Age	2;0	2;1	2;2	2;3	2;4	2;5	2;6	2;3 (avg.)
Utterances	172	342	709	273	656	591	514	3257 (total)
MLU	1.73	1.74	2.11	2.72	3.17	2.96	3.18	

Table 4.4 Information of Didi's data

Age	2;10	2;11	3;0	3;1	3;2	3;3	3;0 (avg.)
Utterances	672	729	891	544	647	882	4274 (total)
MLU	2.34	2.78	2.78	3.57	3.21	3.22	

The development of MLU across age is also presented in Figure 4.1. Assuming the validity of MLU measurement across languages, including Chinese (Erbaugh 1978, Devescovi et al. 2005), Chang (2002) counted the MLU in terms of morphemes, rather than words or Chinese characters.¹⁹ Chang followed the commonly accepted notion that a morpheme is ‘the smallest meaningful linguistic unit’ (Cairns 1996); as a result, for example, *yi-ge xuesheng* ‘one-CL student’ and *chi-le pingguo* ‘eat-PERF apple’ both have three morphemes. (Please see Appendix I for more details about the calculation of MLUs.)

Figure 4.1 MLU of Nana and Didi



For interpretation of the references to color in this and all other figures, the reader is referred to the electronic version of this thesis dissertation.

4.4.2 The GoldVarb Analysis

This study uses a variable rule analysis tool, GoldVarb 2001 (Robinson, Lawrence, and Tagliamonte 2001) to examine which factors contribute, and how strong they are, to the production of bare and non-bare nominals in Mandarin. GoldVarb is considered to be ‘one of the most appropriate methods available for conducting statistical

¹⁹ Each Chinese character corresponds to one syllable.

analysis on natural speech' (Sankoff 1988:987, cited in Tagliamonte 2006:129). It is a multivariate analysis program using binary logistic regression as a statistical model for variation analysis. An analysis of logistic regression uses quantitative data to make predictions about possible relationships among different variables (Paolillo 2002:153-4).

The reason for not using the popular tool of analysis of variance (ANOVA), which will be used in chapter five to analyze the experimental data, is that ANOVA is primarily used to study a single or a few independent variables. In this study, I will investigate 57 factors, excluding the dependent variable. Moreover, there are a number of advantages in using GoldVarb (Bayley and Young, forthcoming). First, GoldVarb helps perform quantitative modeling. It helps make statements about the likelihood of co-occurrence of a variable (e.g. the non-bare nominals in this study) and any contextual features of interest. Second, since no single factor is likely to fully explain the variation that we regularly observe in language, GoldVarb is desirable in examining multiple causes. Furthermore, if a variable turns out to play an insignificant role in the study, it is easy to re-code and re-analyze the data in GoldVarb.

Although GoldVarb was originally designed for studies in sociolinguistics, language variation, and language change, more and more researchers have successfully used it in acquisition studies, e.g. Young and Bayley (1996) for second language acquisition, and De Cat (2002), Roberts (1997), Smith et al. (2006), and Miller (2007) for first language acquisition.

4.4.3 Defining Variables

The data in this study will be analyzed according to one dependent variable and eleven independent variables (a.k.a factor groups), including three variables of nominal

internal structures, that will be examined in distributional analyses, and eight nominal external variables, that will have distributional and multivariate analyses. All the variables and the factors under each variable are presented in Table 4.5. (See Appendix II for more details about the information coded and the coding criteria.)

Table 4.5 Variables in analyzing Mandarin production data

Dependent Variable	Factors
NP (The dependent variable must be binary in GoldVarb.)	Bare: bare noun root, modified bare nominals: adjective+(DE)+noun, noun+(DE)+noun
	Non-bare: possessive+(DE)+noun, relative clause+noun, CL+noun, Num+CL+noun, Dem+CL+noun, etc.
Independent Variables	Factors
I. Variables of Nominal Internal Structures	
1. Adjective-like Modifier (Bare Nominals)	Adjective: <i>xiao</i> 'small' etc.
	N as modifier: the first N is a modifier in N-N sequence
	None
2. Modification above NP (Non-bare Nominals)	Possessive: somebody's
	Relative clauses
	None
3. Classifier	<i>Ge</i> – general classifier
	Count-classifiers, e.g. <i>zhi</i> for animals, <i>ben</i> for books.
	Mass-classifier (Massifiers), e.g. <i>ping</i> 'bottle', <i>zhang</i> 'piece'
	None
4. Determiner	Numeral
	Singular demonstrative
	Plural demonstrative
	Quantifier: <i>yixie</i> 'some', <i>mei</i> 'every', <i>henduo</i> 'many' etc.
	Interrogative determiner: <i>sheme</i> 'what', <i>na</i> 'which' etc.
	None
II. Developmental Variables	
5. Age	Nana 2;0-2;6 (average 2;3)
	Didi 2;9-3;3 (average 3;0)
6. MLU (age) Nana	1.73 (2;0)
	1.74 (2;1)
	2.11 (2;2)
	2.72 (2;3)
	3.17 (2;4)
	2.96 (2;5)
	3.18 (2;6)

Table 4.5 (continued)

MLU (age) Didi	2.34 (2;10)
	2.78 (2;11)
	2.78 (3;0)
	3.57 (3;1)
	3.21 (3;2)
	3.22 (3;3)
III. Semantic Variable	
7. Interpretation	Generic (true world-wide; characterizing statements)
	Existential
	Predicative (copular context; naming – here and now; after stative verbs e.g. <i>dang</i> ‘take as’, <i>xiang</i> ‘resemble’, <i>jiao</i> ‘called’)
	Unclear
IV. Variables of Nominal External Structures	
8. Syntactic Position	Subject
	Object, including direct objects and indirect objects
	Object of preposition
	Predicate (Post-copular)
	Topic
	Appositive (of subject, unless specified)
	Utterances with nominals only or no explicit verb
9. Verb Type	Stative
	Activity (can appear with the progressive marker <i>zai</i>)
	Resultative (Telic verbs that cannot appear with the progressive marker <i>zai</i>)
	Copula
	Utterances with noun phrases or preposition phrases only
10. Aspect Marker	<i>Le</i> ‘perfective’ (verb- <i>le</i> and sentence final - <i>le</i>)
	<i>Zai</i> ‘progressive’
	<i>Zhe</i> ‘durative’
	<i>Guo</i> ‘experiential’
	<i>You</i> ‘perfective’ (Southern dialects)
	<i>Mei</i> ‘perfective negation’
	None
11. Clause Force	Declarative
	Interrogative
	Imperative

The dependent variable is the bare or non-bare form of Mandarin nominals. Bare nominals refer to bare noun roots, such as *shu* ‘book’ or *shui* ‘water,’ different from the bare nouns in some languages that may include inflection such as *books*, *pencils*. Bare

nominals also include modified bare nominals, such as adjective-modified nominals, *hong shu* ‘red book,’ or noun-modified nominals, *xiaopengyou de xiezi* ‘shoes for children.’ Non-bare nominals refer to any nominal that includes a bare nominal and something else, e.g. *zhe yi-ben shu* ‘this one-CL book’ or *san-ben hong shu* ‘three-CL red book.’ Only common nouns in the head noun position will be coded as tokens. Pronouns and proper names (which usually appear as bare nominals) are excluded. Including the pronouns and proper names will greatly increase the number of bare nominals and skew the result.²⁰ Bare kinship terms, which are commonly used as proper names in the Chinese culture, are also excluded due to the ambiguity between common nouns and proper names. However, modified kinship terms are included because the kinship terms in these cases are clearly used as common nouns. (More transcription and coding details can be found in Appendix I and II.)

The independent variables investigated in the study are the following: (i) Nominal internal variables: Adjective-like Modifier, Modification above NP, Classifier, Determiner; (ii) Developmental variables: Age, MLU; (iii) Semantic variable: Interpretation; (iv) Nominal external variables: Syntactic Position, Verb Type, Aspect Marker, and Clause Force.^{21,22,23}

²⁰ Although proper names in the head noun position are excluded in the study, proper names in other positions will still be included. For example, in *Babi de qunzi* ‘Barbie’s skirt’ – the head noun is *qunzi* ‘skirt’ and *Babi de* ‘Barbie’s’ is a possessive modifier.

²¹ Four more variables were also coded – Sentence Type (affirmative or negative), Clause Type (main or embedded clause), *Ba* Sentence, and *Dou* Sentence – but discarded for the study for different reasons. The Sentence Type variable has no effect on the variation of bare and non-bare nominals (more details will be discussed later in the text). More details about the other three variables are provided in the following footnotes.

²² Clause Type is divided into main and embedded clauses. Relative clauses are illustrated in (i) and the emphatic *shi...de* ‘be...DE’ construction in (ii).

-
- (i) Embedded clause: Relative clause (Nana 2;6)

Fang zai **erduo** shangmian de
Put at ear on top DE
'(something) that is put on the top of the ear' (The head noun is missing.)

- (ii) Embedded clause: *Shi...de* construction (Nana 2;4)

Zhe shi **xiaopengyou** zhong de
This be child plant DE
'This is (something that) children planted.'

Note that the emphatic *shi...de* construction is different from the copular *shi* 'be' construction, and they are coded differently. The nominal, *xiaopengyou* 'children,' in sentence (ii) is coded as in the subject position for the activity verb *zhong* 'plant' in an embedded clause. In other words, *xiaopengyou* 'children' is not coded as a post-copular predicate in this example.

Examples below illustrate that different types of clauses accept different types of nominals in the respective clauses.

- (iii) Main clause:

Xuesheng/*ge xuesheng/*yi-ge xuesheng/zhe-ge xuehsng xie zuowen
student/CL student/one-CL student/this-CL student write composition
'Students/This student write(s) composition.'

- (iv) Embedded clause: relative clause

Xuesheng/*ge xuesheng/*yi-ge xuesheng/zhe-ge xuehsng xie de
student/CL student/one-CL student/this-CL student write-DE
zuowen hen youyisi
composition very interesting
'Composition written by students/this student is very interesting.'

- (v) Embedded clause: relative clause

Lisi xihuan kan **xuesheng/*ge xuesheng/yi-ge xuesheng/zhe-ge xuehsng**
Lisi like read student/CL student/one-CL student/this-CL student
xie de zuowen
write DE composition
'Lisi likes to read composition written by students/a (certain) student/this student.'

- (vi) Embedded clause: *Shi...de* construction

zhe shi **xuesheng/ge xuesheng/yi-ge xuesheng/zhe-ge xuesheng** xie de
This be student/CL student/one-CL student/this-CL student write DE
'This is (something that) written by students/a student/one student/this student.'

4.4.3.1 Developmental and semantic variables

Two developmental variables, Age and MLU, were coded. Age is an indispensable variable when studying language acquisition. However, we know from earlier work that children acquire language at widely varying rates – they usually reach a certain stage of language development at different ages – thus Age is not always an objective indicator when studying children’s language development.²⁴ Therefore, MLU was also coded. MLU is an excellent ‘index of grammatical development because almost

However, a closer look of these examples reveals that it seems to be the syntactic position that affects the nominal forms, not the clause types. The subject position in the main clause (iii) and in the embedded clause (iv) accept the same types of nominals. Contrastively, the lexically governed position in (v) and the emphatic *shi...de* construction in (vi) accept different types of nominals. According to the observation that the Clause Type variable seems to be confounded with the Syntactic Position variable and only 24 non-bare nominal tokens are found in embedded clauses, the Clause Type variable is discarded in the study.

²³ Tokens were coded if they appear in sentences with or without *ba* and *dou*. *Ba* is the marker that changes the canonical Mandarin SVO structure (ia) to SOV structure (ib).

- | | |
|---------------------------------|-----------------------------------|
| (i) a. Lisi chi-le pingguo | b. Lisi ba pingguo chi-le |
| Lisi eat-PERF apple | Lisi BA apple eat-PERF |
| ‘Lisi ate the apple.’ | ‘Lisi ate the apple.’ |

Dou ‘both, all’ is an adverb that interacts with nominals and brings out different interpretations – sentence (ii) is an example of how *dou* may influence the semantics of a sentence ((ii) is adopted from Li and Thompson 1981:336).

- (ii) zhexie xuesheng, women dou xihuan
 these student we DOU like
 ‘We like all these students.’ (Dou associates with the topic)
 ‘We all like these students.’ (Dou associates with the subject)
 ‘We all like all these students.’ (Dou associates with both the topic and the subject)

Given the semantic interaction between *dou* and nominals, I coded tokens according to whether they appear in a sentence with or without *dou*. However, *ba* and *dou* variables are excluded due to small number of tokens. Only 15 nominals appear in *ba* sentences; 3 of them are non-bare. Only 10 nominals appear in *dou* sentences; 1 of them are non-bare.

²⁴ In the Brown corpora (1973), Eve was 18-month-old when she achieved an MLU of 2, whereas Adam and Sarah were 27- and 29-month-old, respectively.

(16d), the nominal can be counted as a predicate in a post-copular position or a 'naming' utterance.

4.4.3.2 Variables of the nominal internal structures

The four variables of nominal internal structure are Adjective-like Modifier, Modification above NP, Classifier, and Determiner. Nominals possess adjective-like modifiers are coded as bare NPs, while nominals have Modification above NP, Classifier, and/or Determiner elements are non-bare NPs.

For the Adjective-like Modifier variable, the tokens were coded as having adjectives or noun modifiers. For the Modification above NP variable, the tokens were coded as having possessives or relative clauses. Nominals modified by adjectives or nouns are bare NPs and can appear with other elements to form non-bare NPs. Nominals modified by possessives or relative clauses are non-bare NPs. Noun modifiers refer to the first noun in noun-noun sequences, such as *Bani* in (17). Although there are few nominals modified by relative clauses produced by young children, some examples are found, as presented in (18). Some cases were found where children used more than one modifier. Most of these cases include a possessive and an adjective, as shown in (19).

(17) *Bani xiexie* Bare NP (Nana 2;0)
 ‘Barnie shoe’
 (Shoes that have the image of the cartoon character, Barnie, on them.)

(18) *Baba mai-de mianbao* Non-Bare NP (Didi 2;11)
 Dad buy-DE bread
 ‘the bread that Dad bought’

(19) [NON-BARE *wo-men-de* [BARE *xiaoxiao-de* *jiandao*] Non-Bare NP
 I-PL-DE small-DE scissor (Nana 2;4)
 ‘our small scissors’

Classifiers are the elements that head the Classifier Phrase projections, as argued in Chang (2007) and Huang et al. (2009). I divided the Classifier variable into three categories: the general classifier *ge*, count-classifiers for count nouns (*zhi gou* ‘CL dog’), and massifiers for mass nouns (*ping shui* ‘bottle water’).²⁵ The general classifier *ge* is believed to be a place holder whenever the speakers are unsure about which classifier to use (Hu 1993). Among all three categories of classifiers, *ge* is acquired the earliest and used most frequently. In Hu’s elicited production study of count-classifiers, children aged three to six produced *ge* 73% of the time, but other count-classifiers only 24% of the time. Therefore, *ge* should be separated from count-classifiers and coded individually. (For more details about the sub-categorization of Mandarin classifiers, see Cheng and Sybesma (1998), Chien et. al (2003), P. Li et al. (2008), and Liao and Wang (2008).)

The coding of Determiner variable includes all the elements above CIP in the data: numerals, singular demonstratives, plural demonstratives, quantifiers, and interrogative determiners. Following Huang et al. (2009) I assume that these elements head their own projections, such as NumP, DP, and QP. Since the subjects were two- and three-year-olds and did not produce complex noun phrases very often, all elements above CIP are grouped under the Determiner variable for simplicity.

4.4.3.3 Variables of the nominal external structures

There are eight variables concerning the context in which the nominals appear. For Syntactic Position, the tokens were coded based on the positions of the nominals in the clausal structure – including subjects, objects, objects of preposition/postposition, post-copular positions, topics, and appositives. Different positions will require different

²⁵ Adopting Cheng and Sybesma (1998), classifiers for mass nouns are called ‘massifiers.’

types of nominals. In section 4.2, we already saw subject-object asymmetry examples, which demonstrate that subject and object positions have different preferences for nominal types. Examples concerning other types of syntactic positions are presented below: object of preposition/postposition (20), post-copular position (21), topic (22), and appositive (23).

Examples in (20) show that bare NP, ClP, NumP, and DP are all allowed in the object position of a preposition/postposition. Although ClP without a numeral or demonstrative is not found in the prepositional phrases in the children's data, it is legitimate in the adult grammar, as shown in (20a). In Mandarin, there are prepositions (e.g. *zai* 'at,' *gen* 'with') and postpositions (e.g. *shang* 'on; on the top of'). Some prepositions and postpositions are simple, as in (20c), and some are complex, as shown in (20a, b, and d). This factor will include the objects of preposition and postposition.

- (20) a. Lisi zhu zai **sushe/jian sushe/yi-jian sushe/zhe-jian sushe** li
 Lisi live in dorm/CL dorm/one-CL dorm/this-CL dorm in
 ‘Lisi live in the dorm/a dorm/one dorm/this dorm.’
- b. gua zai **shu** shang (Didi 3;0)
 hang on tree on
 ‘Hang on the tree.’
- c. songshu zai gen **yi-ge** **ji dan** jianghua (Nana 2;5)
 squirrel PROG with one-CL egg talk
 ‘The squirrel is talking with an egg.’
- d. wo diao zai **zhe-ge** **shu** shang (Didi 3;0)
 I hang on this-CL tree on
 ‘I hang on this tree.’

In the post-copular positions, as exemplified in (21a), bare nominals are acceptable with a proper name interpretation (a common noun functions as someone's nickname or title)²⁶ and ClPs convey the predicative reading of the subject. Bare NP and ClP are found in the children's data, as shown in (21b) and (21c). NumP and DP are generally not acceptable in post-copular positions. As indicated in (21a), it is odd to use the NumP, *yi-ge xiaozhang* 'one principal' in the context. Yet a child's example of NumP is found, as shown in (21d), which sounds odd. No DP is found in the post-copular position in the child data.

- (21) a. Na-ge ren shi **xiaozhang/ge-xiaozhang/**
 that-CL person be principal/CL-principal/

?yi-ge xiaozhang/*zhe-ge xiaozhang
 one-CL principal /this-CL principal

'That person is Principal/a principal/one principal/this principal.'

- b. na shi **yaoshi** (Nana 2;5)
 that be key
 'That is Key.'

- c. zhe shi **-ge zhu** (Nana 2;5)
 this be -CL bead
 'This is a bead.'

²⁶ A bare noun has the proper name interpretation in the post-copular position in Mandarin. In the examples below, *taiyang* 'sun' has the proper name interpretation in (i), but predicative interpretation in (ii) because of the classifier. The proper name interpretation can only be conveyed by a common noun in its bare form.

- (i) Zhe shi taiyang
 this be sun
 'This is Sun.'
- (ii) Zhe shi ge taiyang
 this be CL sun
 'This is a sun.'

- d. ?zhe shi **yi-ge** **fengche** (Nana 2;4)
 this be one-CL windmill
 'This is one windmill.'

Topics are part of the common ground and usually allow only specific or definite interpretations. Bare nominals, adjective-modified nominals, and demonstrative nominals can serve as topics, as shown in (22a, b, c). Indefinite nominals, such as CIP and NumP, are not allowed in the topic position, as illustrated in (22a).

- (22) a. **Shu/*ben-shu/*yi-ben shu/zhe-ben shu**, wo kan-le
 Book/CL book/ one-CL book/this-CL book, I read-PERF
 '(As for) The book/a book/one book/this book, I read (it).'
- b. **xiao laoshu**, wo zhao-dao le (Nana 2;2)
 small mouse, I search-reach PREF
 '(As for) The small mouse, I found it.'
- c. **na-ge hua**, song gei wo baba (Didi 3;3)
 that-CL flower, give to I dad
 '(As for) That flower, (it is) for my dad.'

Not much is known about Mandarin appositives. In line with Chao (1968), I define appositives as an inserted phrase referring to the same entity as the arguments, for example, the subject in the cases of (23).

- (23) a. Zhangsan, ***linju/?xin linju/*ge linju/*yi-ge linju/**
 Zhangsan, neighbor/new neighbor/CL neighbor/one-CL neighbor/
zhe-ge linju/wo-de linju yao lai kan wo
 this-CL neighbor /my neighbor want come visit I
 'Zhangsan, a new neighbor/this neighbor/a neighbor of mine, will come visit me.'
- b. ta, **xiao beibi**, mei-you chuan yifu (Nana 2;2)
 he, little baby, not-have wear clothes
 'He, the little baby, does not have clothes on.'

Omitting the appositives, the sentence is still a complete sentence and the noun phrase maintains its reference. Example (23a) shows that demonstrative and possessive

nominals are acceptable as appositives (adjective-modified nominals are marginally acceptable), while other types of nominals cannot appear in the appositive position. In the child's example (23b), the adjective-modified nominal is readily acceptable possibly because the subject pronoun prompts the appositive nominal to have a deictic and definite interpretation.

For Verb Type, the tokens were coded based on the types of verb they appear with: stative, activity, resultative, or copula. This categorization is not the same as the commonly used categorization of situation types, i.e. statives, activities, accomplishments, and achievements (Vendler 1967). Instead, the categorization in Chang (2002) was adopted to better capture the commonly used verbs in young Mandarin-speaking children. Verbs are coded as activity verbs if they can co-occur with the progressive marker *zai*. In addition to the resultative verb compounds, such as *da-po* 'hit-break,' which are commonly seen in Mandarin, resultative verbs include telic verbs that cannot appear with the progressive marker *zai*. (For more details about the categorization, see Chang 2002.)

The examples below illustrate the interaction between verb types and object nominals: a stative verb (24), an activity verb (25), and a resultative verb (26) (examples of copular verbs have already been shown in (21)).

- (24) a. Zhangsan xihuan **gou/*zhi-gou/*yi-zhi gou/zhe-zhi gao**
 Zhangsan like dog/CL-dog/one-CL dog/this-CL dog
 'Zhangsan likes dogs/this dog.'

- b. wo yao **wawa** (Nana 2;1)
 I want doll
 'I want the doll.'

- (25) a. Lisi xie **xiaoshuo/*ben xiaoshuo/*yi-ben xiaoshuo/*zhe-ben xiaoshuo**
 Lisi write novel/CL-novel/one-CL novel/this-CL novel
 'Lisi writes novels.'

b. daren chi **jirou** (Nana 2;0)
 adult eat chicken
 'Adults eat chicken.'

(26) a. Lisi xie-wan **lunwen/*pian lunwen/yi-pian lunwen/zhe-pian lunwen**,
 Lisi write-finish thesis/CL-thesis/one-CL thesis/this-CL thesis,

jiu qu Ouzhou dujia
 as soon as go Europe vacation

‘As soon as Lisi finishes writing the thesis/one thesis/this thesis, he (will) go to Europe for vacation.’

b. chi-diao **tutu** (Nana 2;0)
 eat-up rabbit
 ‘Ate up the rabbit.’

In (24a), the stative verb accepts bare nominals and demonstrative nominals. Example (25a) shows that activity verbs only accept bare nominals and reject non-bare nominals. Resultative verbs, as shown in (26a) can co-occur with bare NP, NumP, or DP, but not with determiner-less CLP. The children’s data in (24b), (25b), and (26b), also show that they start to use different types of verbs at a very young age. It is clear that different verb types select nominals with different structures as their complements. After an aspect marker is added to these bare verb examples, the acceptable object types change.

The variable Aspect Marker consists of the four Mandarin aspect markers: perfective marker *le*, progressive marker *zai*, durative marker *zhe*, and the experiential marker *guo*, as well as the perfective marker *you* in the southern dialects²⁷ (Li and Thompson 1981:431) and the perfective negation marker *mei(you)*. Examples of *le* were

²⁷ Although at the time of the recording, both subjects were Mandarin monolinguals living in the United States, the inclusion of *you* is necessary because their parents grew up in Taiwan, where even monolingual Mandarin-speakers often use *you* as the perfective marker — a clear influence of Taiwanese, a Southern Min dialect.

already discussed in section 4.2 and repeated here as (27a) and (27b). Sentences (27c) and (27d) are children's examples of *le*.

(27) Perfective marker *le*

- a. Zhe-zhi mao chi-le **yu/guantou yu/tiao yu/san-tiao yu/zhe-tiao yu/**
 This-CL cateat-PERF fish/can fish/CL fish/three-CL fish/this-CL fish

quanbude yu
 all fish

‘This cat ate (the) fish/canned fish/a fish/three fish/this fish/all the fish.’

- b. zhe-zhi mao chi **yu/guantou yu/*tiao yu/*san-tiao yu/*zhe-tiao yu/**
 This-CL cat eat fish/canned fish/CL fish/three-CL fish/this-CL fish

***quanbude yu**
 all fish

‘This cat eats fish/canned fish/a fish/three fish/this fish/all the fish.’

- c. wo chi-le **mianbao**, haiyou, yi-zhi hongluobo (Didi 2;10)
 I eat-PERF bread, additionally, one-CL carrot
 ‘I ate bread, and a carrot as well.’

- d. diao-le **yi-zhi yu** (Didi 2;10)
 catch-PERF one-CL fish
 ‘Caught one fish.’

Examples of other aspect markers are presented below. The principle found for *-le* also applies to the progressive marker *zai*. A verb occurring with the progressive *zai* can co-occur with both bare and non-bare nominals, but a bare verb can only appear with bare NP and modified NP, as shown in (28).

(28) Progressive marker *zai*

- a. ta zai-yanjiu **xiaoshuo/waiguo xiaoshuo/ben xiaoshuo/**
 he PROG-study novel/foreign novel/CL novel/

san-ben xiaoshuo/zhe-ben xiaoshuo/quanbude xiaoshuo
 three-CL novel/this-CL novel/all novel

‘He is studying novels/foreign novels/a novel/three novels/this novel/all the novels.’

- b. ta yanjiu **xiaoshuo/waiguo xiaoshuo/*ben xiaoshuo/**
 he study novel/ foreign novel/CL novel/

***san-ben xiaoshuo/*zhe-ben xiaoshuo/?quanbude xiaoshuo**
 three-CL novel/this-CL novel/all novel

‘He studies novels/foreign novels/a novel/three novels/this novel/all the novels.’

- c. ta zai-he **tangtang** (Nana 2;4)
 he PROG-drink soup
 ‘He is drinking soup.’

- d. ta-men zai-zuo **yi-ge xiexie** (Nana 2;4)
 he-PL PROG-make one-CL shoe
 ‘They are making a pair of shoes.’ (The child used the general classifier *ge* instead of the correct classifier, *shuang* ‘pair.’)

In (29) below, the experiential marker *guo* appears with bare nominals or NumP in the object position to express past experience. Sentences with the durative marker *zhe* or the perfective marker *you* in southern China or Taiwan, (30) and (31), can take all types of nominals as the verb complements.

- (29) a. Lisi xie-guo **xiaoshuo/*ben xiaoshuo/yi-ben xiaoshuo/*zhe-ben xiaoshuo**
 Lisi write-EXP novel/CL-novel/one-CL novel/this-CL novel
 ‘Lisi has the experience of writing novels/one novel.’

- b. wo you kan-guo **qier** (Didi 2;11)
 I have see-EXP penguin
 ‘I have seen penguins.’

- (30) a. Lisi zheng kan-zhe **jiemu/ge jiemu/yi-ge jiemu/zhe-ge jiemu** ne
 Lisi exactly watch-DUR show/CL-show/one-CL show/this-CL show SFP
 ‘Lisi is watching the show/a show/one show/this show.’²⁸

²⁸ The use of *zheng* and *ne* makes a sentence with the durative marker *zhe* more natural. *Zheng* means ‘exactly or indicating the progress of an action’ and *ne* is ‘a sentence final particle to emphasize a statement’ (Lü 1980: 412-413, 670-671).

- b. qier fang-zhe
 penguin put-DUR
 ‘The penguin is put (here).’

- (31) a. Liming you-chi **dangao/ge dangao/yi-ge dangao/zhe-ge dangao**
 Liming PERF-eat cake/CL cake/one-CL cake/this-CL cake
 'Liming ate the cake/a cake/one cake/this cake.'

- b. wo jiejie ye you-mo **yang** (Nana 2;4)
I older:sister also PERF-stroke sheep²⁹
'My older sister stroked the sheep, too.'

With respect to the Sentence Type variable, affirmative or negative, the two most commonly used negation markers in Mandarin are *bu* ‘not’ and *mei(you)* ‘didn’t; haven’t.’ Examples in (32a, b) show that a bare verb and its negative counterpart take the same types of nominals; hence causing no variation of nominal forms. Sentence (32c) shows that children have started to use the negation marker *bu* as early as at age two.

- (32) a. Liming xiu **che/*liang che/?yi-liang che/zhe-liang che**
 Liming fix car/CL car/one-CL car/this-CL car
 'Liming fixes cars/this car.'

- b. Liming bu xiu **che/*liang che/?yi-liang che/zhe-liang che**
 Liming not fix car/CL car/one-CL car/this-CL car
 'Liming doesn't fix cars/this car.'

- c. bu shi **niao** (Nana 2;0)
not be bird
'(It) is not a bird.'

However, affirmative sentences with a verb and a perfective marker, as in (33a), behave differently from their negative counterparts, i.e. the sentences with the perfective negation marker *mei(ou)* (33b). As discussed before, the perfective marker may prompt the use of more types of non-bare nominals, compared to the verb without any aspect. Sentences with *-le* can appear with NP, CIP, NumP, or DP. However, sentences with the

²⁹ In this dissertation, if the English gloss for a Mandarin word is a compound, a colon is used to save space, e.g. *gege* ‘elder:brother.’ (Li and Thompson 1981).

negation marker *mei(you)* display the same behavior as sentences with the negation marker *bu* – accepting only bare NP and DP.

- (33) a. Liming xiu-le **che/liang che/yi-liang che/zhe-liang che**
 Liming fix-PERF car/CL car/one-CL car/this-CL car
 ‘Liming fixed the car/a car/one car/this car.’
- b. Liming mei(you) xiu **che/*liang che/?yi-liang che/zhe-liang che**
 Liming not(have) fix car/CL car/one-CL car/this-CL car
 ‘Liming didn’t fix the car/this car.’
- c. Nana meiyou hua **xiaopengyou** (Nana 2;1)
 Nana not:have draw child
 ‘Nana did not draw children.’

In other words, Sentence Type – affirmative or negative – makes no effect on the structure of nominals in the case of regular affirmative vs. regular negative (32a, b), but it does play a role in the case of perfective affirmative vs. perfective negative (33a, b). Obviously, it is the perfective aspect of these sentences that causes the variation of nominal forms. Therefore, the perfective negation marker *mei(you)* is coded under the variable of Aspect Marker because Sentence Type does not have an effect on the variation of bare and non-bare nominals. Therefore, Sentence Type needs not to be considered as a variable for the current study.

The last nominal external variable concerns the clausal structure.³⁰ The variable Clause Force categorizes the data into declaratives, interrogatives, and imperatives. The following sentences present the effect of the clause force on the nominal structure in object position.

³⁰ The other four clausal variables coded in the data – Sentence Type (affirmative or negative), Clause Type (main clause or embedded clause), *Ba* Sentence, and *Dou* Sentence – are discarded for the current study for different reasons. See footnotes 21, 22, 23 for details.

- (34) a. Xiaoming he **tang/*wan tang/*yi-wan tang/zhe-wan tang**
 Xiaoming drink soup/CL soup/one-CL soup/this-CL soup
 ‘Xiaoming drinks soup/this soup.’
- b. chi **yaya** (Nana 2;0)
 eat duck
 ‘(I) eat duck.’
- (35) a. Xiaoming he **tang/*wan tang/*yi-wan tang/zhe-wan tang** ma?
 Xiaoming drink soup/CL soup/one-CL soup/this-CL soup MA
 ‘Does Xiaoming drink soup/this bowl of soup?’
- b. **shou** lei? (Nana 2;1)
 hand how:about
 ‘How about hands?’
- (36) a. Xiaoming, he **tang/wan tang/yi-wan tang/zhe-wan tang**
 Xiaoming, drink soup/CL soup/one-CL soup/this-CL soup
 ‘Xiaoming, have soup/some soup/one bowl of soup/ this bowl of soup!’
- b. qu na **panzi** (Nana 2;0)
 go get plate
 ‘Go get plates.’

The declarative and interrogative sentences, (34a) and (35a), can only appear with bare or demonstrative nominals, while the imperative sentence (36a) allows all four types of nominals in the example. The child data shows that children produce sentences with different clause force at a very young age.

4.5 Results

This section presents the study results of 1969 nominal tokens extracted from thirteen recordings.

Table 4.6 Nana’s tokens of coded nominals across age

Age	2;0	2;1	2;2	2;3	2;4	2;5	2;6	
Tokens	38	41	126	47	169	110	121	652 (total)

Table 4.7 Didi’s tokens of coded nominals across age

Age	2;10	2;11	3;0	3;1	3;2	3;3	
Tokens	163	277	287	147	194	249	1317 (total)

While there is little research about the amount of tokens necessary to carry out a variable rule analysis on children's speech, Guy (1980) finds that a sufficient number of tokens should be approximately 20 per cell (cited in Bayley and Young forthcoming: 9). The variables in this study, excluding the MLU variable, have up to seven factors in each group. Based on Guy's findings, 140 tokens (7x20) from each subject should be sufficient. However, we know that some cells may have less than 20 tokens, and therefore 140 tokens per subject may be insufficient. In order to reach a statistically significant result, usually 500 tokens from each subject should be the minimum (Bayley and Young, forthcoming: 8). The numbers of tokens from both Nana (652) and Didi (1317) meet this criterion.

Inter-rater reliability between the two Mandarin native speaking coders (researcher A coded 10% of researcher B's coding and vice versa) is found to be 98.22% and 98.5% respectively. The discrepancy is mainly found in the coding of the Interpretation variable. All differences were discussed and corrected before running the GoldVarb analysis.

4.5.1 Distributional Analyses

4.5.1.1 General distribution and developmental variables

Table 4.8 presents the distribution of the dependent variable. Among all 1969 tokens analyzed, there are more bare nominals (76%) than non-bare nominals (24%).

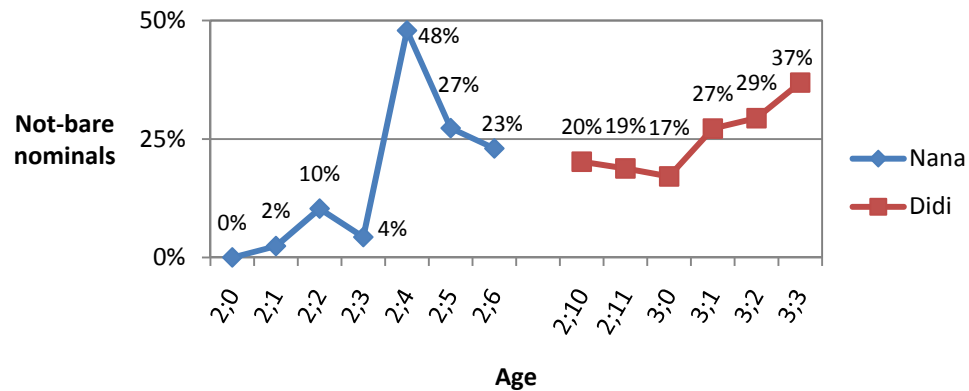
Table 4.8 Overall distribution of bare and non-bare nominals

Dependent Variable	Factors	
NP	Bare	76% (1491)
	Non-bare	24% (478)
	<i>Total</i>	<i>100% (1969)</i>

Figure 4.2 shows the percentage of non-bare nominals at different ages, which

shows that the percentage of non-bare nominals increases as the children grow older.³¹

Figure 4.2 Distribution of non-bare nominals across age



In Figure 4.2, the production of non-bare nominals was very low, 0-10%, when Nana just turned two years old (2;0 - 2;3). When she approached two and a half years old, the percentage jumped to 23-27%. The highest percentage produced was at two-year-and-four-month old, 48%. It seems that the major increase is between 2;3 and older. The situations of that specific recording include story-telling and sitting and chatting on the corridor outside her apartment. Since telling stories and playing outdoors are also the

³¹ The distribution of non-bare nominals across MLU is presented below. The length of MLU is mainly in accord with the age. The distribution of non-bare nominals ordered by age (Figure 4.2) and that across MLU (below) are almost identical, except in age 2;5 and 3;1 (indicated in bold and italic fonts). Accordingly, the developmental trend discussed in this study will be mainly presented across age. MLU will be discussed when necessary.

(i) Distribution of non-bare nominals across MLU

MLU	1.73	1.74	2.11	2.72	2.96	3.17	3.18
(Age)	(2;0)	(2;1)	(2;2)	(2;3)	(2;5)	(2;4)	(2;6)
	0%	2%	10%	4%	27%	48%	23%
MLU	2.34	2.78	2.78	3.21	3.22	3.57	
(Age)	(2;10)	(2;11)	(3;0)	(3;2)	(3;3)	(3;1)	
	20%	19%	17%	29%	37%	27%	

situations in other recordings, what makes Nana utilize more non-bare nominals in that recording is unclear. As for Didi's data, the production of non-bare nominals does not change much with time.

4.5.1.2 Variables of the nominal domain

In this section, I present the distributional analyses of the variables that could not go through the multivariate regression analysis, namely, the nominal internal variables: Adjective-like Modifier, Modification above NP, Classifier, and Determiner. Nominal internal factors have a 'knock out' result – the value in a cell is either 0 or 100 % – which means these factors do not affect the variation (either bare or non-bare) of the nominal phrases. Factors that have a 'knock out' cell are not able to go through the regression analysis of GoldVarb. Therefore, these nominal internal factors are analyzed through distributional analysis.

Table 4.9 presents the distribution of adjective-like modifiers.

Table 4.9 Distribution of adjective-like factors in bare and non-bare nominals

Independent Variable	Factors	% of all non-bare nominals (n = 478)	% of all bare nominals (n = 1491)	% of all nominals (n = 1969)
	<i>without Adjective</i>	88% (422)	77% (1147)	80% (1569)
Adjective-like Modifier	Adjective	10% (48)	18% (264)	16% (312)
	N as modifiers	2% (8)	5% (80)	5% (88)
	<i>Total</i>	<i>12% (56)</i>	<i>23% (344)</i>	<i>20% (400)</i>

Table 4.9 shows that a total of 20% (400/1969) nominals contain an adjective-like element. Among these adjective- or noun-modified nominals, 86% (344/400) of them remain as bare nominals. The other 14% (56/400) appear with other elements to form non-bare nominals. Overall, there are more nominals modified by adjectives, 16% (312/1969), than by noun-modifiers, 5% (88/1969). Examples from children's data that

have adjective-like modifiers are presented below.

- (37) a. **xiao niao** si-diao
 little bird die-off
 ‘The little bird died.’

Adjective (Nana 2;0)

- b. **hong se dangao**
 red color cake
 ‘cake in red color’

N as a modifier (Nana 2;2)

Table 4.10 presents the distribution of the nominal internal factors in non-bare nominals.

Table 4.10 Distribution of nominal internal factors in non-bare nominals

Independent Variables	Factors	% of all non-bare nominals (n = 478)	N
	<i>without Modification above NP</i>	49%	(232)
Modification above NP	Possessive	49%	(236)
	Relative clauses	2%	(10)
	<i>Total</i>	51%	(246)
	<i>without Classifier</i>	55%	(264)
Classifier	Ge – general classifier	37%	(176)
	Count-classifiers	8%	(36)
	Massifiers	0.4%	(2)
	<i>Total</i>	45%	(214)
	<i>without Determiner</i>	52%	(247)
Determiner	Numeral	27%	(129)
	Singular demonstrative	17%	(79)
	Plural demonstrative	-	-
	Quantifier	3%	(15)
	Interrogative determiner	2%	(8)
	<i>Total</i>	48%	(231)

As shown in Table 4.10, a total of 51% non-bare nominals appear with possessives or relative clauses, 45% with classifiers, and 48% with determiners. Some of the non-bare nominals appear with two or all three types of these variables. I will discuss the intersection among these variables later.

Among non-bare nominals, 49% of them have possessives and 2% of them, relative clauses. Examples from children's data that have possessives or relative clauses are presented below.

(38) a. na shi **wo-de** **shoutao** Possessive (Nana 2;2)
that be I-DE glove
‘That is my glove.’

b. **wo-men de** **xiaoxiaode jiandao** ne? Possessive + Adjective
I-PL DE small DE scissor where:about (Nana 2;4)
‘Where is our small scissor?’

c. **baba mai de mianbao** gei wo Relative clause (Didi 2;11)
dad buy DE bread give I
‘Give me the bread that Dad bought.’

A total of 37% non-bare nominals co-occur with the general classifier *ge*, 8% with count-classifiers, and 0.4% with massifiers. The massifier tokens were only found in the data of the older subject, not the younger one. The following sentences are children’s examples of classifiers.

(39) a. wo hui pa **zhe-ge** **xiaochou** classifier *ge* (Nana 2;2)
I will scare this-CL clown
‘I will be scared of this clown.’

b. haiyou **yi-zhi** **maotouying** count-classifier (Nana 2;2)
in:addition one-CL owl
‘An owl as well.’

c. **hao da -ping de niunai** gei wo massifier (Didi 2;11)
very big -bottle DE milk give me
‘Give me a very big bottle of milk.’

A total of 27% non-bare nominals appear with numerals, 17% with singular demonstratives, 3% with quantifiers, and 2% with interrogative determiners. Sentences in (40) present examples of determiners from children's data.

(40) a. you **san-ge** **huahua** numeral (Nana 2;2)
have three-CL flower
‘There are three flowers.’

- b. wan **na-ge** **xiaopengyou** singular demonstrative (Nana 2;2)
 play that-CL child
 'Play with that child.'
- c. **haoduo** **xiao niao** yao qu bie difang quantifier (Nana 2;4)
 very:many small bird want go other place
 'Many small birds want to go to other places.'
- d. zhe shi **shenme shengyin** interrogative determiner (Nana 2;6)
 this be what sound
 'What sound is this?'

Examining all the nominal internal factors, adjectives, possessives, the general classifier *ge*, and numerals are the elements that most frequently added to nouns.

4.5.1.3 Cross-tabulation between nominal internal variables and age

The following subsections present the cross-tabulations between the nominal internal variables and age. The cross-tabulation analysis provided by the GoldVarb program shows the intersection of two factors (Tagliamonte 2006:182).

4.5.1.3.1 Adjective-like Modifier/Modification above NP and age

First, the cross-tabulation of adjective-like modifiers and age among all nominals in Table 4.11 confirms the finding that adjectives, which occupy the NP adjunct position, are acquired before other modification, at the age of 2;0. The single token of 'noun as modifier' produced at age 2;0 is arguably a fixed term learned from adults, as shown in (41).³² Therefore, the only type of modifications used in the file of Nana 2;0 are adjectives.

³² Although modified nominals may be transcribed as one inseparable morpheme in young children's speech, if evidence exists that the children voluntarily use the modifier and the noun separately in the same file or in a previous file, the modifier and the noun will be counted independently. The bare noun *xiexie* 'shoe' is used in Nana 2;0; therefore, *Barnie xiexie* 'Barnie shoe' is coded as a nominal with a noun modifier.

Table 4.11 The use of adjective-like modifiers in all nominals across age

(tokens)	2;0	2;1	2;2	2;3	2;4	2;5	2;6	2;10	2;11	3;0	3;1	3;2	3;3	Total
Without adjective-like modifiers	34	38	119	43	139	90	110	118	207	238	100	152	181	1569
Adjectives	3	3	6	4	25	13	7	40	54	39	34	27	57	312
Noun as modifiers	1	-	1	-	5	7	4	5	16	10	13	15	11	88

- (41) Mama mai **Barnie xiexie** gei wo chuan (Nana 2;0)
 Mom buy **Barnie shoe** for I wear
 ‘Mom bought shoes with the **Barnie** image for me to wear.’

Table 4.12 presents the cross-tabulation of modification above NP and age in non-bare nominals.

Table 4.12 The use of modification above NP in non-bare nominals across age

(tokens)	2;0	2;1	2;2	2;3	2;4	2;5	2;6	2;10	2;11	3;0	3;1	3;2	3;3	Total
Without modification	-	-	5	1	49	12	13	7	21	28	18	19	59	232
Possessives	-	1	8	1	32	17	15	25	29	18	22	38	30	236
Relative clauses	-	-	-	-	-	1	-	1	2	3	-	-	3	10

As shown in Table 4.12, possessives started to emerge at age 2;1. Adjectives remain the first emerged and most commonly used modification across ages, and possessives, the second. As the age increases, children use more types of modifications.

4.5.1.3.2 Classifiers and age

The use of classifiers across age is presented in Table 4.13.

Table 4.13 The use of classifiers in non-bare nominals across age

(tokens)	2;0	2;1	2;2	2;3	2;4	2;5	2;6	2;10	2;11	3;0	3;1	3;2	3;3	Total
Without classifiers	-	1	8	2	33	19	17	27	34	29	22	39	33	264
General CL: <i>ge</i>	-	-	4	-	46	11	11	1	10	14	14	16	49	176
Classifiers	-	-	1	-	2	-	-	5	7	5	4	2	10	36
Massifiers	-	-	-	-	-	-	-	-	1	1	-	-	-	2

As shown in Table 4.13, of all non-bare nominals, most do not appear with any classifiers at the beginning. As age increases, the tokens of the general classifier *ge* and the count-classifiers increases. Table 4.13 supports that *ge* is mastered before other classifiers. The

three count-classifiers that were used before age two-and-a-half are one token of *zhi* ‘classifier for animals’ and two tokens of *zhong* ‘kind,’ which would be a special classifier to take note for future study, as shown below.³³

- (42) a. **yi-zhi maotouying** (Nana 2;2)
 one-CL owl
 'an owl'
- b. **zhei-zhong chezi** (Nana 2;4)
 this-CL car
 'this type of car'
- c. yong **zhei-zhong tangchi** he tangtang (Nana 2;4)
 use this-CL spoon drink soup
 'Use this type of spoon to drink soup.'

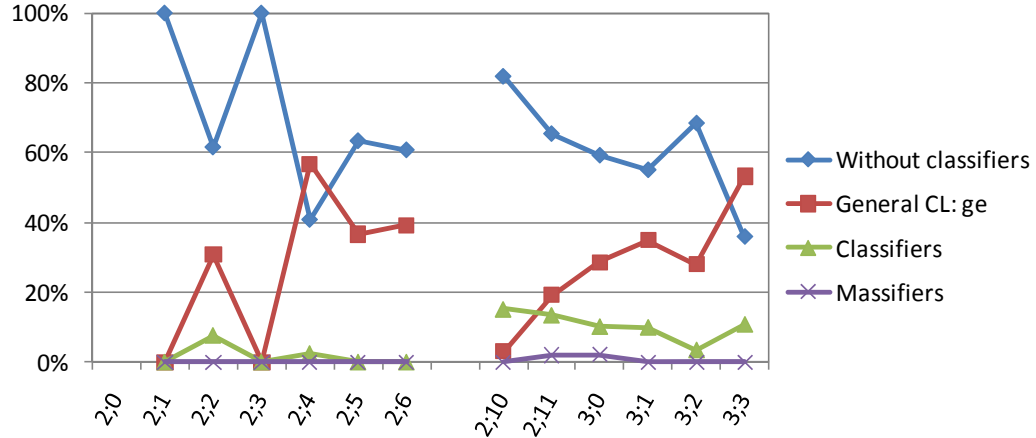
The data shows that massifiers emerge around age three although the tokens of the massifiers are too small to display any development trend. The two examples of massifiers are presented below.

- (43) a. **hao da -ping de niunai** gei wo (Didi 2;11)
 very big -bottle DE milk give me
 'Give me a very big bottle of milk.'
- b. chi **yi-kou mianbao** (Didi 3;0)
 eat one-mouthful bread
 'Eat a mouthful of bread.'

³³ Since *zhong* ‘kind’ can appear with count nouns, it is counted as a count-classifier. However, it can also appear with all types of nouns, including mass and abstract nouns, such as *ai* ‘love’ and *xiangfa* ‘thought’. Liao and Wang (2008) argue that kind-classifiers are syntactically different from count-classifiers and massifiers and should be treated as an independent category. If they are in the right track, it would be interesting to see whether children differentiate kind-classifiers from other classifiers in the development process. For example, do children acquire kind-classifiers before massifiers as the data suggest here? Studies of Chien et al. (2003) and P. Li et al. (2008) shows that children differentiate count-classifiers from massifiers in Mandarin. To my knowledge, no acquisition studies have been conducted about the difference among kind-classifiers, count-classifiers, and massifiers in Mandarin.

Figure 4.3 presents the percentage of each classifier type in non-bare nominals across age (including non-bare nominals without classifiers).

Figure 4.3 Percentage of each classifier type in non-bare nominals



As shown in Figure 4.3, non-bare nominals with zero classifiers and nominals with the general classifier *ge* both move toward 50% by age 3;3, which means the percentage of nominals with classifiers increases as children grow. The increase of the percentage of CIP and NumP among non-bare nominals also indicates that the structure of non-bare nominals becomes more complicated with the phrasal projections of CIP and NumP. (All NumPs in the data dominate an overt CIP, which will be discussed in the next section).

4.5.1.3.3 Determiners and age

Third, the use of determiners across age is presented in Table 4.14.

Table 4.14 The use of determiners in non-bare nominals across age

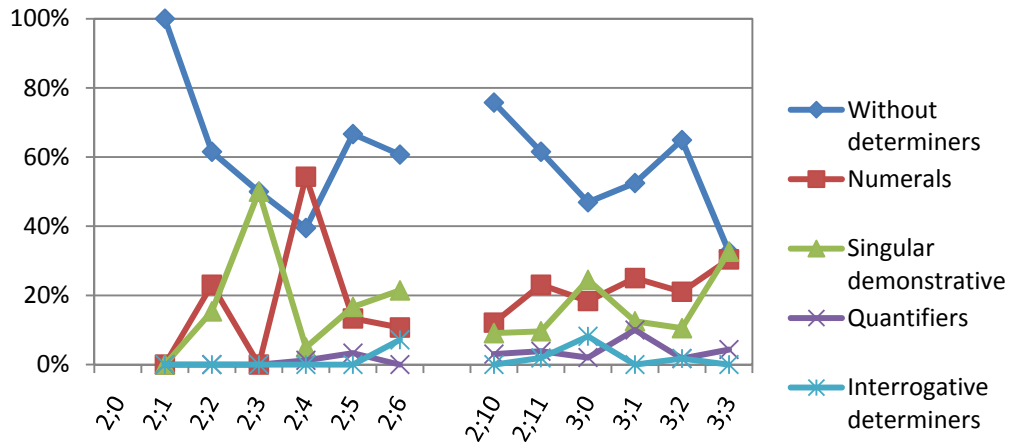
(tokens)	2;0	2;1	2;2	2;3	2;4	2;5	2;6	2;10	2;11	3;0	3;1	3;2	3;3	Total
Without determiners	-	1	8	1	32	20	17	25	32	23	21	37	30	247
Numerals	-	-	3	-	44	4	3	4	12	9	10	12	28	129
Singular demonstratives	-	-	2	1	4	5	6	3	5	12	5	6	30	79
Quantifiers	-	-	-	-	1	1	-	1	2	1	4	1	4	15
Interrogative determiners	-	-	-	-	-	-	2	-	1	4	-	1	-	8

Table 4.14 shows that numerals and singular demonstratives emerge in children's speech at the same time at age 2;2. The numerals used at age 2;2 are *yi* 'one' and *san* 'three.' Nana used both singular demonstratives, *zhe* 'this' and *na* 'that' at age 2;2 and both co-occurred with the general classifier *ge*. Quantifiers do not appear until 2;4, and interrogative determiners not until 2;6. The two tokens of quantifiers in Nana's data are *henduo* 'many' and the two interrogative determiners are *shei* 'who' and *shenme* 'what'. Some examples are provided below.

- (44) a. you **san-ge** **huahua** (Nana 2;2)
have three-CL flower
'There are three flowers.'
- b. wan **na-ge** **xiaopengyou** (Nana 2;2)
play that-CL child
'Play with that child.'
- c. wo hui pa **zhe-ge** **xiaochou** (Nana 2;2)
I will fear zhe-CL clown
'I'll be scared of this clown.'
- d. **haoduo** **xiao niao** yao qu bie difang (Nana 2;4)
very:many small bird want go other place
'Many small birds want to go to another place.'
- e. zhe shi **shenme** **shengyin** (Nana 2;6)
this be what sound
'What sound is this?'

Among all non-bare nominals, the percentage of different determiners used at different age is presented in Figure 4.4.

Figure 4.4 Percentage of each determiner type in non-bare nominals



It shows that when children become older, the percentage of non-bare nominals without any determiners decreases, while the percentage of those with numerals and singular determiners slightly increases.

To summarize, the findings according to the cross-tabulation of nominal internal variables and age are as follows. Adjectives emerges first among modification elements (age 2;0). The general classifier *ge* and a count-classifier emerge at the same time (age 2;2), while mass-classifiers are only produced by the older child (starting from age 2;11). Numerals and demonstratives also first appear in the same file (age 2;2), but numerals are overall used more often than demonstratives.

4.5.1.4 Cross-tabulation among nominal internal variables

Nominals may have very complex structure and include elements of different syntactic categories. For example, a nominal can have a demonstrative, a numeral, a classifier, a relative clause, and an adjective, as demonstrated in (45).

- (45) **zhe san-fu zai Faguo mai de hao hua**
 this three-CL in France buy DE good painting
 'these three good paintings bought in France'

Since these nominal internal elements may co-occur, it is important to examine their co-occurrence in the child data. Table 4.15 and Table 4.16 present the intersection of adjective-like modifiers and classifiers and that of modification above NP and classifiers respectively.³⁴

Table 4.15 Cross-tabulation of adjective-like modifiers and classifiers in non-bare nominals

	Without classifiers		General CL: <i>ge</i>		Count-classifiers		Massifiers		Total (478)	
Without adjective-like modifiers	57%	(241)	36%	(151)	7%	(29)	.2%	(1)	100%	(422)
Adjectives	42%	(20)	42%	(20)	15%	(7)	2%	(1)	100%	(48)
Noun as modifiers	38%	(3)	62%	(5)	-		-		100%	(8)

Table 4.16 Cross-tabulation of modification above NP and classifiers in non-bare nominals

	Without classifiers		General CL: <i>ge</i>		Count-classifiers		Massifiers		Total (478)	
Without modification above NP	11%	(25)	73%	(170)	15%	(35)	.9%	(2)	100%	(232)
Possessives	97%	(230)	2%	(5)	.4%	(1)	-		100%	(236)
Relative clauses	90%	(9)	10%	(1)	-		-		100%	(10)

Focusing on the tokens in the cells shown in bold in Table 4.15 and Table 4.16, we can see that in the columns of classifiers, the general classifier *ge* appears with all types of adjective-like modifiers and modification above NP, count-classifiers appear with adjectives and possessives, while massifiers with adjectives only. From the perspective of modification in general, only adjectives appear with all types of classifiers. Possessives are used with both *ge* and count-classifiers. All types of adjective-like modifiers and modification above NP are found to co-occur with the general classifier *ge*. The result shows that adjectives and possessives are not only the two most commonly-used modifications, they are acquired earlier as shown in the previous section, and they

³⁴ This section concerns complex non-bare nominals. Accordingly, bare nominals with adjective-like modifiers are irrelevant and not discussed.

also allow more varieties of classifiers to co-occur with them.

The cross-tabulation of determiners and classifiers in Table 4.17 also presents some interesting results.

Table 4.17 Cross-tabulation of determiners and classifiers in non-bare nominals

	Without classifiers	General CL: <i>ge</i>	Count-classifiers	Massifiers	Total (478)
Without determiners	97% (240)	2% (5)	.4% (1)	.4% (1)	100% (247)
Numerals	-	83% (107)	16% (21)	1% (1)	100% (129)
Singular demonstratives	6% (5)	78% (62)	15% (12)	-	100% (79)
Quantifiers	73% (11)	13% (2)	13% (2)	-	100% (15)
Interrogative determiners	100% (8)	-	-	-	100% (8)

In the first row of Table 4.17, which represents non-bare nominals without any determiners, there are six tokens of non-bare nominals appearing only with classifiers without any determiners – in other words, these are CL-Ns (CLPs). These bare CL-Ns are interesting in that they contradict the common belief that a classifier must occur with a number, a demonstrative, or a quantifier (Li and Thompson 1981:104).³⁵ Among these tokens, the two-year old girl, Nana, produced three of them using the general classifier *ge*, as shown below.

- (46) a. zhe shi **-ge yaoshi** (Nana 2;5)
 this be **-CL key**
 'This is a key.'
- b. zhe shi **-ge zhu**, xiaoxiao de zhu (Nana 2;5)
 this be **-CL pig**, small DE pig
 'This is a pig, a small pig.'
- c. haiyou **-ge ice cream** (Nana 2;6)
 in: addition **-CL ice cream**
 '(There's) an ice cream left.'

The three year-old boy, Didi, produced the other three tokens and each appears

³⁵ Chang (2007) argues that CLPs without numerals (NumP) or determiners (DP) is a legitimate structure in Mandarin.

with a different type of classifiers: the general classifier *ge*, a classifier *pian* ‘piece’ (which can be a count-classifiers for leaves or pedals or an arguably massifier for bread or pizza), and a massifier *ping* ‘bottle,’ as shown below.

- (47) a. chi **hao da -pian de bisa** (Didi 2;10)
 eat very big -CL(piece) DE pizza
 ‘Eat a very big slice of pizza.’
- b. **hao da -ping de niunai** gei wo (Didi 2;11)
 very big -CL(bottle) DE milk give me
 ‘Give me a very big bottle of milk.’
- c. hua **-ge hao da de shan** (Didi 3;1)
 draw -CL very big DE mountain
 ‘Draw a very big mountain.’

These six sentences are all grammatical sentences in the adult language, except in (46b), where the correct classifier for *zhu* ‘pig’ should be *zhi* ‘classifier for animals,’ and in (47b), where ‘give me’ will precede the noun phrase in adult language. Nonetheless, the structure of these noun phrases are all grammatical.

Examining Table 4.17 from the perspective of classifiers, we can see that their intersection with determiners is similar to that with modification. The general classifier *ge* and count-classifiers both co-occur with singular demonstrative, numerals, and quantifiers, while the massifiers appear with numerals only. For the distribution of determiners, 100% non-bare nominals with a numeral co-occur with a classifier and they appear with all types of classifiers. Of the non-bare nominals with a singular demonstrative (no plural demonstrative is found), a total of 93% appear with a classifier and they only appear with *ge* (78%) and count-classifiers (15%). For non-bare nominals with quantifiers, 26% of them (4 tokens) appear with a classifier. Quantifier tokens, like demonstrative ones, do not co-occur with massifiers. Interrogative determiners tokens never interact with classifiers.

At last, the concurrence of determiners and adjective-like modifiers and other modification is presented in Table 4.18 and Table 4.19 respectively.

Table 4.18 Cross-tabulation of determiners and adjective-like modifiers in non-bare nominals

	Without det.		Numerals		Sing. dem.		Quantifiers		Inter. det.		Total (478)	
Without adjective-like modifiers	54%	(226)	27%	(113)	16%	(66)	2%	(10)	2%	(7)	100%	(422)
Adjectives	40%	(19)	27%	(13)	23%	(11)	10%	(5)	-		100%	(48)
Noun as modifiers	25%	(2)	38%	(3)	25%	(2)	-		13%	(1)	100%	(8)

Table 4.19 Cross-tabulation of determiners and modification above NP in non-bare nominals

	Without det.		Numerals		Sing. dem.		Quantifiers		Inter. det.		Total (478)	
Without modification above NP	4%	(10)	56%	(129)	30%	(70)	6%	(15)	3%	(8)	100%	(232)
Possessives	97%	(228)	-		3%	(8)	-		-		100%	(236)
Relative clauses	90%	(9)	-		10%	(1)	-		-		100%	(10)

With respect to the distribution of modification, both adjectives and noun-modifiers co-occur with three types of determiners. Same as what has been found in the intersection of modification and classifiers, adjectives remain the most used modifiers when co-occurring with other nominal internal elements. For the distribution of determiners in Table 4.18 and Table 4.19 (presented in the columns), singular demonstratives appear with four types of modification, numerals co-occur with two types, while quantifiers and interrogative determiners only interact with one type. An interesting observation is that the total number of tokens of numerals (n=129) is more than that of demonstratives (n=79), but numerals appear with fewer types of modification than demonstratives do. It is grammatical for singular demonstratives and numerals to co-occur with all these types of modification and all of the data show grammatical combinations.

4.5.1.5 Summary and Discussion

This section summarizes and discusses the distributional analysis of the data of the internal nominal structure across age. Among the nominal tokens analyzed, 24% (478/1969) are non-bare nominals. The percentage of non-bare nominals increases when the children grow older: from 0-10% at age two (Nana 2;0-2;3, MLU 1.73-2.72) to 27-37% at age three (Didi 3;1-3;3, MLU 3.21-3.57).

The order of emergence of the nominal internal elements is as follows. The use of adjectives is found at age 2;0 before possessives (2;1) and all other types of modifiers. Classifiers first appear at age 2;2 – the general classifier *ge* and a count-classifier for animals *zhi* both appear in this recording. As for determiners, both numerals and singular demonstratives first appear at age 2;2. The order of emergence of the early elements used within the nominals is adjectives, possessives, and then classifiers/numerals/demonstratives. Except adjectives and possessives, the others all first appear at age 2;2; therefore, a detailed acquisition order cannot be obtained.

Examining all of the nominal internal elements, the study finds that adjectives are used most frequently (n=312). Among non-bare nominals, 51% of them appear with modification above NP, 45% with classifiers, and 48% with determiners – possessives (49%), the general classifier *ge* (37%), numerals (27%), and singular demonstratives (17%) are the elements appearing most frequently. Frequency-wise, the order is the following (from most to least frequent): NP with adjectives (NP with adjuncts), possessives (DP), classifiers (CIP), numerals (NumP), and demonstratives (DP). Adjectives are the first emerged and the most commonly seen elements in all nominals. This is expected since modifications are usually added to the NP structure through adjunction to form a larger NP and no higher phrasal projections are needed (consistent

with Cinque 2005). Assuming Roeper (2006), this would be because only minimal amount of structure is needed to use adjectives.

Although the data do not directly argue for Roeper's proposal (2006) that children build up nominal structures from bottom up (because a detailed acquisition order cannot be found), the finding of the most frequently used Mandarin nominal structure does nicely accord with his bottom-up argument and support another argument of his that D (such as possessives and demonstratives in the data) is one of the defaults of nominals and appears early.

To sum up, the distribution of nominal internal variables – Adjective-like Modifier, Modification above NP, Classifier, and Determiner – and their intersection among each other, are generally consistent with the Mandarin nominal structure proposed in Chang (2007) and Huang et al. (2009), the universal nominal structure in Cinque (2005), and the acquisition path proposed in Roeper (2006).

The finding that adjectives are added first to bare noun roots is predicted given that they appear low in the Mandarin nominal structure, as NP adjuncts. The first phrasal projection added to bare NPs is possessives (age 2;1), and then classifiers, numerals, and demonstratives (age 2;2). Given these two observations, it can be concluded that lexical elements, such as adjectives, are added to bare noun roots before other functional elements (Radford 1990). Besides, according to the frequency of appearance, possessive DP (n=236) and ClP (n=176) are the two most common phrasal projections produced by children, which respectively supports Roeper's DP as one of the defaults and indirectly supports my prediction about ClP. Following Roeper's theory, it is not surprising that NumP (n=129) is the next most commonly seen phrasal projection added to the structure. The frequency order of ClP and then NumP is also consistent with the fact that all NumPs

in the data include a classifier, but not the other way around. The order of emergence of demonstratives and quantifiers is currently found to be D before Q, at age 2;2 and 2;4 respectively.

4.5.2 Multivariate Analyses

4.5.2.1 General result

This section presents the multivariate analyses of the development factors, the semantic factors, and the nominal external factors, including the verbal and clausal domains. Table 4.20 presents the multivariate analyses of the contribution of each factor to the probability of non-bare nominal occurrence. The contribution is presented in terms of factor weight. Factor weight indicates the probability of the occurrence of a non-bare nominal when the factor is present. For factors found to be non-significant by the regression analysis, their factor weight is shown in square brackets. The percentage, taking the ‘subject’ factor under syntactic positions as an example, shows the number of non-bare nominals appearing in subject position out of the total number of subjects – i.e. 27% of the total of 298 subject nominals are non-bare.^{36,37}

³⁶ The format of Table 4.20 is a typical report format for multivariate analysis (Tagliamonte 2006:247-278).

³⁷ More details about some of the terms and data presented in Table 4.20 are as follows.

- (i) Log likelihood is the measure of the goodness of fit of an analysis; figures closer to zero represents better models than those away from zero (Tagliamonte 2006:265). The more variables we put into an analysis, the bigger the log likelihood. A run with only two or three variables may produce a log likelihood of -456.24. A run with six variables might produce a log likelihood of -1456.24 (Tagliamonte 2006:226). Accordingly, given that four significant variables are included in the final model, the number of the log likelihood is reasonable.
- (ii) The range (relative strength) of each statistically significant effect is measured by subtracting the lowest value from the highest (Robinson et al. 2001:27). It is a number and should not appear with a decimal (Tagliamonte 2006:251).

Table 4.20 Multivariate analysis of the contribution of each factor to the probability of non-bare nominal occurrence

Input probability		.22		
Log likelihood ⁱ		-1010.911		
Total N		1969		
	Factor	Weight	%	N
Age	Age: 2 (Nana)	[.45]	24%	155/652
	Age: 3 (Didi)	[.52]	25%	323/1317
MLU	MLU < 2.5	.34	13%	47/368
	MLU 2.5~3.0	.43	18%	133/721
	MLU 3.0~3.5	.62	34%	298/880
	<i>range</i> ⁱⁱ	.28		
Interpretation	Predicative	.40	19%	122/637
	Existential	.58	28%	352/1243
	Generic	.10	3%	2/78
	Unclear	.42	18%	2/11
	<i>range</i>	.48		
Verb Type	Stative	.51	30%	103/345
	Activity	.40	20%	95/479
	Resultative (telic)	.59	37%	67/183
	Copula	.68	29%	73/249
	No verbs	.47	20%	140/713
	<i>range</i>	.28		
Aspect Marker ⁱⁱⁱ	Perfective: <i>le, you, mei, guo</i>	.37	22%	27/122
	Imperfective: <i>zai, zhe</i>	.44	20%	8/40
	Bare verb: no asp. markers	.51	25%	443/1807
	<i>range</i>	.14		
Syntactic Position	Subject	[.50]	27%	80/298
	Object	[.45]	25%	155/619
	Object of preposition	[.51]	31%	35/114
	Post-copular	[.71]	30%	69/229
	Topic	[.49]	30%	9/30
	Appositive of subject	[.57]	25%	1/4
	No verbs	[.47]	19%	129/675

(iii) The tokens of aspect markers are small or have a ‘knock out’ result – for example, the durative *zhe* only appears with bare nominals – thus, the tokens of different aspects have to be combined before running the multivariate analysis and to obtain a more parsimonious model. The tokens of non-bare nominals appearing with aspect makers are as follows: 24 perfective *le*, 1 perfective *you*, 1 negative perfective *mei*, 1 experiential *guo*, 8 progressive *zai*, and 0 durative *zhe*.

Table 4.20 (continued)

Clause Force	Declarative	[.50]	24%	456/1891
	Interrogative	[.48]	29%	21/73
	Imperative	[.40]	20%	1/5

As shown in the top row of Table 4.20, the logistic regression analysis has calculated the ‘input probability’ as 0.22. Input probability is ‘an average frequency of occurrence of the application value of the dependent variable’ (Paolillo 2002:79).

The statistics reveals that four variables – MLU, Interpretation, Verb Type, and Aspect Marker – have significant effects ($p < 0.05$) on the dependent variable. Three variables – Age, Syntactic Position, and Clause Force – do not play a statistically significant role in the bare and non-bare nominal variation.

The factor weights are reported as probabilities between zero and one. Factor weights lingering near .50 indicate that there is little tendency either way for these factors (Tagliamonte 2006:252). A factor weight above .50 can be considered as favoring the occurrence of non-bare nominals, while a factor weight below .50 as disfavoring the occurrence of the non-bare nominals. The factor weights show that the non-bare nominal is unsurprisingly favored in the recordings with longer MLUs (3.0-3.5, weight .62) (and is slightly favored in the child of older age (age: 3, weight .52)). Non-bare nominals are favored with existential interpretation (.58) and disfavored in generic interpretation (.10) as expected. Predicative interpretation (.40) favor bare nominals. Sentences with resultative (telic) verbs (.59), and copulas (.68) favor non-bare nominals, while activity verbs (.40) favor bare nominals. Among the aspect factors, factor weights indicate that bare verbs without aspect markers (.51) slightly favor non-bare nominals, while perfective aspect (.37) and imperfective aspect (.44) favor bare nominals. According to the range (relative strength) of the significant variables as shown in Table 4.20, the

strength of the effect of each variable (from the one with the strongest effect to the one with the weakest) is as follows: Interpretation (48), MLU and Verb Type (28), Aspect Maker (14).

4.5.2.2 Effect of the variables in the verbal domain

4.5.2.2.1 Verbs

Both variables in the verbal domain, verb types and aspect markers, have significant effect on the dependent variable. Copular verb is the verb type that has the highest factor weight (.68), favoring non-bare nominals. The cross-tabulation of post-copular position and different nominal interpretations in Table 4.21 may help shed some light on copula's preference for non-bare nominals.

Table 4.21 Cross-tabulation of post-copular position and interpretations

		Predicative	Existential	Generic	Total
Post-Copular position	Non-bare	63	6	0	69
	Bare	159	1	0	160
	Total	222	7	0	229

Only nominals with predicative or existential definite interpretation are found in the post-copular position. Nominals in the post-copular position, a type of predicate position, can have predicative interpretation, as shown in (48).

(48) a. zhe shi **yi-ge fengche** (Nana 2;4)
 this be one-CL kite
 'This is a kite.'

b. zhe shi **hao piaoliang-de zhuzi** (Nana 2;5)
 this be very pretty-DE bead
 'This is a very beautiful bead.'

As discussed in section 4.2 Linguistic Background, nominals appearing in naming context and predicate positions are interpreted as having predicative interpretation, which mainly concerns property, not individuals.

Example (49) illustrates a post-copular nominal with existential definite interpretation.

- (49) Adult: 'Not your brother?'
 Didi (3;3): shi, shi **wo didi**
 yes, is I younger:brother
 'Yes, it's my younger brother.'

As shown in Table 4.21, among the non-bare nominals appearing in post-copular position, 91% (63/69) have predicative interpretation. Copular verb is strongly related to predicative interpretation – 97% (222/229) of the nominals in the post-copular position have the predicative reading. The co-occurrence of copular verbs and predicative nominals (in post-copular position) may account for the copula's preference for non-bare nominals.

Activity verb is the only verb type that disfavors non-bare nominals. An obvious reason is that activity predicates can be either telic or atelic based on the arguments they take. An activity verb with a bare noun, such as *xie shu* 'write book(s)', is an atelic predicate. An activity verb combined with a non-bare nominal, such as *xie san-ben shu* 'write three books(s)', is telic.³⁸ Therefore, activity verbs are ready to appear with either

³⁸ Mandarin bare nominals can have an existential definite reading given certain context, which results in a telic predicate, as shown below.

- (i) wo zai xie lunwen
 I PROG write dissertation
 'I am writing the dissertation.'

In general, Mandarin activity verbs can combine with a bare noun and a quantified noun phrase to form atelic and telic predicates respectively. A telicity test (Verkuyl 1989), *for ten years* and *within ten years*, can illustrate the difference.

- (ii) a. ta **xie shu** xie-le shi nian (atelic)
 he write book write-PERF ten year
 'He has been writing books for ten years.'

bare or non-bare nominals and should not show any favoritism toward either of them. A closer look at activity verbs and their intersection with other factors will help us better understand the findings.

The cross-tabulation of activity verbs and different interpretations, Table 4.22, shows that existential and generic interpretations are highly associated with activity verbs with bare nominals. The existential nominals occurring with activity verbs appear as bare nominal 78% (324/416) of the time, and the generic nominals occurring with activity verbs appear as bare nominal 96% (49/51) of the time.^{39,40}

-
- b. *ta zai shi nian nei **xie** **shu**
 he in ten year within write book
 '*He write books within ten years.'
- c. ta zai shi nian nei **xie** **san-ben** **shu** (telic)
 he in ten year within write three-CL book
 'He writes three books within ten years.'

³⁹ The predicative nominals appearing in sentences with activity verbs are all produced by the three-year-old, Didi, as illustrated below. These are nominals appearing after *dang* 'work as, act as,' which is an activity verb in Mandarin and can be modified by the progressive marker *zai*. Being able to appear with *zai* is a test to categorize verbs as activity verbs (Chang 2002).

- (i) yao dang xiao yu (Didi 3;0)
 want act:as small fish
 '(I) want to act as a fish.'

⁴⁰ In Table 4.16, the generic nominals with non-bare forms seem puzzling because generic nominals are typically bare nouns in Mandarin. It turns out that 6 of the 8 non-bare generic tokens are modified 'bare' nominals, such as *xiao laohu* 'small tigers' or *da che* 'big car' (as shown in (i)), which are legitimate to express genericity (all these modified generic tokens are produced by Didi). The two-year-old Nana produced the other two non-bare generic tokens, *yi-ge dongxi* 'one thing' and *yi-ge xiaxia* 'one shrimp' (as shown in (ii)), which are erroneous forms for generic expression in Mandarin.

- (i) Context: Didi (3;3): 'That car does not have many seats.'
 Didi (3;3): 'There is one that has three seats.'

Table 4.22 Cross-tabulation of activity verbs and interpretations

		Predicative	Existential	Generic	Total
Activity verb	Non-bare	1	92	2	95
	Bare	11	324	49	384
	Total	12	416	51	479

Most generic nominals tend to appear in the subject position, while existential nominals are in either the subject or object position. Therefore, the variable of syntactic positions should also be considered in order to support the argument that the association between these two interpretations and activity verbs may account for the activity verb's preference for bare nominal. Cross-tabulation of activity verbs and the relevant syntactic positions is presented in Table 4.23 (excluding the objects of prepositions and post-copular position).

Table 4.23 Cross-tabulation of activity verbs and syntactic positions

		Subject	Appositive (subject)	Object	Topic (object)	Total
Activity verb	Non-bare	8	1	74	3	86
	Bare	78	0	268	7	353
	Total	86	1	342	10	439

The result in Table 4.23 shows that 80% (353/439) of the time in either the subject or object position with an activity verb, a bare nominal is found. Concentrating on subjects,

Adult: 'Why only three seats?'

Didi (3;3): yinwei **da** **che** cai you henduo henduo -ge weizi
 because big car only have very:many very:many -CL seat
 'Because only big cars have many many seats.'

(ii) Context: Adult: 'Does Nana like to eat vegetables?'

Nana (2;4): 'No.'

Adult: 'Then, what does Nana like to eat?'

Nana (2;4): wo xihuan chi **yi-ge** **xiaxia**

I like eat one-CL shrimp

'I like to eat shrimps.'

(The correct way to say is to use 'shrimp' without 'one-CL.')

we can see that subject nominals appearing with an activity verb occur in the bare form 91% (78/86) of the time, and object nominals, 78% (268/342) of the time.

Activity verbs favor bare nominals and disfavor non-bare nominals. This can be attributed to at least three facts. First, activity verbs favor bare nominals possibly because they are compatible with atelic interpretations. Second, most (65%) generic nominals appear with activity verbs and generic nominals are typically bare. Last, in sentences with activity verbs, 80% of the subjects and objects appear as bare nominals.

4.5.2.2.2 Aspect Markers

As for the other verbal variable, the aspect marker, its factor weight indicates that bare verbs without aspect markers (.51) slightly favor non-bare nominals, while verbs with perfective aspect markers (.37) and imperfective aspect markers (.44) disfavor non-bare nominals. As discussed before, in general, verbs with perfective or imperfective aspect markers allow more nominals types in their object position than bare verbs do. Hence, aspect markers may prompt the use of non-bare nominals in the object position. Since all the data with aspect markers are acceptable sentences in adult language, the result is somewhat unexpected. However, note that when analyzing the effect of aspect marker on the forms of nominals, the analysis includes all nominals, not just the object nominals that co-occur with aspect markers. Therefore, the cross-tabulation of object position and aspect markers, Table 4.24, should be examined.

Table 4.24 Cross-tabulation of object position and aspect

		Bare Verb	Perfective	Imperfective	Total
Object position	Non-bare	138	11	6	155
	Bare	399	42	23	464
	Total	537	53	29	619

The result reveals that bare nominals appear predominately with bare verbs and with both types of aspect markers. Sentences with perfective markers co-occur with object bare nominals 79% (42/53) of the time. Those with imperfective markers co-occur with object bare nominals 79% (23/29) of the time. The result that the occurrence of aspect markers favor bare nominals does not support the prediction that aspect markers will prompt the appearance of non-bare nominals.

Children's preference for bare nominals when an aspect marker is present is reasonable given that perfective (*le*, *you*, *mei*, *guo*) markers are usually used to comment on hearer-old or discourse-old information, which correlates with definite nominals (Abbott 2004, Prince 1992). Definite nominals may surface as bare nominals in Mandarin. The imperfective (*zai*, *zhe*) markers are usually used to describe the events that are happening 'here and now;' therefore, given the nonlinguistic information in the context, using bare nominals is sufficient to carry on the conversation. This may account for the reason why perfective and imperfective markers favor bare nominals.

What is somewhat unexpected is that bare verbs are more associated with non-bare nominals in the object position than verbs with aspect markers are. Table 4.24 shows that 89% (138/155) of the non-bare objects appear with bare verbs. After examining the non-bare objects that occur with a bare verb (n=138), I find that 45 tokens are non-bare nominals with possessives or relative clauses only, which are acceptable objects of bare verbs. The other 93 tokens have various forms of nominals – with classifiers, numerals, determiners, and/or modification. According to the context, some of them are expected to appear with an aspect marker, but instead appear with a bare verb. The following example illustrates that children may produce non-bare nominals with bare verbs in the context that requires an aspect marker.

(50) Adult: ta chuan-le yi-shuang shenme?
 he put:on-PREF one-pair what
 'He put on a pair of what?'

Nana (2;4): ta chuan yi-ge tuoxie
 he put:on one-CL sandal
 'He put on a pair of sandal.'

Comparing the two sentences in (50), we can see that the perfective *-le* is missing and the classifier is incorrect in the child's sentence, even though *-le* and the correct classifier are used in the adult's question. Considering the following examples for the progressive marker *zai*, we can see that *zai* is used in the adult's question in both scenarios, (51) and (52), while the child only answers with *zai* in (52). The child uses incorrect classifier in both cases – the correct classifier should be *shuang* for (51) and *ba* for (52).

(51) Adult: tamen zai-zuo shenme?
 they PROG-make what
 'What are they making?'

Nana (2;4): tamen zai-zuo yi-ge xiexie
 they PROG-make one-CL shoes
 'They are making a pair of shoes.'

(52) Adult: zhe-ge gonggong zai-na shenme dongxi?
 this-CL old:man PROG-take:out what thing
 'What is the old man taking out?'

Nana (2;4): na yi-ge jiandao
 take:out one-CL scissor
 '(He is) taking out a pair of scissor.'

Although it is not easy to identify in which sentence children clearly omit the aspect marker, based on the examples above, it is certain that some of the non-bare nominals appearing with bare verbs should appear with aspect markers in the adult language. The unexpected association between non-bare nominals and bare verbs may be partly due to the aspect marker omission in child Mandarin.

4.5.2.3 Effect of the variables in the clausal domain

Variables in the clausal domain – Syntactic Position and Clause Force – do not make statistically significant contribution to the variation model of nominal forms.

Although the reasons for the non-significance of clause force is unknown, the results are understandable. Declaratives (factor weight .50) and interrogatives (.48) do not show favoritism toward either nominal forms. Imperatives (factor weight .40) favor bare nominals probably because they often involve specific and known entities, which may be expressed by using bare nominals in Mandarin.

The non-significance of Syntactic Position is puzzling, given that syntactic positions play a crucial role in the interpretation and the structure of nominals, such as the subject-object asymmetry discussed in section 4.2. Unfortunately, as far as I know, there is no research focusing on the relation between syntactic positions and nominal forms in Mandarin. Research about bare and non-bare nominals mostly focuses on the effect of syntactic positions on the interpretation of the nominal (Cheng and Sybesma 1999, Huang et al. 2009). What is known is that both bare and non-bare nominals can appear in all kinds of syntactic positions and can have various interpretations although not without limitations. Syntactic positions have restrictions on interpretation and thus have restrictions on nominal forms as well. For example, Mandarin subject positions select definite or generic nominals, but not indefinites;⁴¹ consequently, only bare nominals or non-bare nominals with definite determiners, such as demonstrative or universal quantifiers, can serve as subjects.

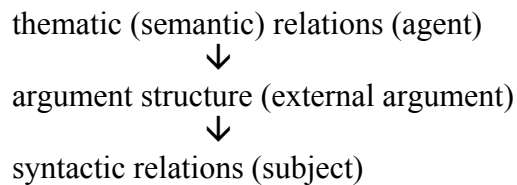
⁴¹ Except the special cases in Li (1998), discussed in footnote 17.

The non-significance of syntactic positions is also found in Liu's case study on the production of demonstratives (Liu 2008). The focus of Liu's study is the acquisition of classifiers. He coded demonstratives because the use of classifiers is obligatory when a demonstrative is present (p.345). Liu's study and the current study are different in that Liu's study coded all utterances with demonstratives and thus skipped bare nouns. The current study coded all utterances that contain nouns and thus left out the cases of bare demonstratives (which may be deictic pronouns or demonstratives with missing nouns). Liu examined the data from a boy (age: 2;2) and analyzed his DP utterances with demonstratives, excluding the bare demonstratives used as pronouns. Although Liu's study presents the data appearing in post-verbal positions, he does not discuss the effect of syntactic position, due to the non-significant effect of pre-verbal and post-verbal positions on the production of DP (Liu p.c.). Both Liu's study of DPs and the current study of nominals find that syntactic positions do not make significant contributions to the various forms of nominal phrases produced by two- and three-year-olds. Possibly the subjects of these two studies were too young (all under age 3;3) and their MLUs were too short (under 3.6).⁴² The short MLU of young children may not provide enough data for syntactic positions to make significant effects. Given the importance of syntactic positions in adult's grammar of Mandarin nominals, more research is needed to examine what their roles are and when they start to actively participate in children's acquisition of nominals.

⁴² The MLU of Liu's subject (Bao-Bao at age 2;2) is 2.6, which is not provided in Liu (2008), but can be found in Liu (2009). The subjects in these two studies are different but Bao-Bao happens to appear in both studies.

In order to map from the lexicon to syntactic positions, children need to know at least thematic roles, argument structure of particular verbs, and syntactic positions, as proposed by Wijnen and Verrips (1998).

(53)(= Wijnen and Verrips 1998 (48))



In addition to the knowledge required to produce nominals in the correct syntactic positions, children also need to know the legitimate forms of nominals appearing in different syntactic positions.

Though syntactic position is a non-significant variable in the current study, some results are still noteworthy. In the prediction section of this chapter, I predicted that the first place for non-bare nominals to appear is the object position because Mandarin object positions allow more types of non-bare nominals than what is allowed in Mandarin subject and predicate positions. The prediction is not supported since Nana used non-bare nominals in both subject and object positions at age 2;0. However, when examining the distribution of non-bare nominals with an identifiable syntactic position, 44% (155/349) of them appear in the object position, while 23% of them in the subject position, and 20%, in the post-copular predicate position. The result shows that non-bare nominals mostly associate with the object position in young children's speech.

4.5.2.4 Effect of the semantic variable

The semantic variable examined in the study, Interpretation, had a significant effect on the dependent variable. The result shows that generic (factor weight .10) and

predicative (.40) interpretation favors bare nominals, while existential interpretation (.58) favors non-bare nominals. The behavior of the generic interpretation factor is predictable since Mandarin generic nominals are typically bare. More details of Mandarin generic nominals will be examined in an experimental study in Chapter 5. A total of 74% (352/478) non-bare nominals have existential interpretation. As predicted, it is the interpretation that most frequently associated with non-bare nominals in child Mandarin.

4.5.2.5 Summary

The study finds that MLU, Interpretation, Verb Type, and Aspect Marker all have a significant effect on the dependent variable, while Syntactic Position and Clause Force do not. According to the factor weights, non-bare nominals are favored in recordings with longer MLUs (3.0-3.5), in existential interpretation, in sentences with copular and resultative verbs, and in sentences with bare verbs without aspect markers. On the contrary, non-bare nominals are disfavored in recordings with shorter MLUs, in generic and predicative interpretations, in sentences with activity verbs, and in sentences with verbs with aspect markers.

4.6 General Discussion and Concluding Remarks

The goal of this chapter was to answer the main research question: what are the contexts that trigger the emergence of non-bare nominals? In other words, what are the factors that determine whether or not to use non-bare nominals?

Among the seven independent variables examined in the multivariate analysis, four are found to have significant effects on the dependent variable, the variation of bare and non-bare nominals: MLU, Interpretation, Verb type, and Aspect Marker. The other

three variables, Age, Syntactic Position, and Clause force, do not have significant effects. However, note that the developmental variables, Age and MLU, simply indicate that children's language production is becoming adult-like when they grow older and they do not help understand the grammar. Therefore, Interpretation, Verb Type, and Aspect Marker are the linguistic variables that significantly shaping the grammar of bare and non-bare nominals in Mandarin acquisition.

The research questions, predictions, and findings are summarized below:

(54) Questions, predictions, and findings:

- a. What are the factors associated with the use (or non-use) of a non-bare nominal?

Finding: Factors associate with non-bare nominals:

- Longer MLU (over 3.0)
- Nominals with existential interpretation
- Sentences with copular or resultative verbs
- Sentences with bare verbs without aspect markers

Factors not associate with non-bare nominals:

- Short MLU (under 3.0)
- Nominals with generic or predicative interpretations
- Sentences with activity verbs
- Sentences with aspect markers

- b. With which element do non-bare nominals first appear beyond the bare nominal stage?

Prediction: (i) Adjective, which occupies the adjunct position of NP in the syntactic tree, is the first element children add to a bare noun root.

(ii) Classifier is the first X^0 element appearing beyond the NP projection.

Finding: Prediction (i) is supported.

As for prediction (ii), possessives first appear at age 2;1 and classifiers, numerals, and demonstratives all first appear at age 2;2. However, the prediction is indirectly supported. Except the possessive DPs, which are arguably part of the default grammar according to Roeper (2006), 37% of all non-bare

nominals contain the general classifier *ge*, while 27% and 17% of them contain a numeral or demonstrative respectively.

- c. In what syntactic position do non-bare nominals first appear?

Prediction: The object position, which allow more types of non-bare nominals to appear than subject and predicate positions do.

Finding: The prediction is not supported. Non-bare nominals appear at age 2;0 in subject, object, and predicate positions. But a total of 44% non-bare nominals with an identifiable syntactic position appear as objects.

- d. With which interpretation are non-bare nominals most frequently associated?

Prediction: The existential interpretation.

Finding: The prediction is supported. A total of 74% non-bare nominals have existential interpretation.

In addition to the findings pertaining to various variables, one may wonder if, overall, there is a sentence construction that requires non-bare nominals. And, if so, in what structural context do non-bare nominals first appear? So far, no known construction has been found to require non-bare nominals. The study results show that early non-bare nominals tend to appear as possessive nominals or nominals with a classifier in sentences with copular or resultative verbs without aspect markers, and they have existential interpretation. The research finding supports my proposal that if child speech reflects the target language, Mandarin-speaking children will systematically produce nominals with particular structures, in particular constructions, and with particular interpretations.

The data presented in this chapter also generally validate the previous research about structure and acquisition of noun phrases: Borer (2005), Chang (2007), Cinque (2005), and Roeper (2006). The current research has indicated that at age three, children's speech have reflected some but not all the constraints governing the variable

rules of bare and non-bare nominals in Mandarin. For example, the rather unexpected results concerning aspect markers ask for further scrutiny.

This is a pioneer and general study about the variables that affect the variation of nominal forms, focusing on the emergence of non-bare nominals in young children (age: 2-3). It is desirable to have more research studying each variable in greater details and with children of different ages.

Finally, among the different interpretations, existential interpretation favors non-bare nominals, while generic and predicative interpretations favor bare nominals. Given the significant effect of the interpretation variable on children's production of bare and non-bare nominals, the next question to ask is: what is the initial interpretation that young children assign to the bare and non-bare nominals? The next chapter attempts to answer this question.

CHAPTER FIVE

COMPREHENSION OF MANDARIN BARE AND DEMONSTRATIVE NOUN PHRASES

This chapter presents the research regarding the second challenge with a focus on the nominals in the subject position.

- Challenge 2. a. How do Mandarin-speaking children interpret bare nominals given that bare nominals can have various interpretations?
b. In contrast, what is the interpretation that children assign to non-bare nominals, such as demonstrative nominals?
c. Do children have the same interpretation as adults do?

In Mandarin, bare nominals can have generic reading; while non-bare nominals, such as nominals with numerals or demonstratives, generally cannot. In Chapter 4, the study of two children's spontaneous speech data has shown that young children express genericity by using bare nominals, such as *shuiguo* 'fruit' or *ningmeng dangao* 'lemon cake' (Didi 2;9). Adults express genericity in the same way.

This chapter presents an experimental study to investigate how children interpret bare and demonstrative nominals in Mandarin. This study is controlled in such a way that only the generic and existential definite readings of the bare nominals are possible. Mandarin lacks definite articles. In order to compare the existential definite reading of bare nominals, demonstrative nominals – a type of definite expressions – are used in the experiment. This study also examines whether children only interpret bare nominals as generic and never interpret demonstrative nominals as generic, as predicted by the adult Mandarin grammar.

5.1 Introduction

GENERICITY is related to the properties associated with kinds (*Lions eat meat*) and the generalizations about entities (*My cat does not eat canned fish*). It is conveyed in both the nominal and the verbal domain. At least two developmental challenges are related to genericity: conceptual and linguistic. Researchers have argued that language helps children learn the conceptual side of genericity – language plays an important role in acquiring generic knowledge (Gelman and Raman 2003, Pappas and Gelman 1998, Prasada 2000, to name a few).⁴³ From the linguistic perspective, generics present a noteworthy acquisition problem – they are not usually unambiguously presented in specific linguistic constructions, but they can be expressed in a variety of structures that may also have non-generic meanings.

There are two nominal constructions in Mandarin that can have a generic interpretation. Bare nominals (such as *ren* ‘human’) are the most common form of generic nominals in Mandarin, while singular nominals (such as *yi-ge ren* (one-classifier human) ‘a human being’) have a very restricted use as generic expressions. Therefore, this structure will not be discussed in this study.⁴⁴

⁴³ For the acquisition challenges of generics in the conceptual part, please see Prasada 2000.

⁴⁴ In addition to bare nominals, Mandarin *yi*+classifier+nouns, the most comparable form to English nominals with an indefinite article, may also be used as kind-referring nominals, as shown below.

- (i) a. **yi-ke** **putaoyou** hanyou henduo weitaming C (generic)
 one-CL grapefruit contain much vitamin C
 ‘A grapefruit contains lots of vitamin C.’

However, the nouns that may be used in the *yi*+classifier+noun construction to express genericity are very limited, e.g. animate nominals usually cannot appear in this

Mandarin bare nominals can occur in various positions, such as subjects, objects, or predicates, and can receive generic, predicative, or existential (indefinite or definite) interpretations. Which interpretation they receive depends on their syntactic position and the predicates they appear with. In contrast to bare nominals, which can have either generic or existential interpretation in argument positions, the demonstrative nominals (there is no definite article in Mandarin) receive almost always the existential definite interpretation.⁴⁵

The goal of this chapter is to examine how children interpret bare and demonstrative nominals in Mandarin – specifically, whether the interpretation that young children assign to bare and demonstrative nominals coincides with how adults interpret them or diverges from these interpretations in a systematic way in particular contexts. I will limit my study to bare and demonstrative nominals in the subject position with a characterizing statement and focus on two major questions – what is Mandarin-speaking children's interpretation of bare and demonstrative nominals, and what are the factors affecting their interpretation of nominals? I will explore these issues by revising and

construction to express genericity, as shown in (iia-b). This structure is excluded in the study.

- (ii) a. ***yi-zhi laohu** zhu zai conglin li
 one-CL tiger live PROG jungle in
 ‘A tiger lives in the jungle.’
- b. ***yi-zhi mao** zhu zai fangzi li huo jie shang
 one-CL cat live PROG house in or street on
 ‘A cat lives either in a house or on the street.’

⁴⁵ More details about the readings for demonstratives nominals in Mandarin will be discussed in Section 5.2.2 Generic nominals.

implementing a comprehension experiment that was used for testing English and Spanish generics (Pérez-Leroux, Munn, Schmitt, and DeIrish 2004).

The organization of this chapter is as follows. Sections 5.2 and 5.3 summarize the linguistic characteristics of various types of generic expressions and the acquisition background of bare, definite, and demonstrative nominals. The research questions and hypotheses will be discussed in section 5.4. Section 5.5 will present the details of the experiment and the analyses of the results. The last section will consist of the conclusion and discussion.

5.2 Linguistic Background

5.2.1 Linguistic Characteristics of Generic Expressions

There are two types of generics (Krifka et al., 1995:2-5). The first type is in reference to a kind. In this type, genericity is a feature of the nominals (nominal domain). Examples in (1) do not denote some particular potato or group of potatoes, but rather the kind, Potato (*Solanum tuberosum*) ((1) = Krifka et al. 1995: (1)).

- (1) a. **The potato** was first cultivated in South America.
- b. **Potatoes** were introduced into Ireland by the end of the 17th century.
- c. The Irish economy became dependent upon **the potato**.

The second type is characterizing statements. Genericity of this type is a feature of the entire sentence (verbal domain). This type of generic sentence does not express a specific event or an isolated fact, but conveys the general property of a kind (2a) or an entity (2b).

- (2) a. **An orange** contains vitamin C.
- b. **Bill** eats an apple every day.

However, not any noun can form a kind-referring nominal (Krifka et al. 1995:11).

The definite nominal in subject position in (3a) can refer to a kind, while (3b) cannot,

because there exists a ‘well-established kind’ for Coke bottles, but not for green bottles

((3) = Krifka et al.: (24)).

- (3) a. **The Coke bottle** has a narrow neck.
- b. ??**The green bottle** has a narrow neck.

Not all predicates can be interpreted as generic. Sentence (4a) has a generic reading, while (4b) has an existential reading.

- (4) a. Dogs are men’s best friends. (generic, ILP)
- b. Dogs are eating canned food in the yard now. (existential, SLP)

The difference between (4a) and (4b) is that the former has an INDIVIDUAL-LEVEL PREDICATE (ILP), which expresses a permanent property of the subject, while the latter has a STAGE-LEVEL PREDICATE (SLP), which expresses a temporary property of the subject. ILPs are inherently generic predicates (Carlson 1977, Chierchia 1995).

However, SLPs can be used to convey habitual reading, which is a type of generic interpretation as well, as shown in (5).⁴⁶

- (5) **Dogs** are eating canned food these days. (generic)

In addition, tense also plays a role. English present tense forces a habitual/generic interpretation in eventive predicates, as in (6a). In contrast, Spanish present tense sentence (6b) is ambiguous between a generic and a here and now reading. Past tense in English favors the existential reading (7a), but the generic reading is also possible under special circumstances – for example, as in the *used to* construction (7b) (Krifka et al. 1995:7), or when discussing extinct species (7c) in generic contexts.

- (6) a. She eats apples.
- b. Come manzanas.
 eat-3rd.Sing. apple-PL
 ‘S/he eats/is eating apples.’

⁴⁶ For more discussion about ILP and SLP, see Carlson (1977, 1980) and Kratzer (1995).

- (7) a. **Dogs** ate vegetables. (usually existential)
 b. **Dogs** used to eat vegetables in those days. (generic)
 c. **Sabre-toothed tigers** ate meat. (generic acceptable)

There is no tense morphology in Mandarin. The predicates with bare verbs allow a generic reading, as in (8a, b). Mandarin verbs that appear with an aspect marker, such as the perfective marker *le* or the progressive marker *zai*, will force an existential reading (8c, d). Although bare subject nominals appearing with characterizing statements in the bare verb form can be interpreted as generic or existential (8a), the bare verb form is the only possible construction for generic sentences in Mandarin, even for the past property of a kind (8e) or for extinct species (8f).

- (8) a. **gou** (meitian) chi bingqilin
 dog (every day) eat ice cream
 ‘Dogs eat ice cream (every day).’ (generic)
 ‘The dog eats ice cream (every day).’ (existential or generic)
 ‘The dogs eat ice cream (every day).’ (existential)
- b. xianzai **gou** dou chi bingqilin
 nowadays dog all eat ice cream (generic only)
 ‘Nowadays dogs eat ice cream.’
- c. **gou** chi-le bingqilin
 dog eat-PERF ice cream (existential)
 ‘Dogs ate ice cream.’
- d. **gou** zai-chi bingqilin
 dog PROG-eat ice cream (existential)
 ‘The dog is eating ice cream.’
 ‘The dogs are eating ice cream.’
- e. **gou** congqian shi yeshengde
 dog past be wild (generic)
 ‘Dogs were wild in the past.’
 (meaning: ‘Dogs used to be wild.’)
- f. **baolong** zhi chi rou
 T-rex only eat meat (generic)
 ‘T-rex only ate meat.’

In sum, genericity results from the interaction between elements from the nominal and verbal domain. This chapter will only discuss characterizing statements with either present tense in English and Spanish or bare verb forms in Mandarin to make nominal types (bare or demonstrative) the focus of the study.⁴⁷

5.2.2 Generic Noun Phrases

In different languages, different structures are used to create kind-referring nominals. English bare plurals, definite singular and indefinite singular can be interpreted as generic, while in Mandarin, generic expressions generally appear as bare nominals. The first part of this section will discuss the structure of generic nominals (focusing on count nouns) in the subject position of characterizing statements in English and Mandarin with some comparison with Spanish examples. The second part will briefly discuss Chierchia's (1998) Nominal Mapping Parameter, which may account for the form-meaning differences of nominals in these languages.

English bare nominals (mass or bare plural count nouns +/- modifiers) and in/definite nominals (*a/an/the* + singular count noun) are all able to express genericity when appearing with a characterizing statement, as illustrated in (9) – (11a).

- (9) a. Dogs eat meat. (generic)
b. *Dog eats meat.
- (10) A grapefruit contains a lot of vitamin C. (generic)
- (11) a. The hummingbird gets its energy from sugar. (generic or existential, ILP)
b. The hummingbird drinks from the bird feeder in my yard. (existential, SLP)
c. The hummingbirds do not drink a lot. (existential)

⁴⁷ Spanish data are introduced in that Spanish definite plurals, like Mandarin bare nominals, may be ambiguous between generic and existential definite readings. Since English bare plurals and definite plurals do not have ambiguous readings, Spanish data are needed to compare with Mandarin data.

In characterizing statements, bare count nouns in subject position must be plural (9a) to be acceptable in English. Bare singular count nouns in subject position will result in an unacceptable sentence (9b). An indefinite singular count nominal can also express genericity, as in (10). English definite singular nominals may be ambiguous between generic and existential readings, given an individual-level predicate, as in (11a), while only the existential reading is possible for those appearing with stage-level predicates, as in (11b). On the contrary, definite plural nouns (11c) always refer to specific objects and cannot be kind-referring.

In Mandarin, genericity is generally expressed by using bare nominals, as shown in (12).

- (12) **gou** chi rou (generic or existential)
 dog eat meat
 ‘Dogs eat meat.’
 ‘The dogs eat meat.’
 ‘The dog eats meat.’

A present-tense characterizing statement always forces English bare plurals in subject position to receive a generic/habitual reading. In contrast, Mandarin subject bare nominals can be interpreted as either generic or existential (Cheng and Sybesma 1999), given a characterizing statement without any aspect markers. Moreover, Mandarin does not have plural morphology, and bare nominals can be interpreted as either singular or plural.⁴⁸ Thus, the bare count noun in (12) can be translated in English as a bare plural, a definite plural, or a definite singular. In order to compare Mandarin bare nominals with

⁴⁸ *-Men* is arguably a definite plural marker in Mandarin; however, it can only be used as the suffix of nouns denoting humans, such as *laoshi-men* ‘teachers,’ and not with other common nouns, such as **zuozi-men* ‘tables.’ For more information about *-men*, please see Li (1999) and Munn, Zhang, and Schmitt (2009).

English bare plurals, the experiment in this study is designed in such a way that only the plural reading of Mandarin bare nominals will be felicitous.

As for definite nominals, which can be interpreted generically in some languages, Mandarin does not have definite articles. In Mandarin, the determiner that is the closest to definite articles is the demonstrative, as has been noted in the literature (S. Huang 1999, Tao 1999). Regardless of their differences, demonstrative and definite nominals are similar in that they are both definite expressions and are presupposed to denote uniquely identifiable entities (Gundel et al. 1993, 2001, cited in Abbott 2010). Besides, studies have revealed that Mandarin demonstratives are developing into definite articles although they are not fully grammaticalized into definite articles yet (Chen 2004, S. Huang 1999). In the current study, Mandarin demonstrative nominals are used as controls – they only have the existential definite reading in the experiment.

Unlike Mandarin demonstratives, English demonstratives are able to convey a generic reading, which refers to certain type of objects, as shown in (13). However, both the plural and singular demonstrative nominals in Mandarin can only receive existential definite reading, as in (14a-b).

- (13) a. **These cars** are very sturdy. (existential or type reading ‘this type of car’)
b. **This car** is very sturdy. (existential or type reading ‘this type of car’)

- (14) a. **zhexie che** hen naiyong
these car very sturdy
‘These cars are very sturdy.’ (existential)
- b. **zhe-liang che** hen naiyong
this CL car very sturdy
‘This car is very sturdy.’ (existential)
- c. **zhe-zhong che** hen naiyong
this type car very sturdy
‘This type of car is very sturdy.’ (type reading)

- d. **zhe che** hen naiyong
 this car very sturdy
 ‘This car is very sturdy.’ (existential)
 ‘This type of car is very sturdy.’ (type reading – very marginal)

In Mandarin, the word *zhong* ‘type’ must appear to express the generic ‘type’ reading, as shown in (14c). Singular demonstrative nominals appearing without classifiers may be ambiguous between existential and type readings as in (14d). However, native speakers accepting the type reading of (14d) indicate that such a reading is very marginal and they would prefer to use (14c) to express ‘types of objects.’ Although the capability of *zhe che* ‘this car’ in (14d) to express the type reading is unclear, it is possible that it may be the abbreviated form of either *zhe liang che* (14b) or *zhe zhong che* (14c) and thus is capable of conveying the meaning of these two forms.

In contrast to its English and Mandarin counterparts, Spanish bare nominals cannot be interpreted as generic, and definite articles are required to construct generic nominals as in (15) (Contreras 1986, Pérez-Leroux et al. 2004). Spanish definite nominal subjects, similar to Mandarin bare nominal subjects, can have generic or existential readings depending on the context.

- (15) ***(Los) tigres** comen carne. (generic or existential)
 ‘The tigers eat meat.’

Like English demonstratives, Spanish demonstratives may convey the existential reading or the ‘type’ reading, as illustrated in (16).

- (16) a. **Este coche** es muy firme. (existential or type reading ‘this type of car’)
 ‘This car is very sturdy.’
 b. **Estos coches** son muy firmes. (existential or type reading ‘this type of car’)
 ‘These cars are very sturdy.’

As already mentioned in this section, the current study will focus on plural readings to avoid the ambiguity of Mandarin nominals, which can have plural or singular

readings. Also, the Mandarin demonstrative will be used as a parallel to the definite in other languages. Based on these requisites, summarized below are the semantic mappings of bare, definite, and demonstrative (plural) subject nominals in sentences with characterizing statements in English, Spanish, and Mandarin.

Table 5.1 Interpretation of subject nominals with characterizing statements in English, Spanish, and Mandarin⁴⁹

	English	Spanish	Mandarin
Bare (plural)	generic	*	generic/ existential definite
Definite (plural)	existential definite	generic/ existential definite	*
Demonstrative (plural)	type/ existential definite	type/ existential definite	existential definite

One possible explanation for the distinct form-meaning distributions in these languages lies in Chierchia's (1998) Nominal Mapping Parameter. As discussed in Chapter 2, Chierchia proposes two features [\pm arg(ument)] and [\pm pred(icate)], which can provide a classification of nouns in different languages.

- (17) a. [+arg, -pred] e.g. Mandarin
- generalized bare arguments
 - the extension of all nouns is mass
 - no plural morphology
 - generalized classifier system
- b. [-arg, +pred] e.g. French, Spanish
- no bare nouns in argument position
 - mass/count distinction
 - plural morphology

⁴⁹ In the English, Spanish, and Mandarin experiments discussed in this chapter, given the plural context and the experimental stories, 'the existential definite reading' refers to an existential and anaphorically definite reading. For the ease of discussion and to differentiate it from the definite reading of the regular definite-articled nominals, I will refer it as 'the existential definite reading' in this chapter.

- c. [+arg, +pred] e.g. English
- bare mass nouns can be arguments;
 - singular count nouns cannot
 - mass/count distinction
 - ([+arg] N: mass, [+pred] N: count)
 - plurals can be arguments
 - plural morphology

- d. [-arg, -pred] non-existent

English is a [+arg, +pred] language. English mass nouns are [+arg] and thus can appear as bare arguments. Count nouns are [+pred], but given the available type-shifting in English, count plurals can be mapped into arguments. As discussed in Chapter 2 and 3, Chierchia (following Carlson 1977) argues that bare arguments are kind-referring by default; therefore, English bare mass nouns and bare count plurals can function as generic expressions. However, why cannot definite plurals be kind-referring in English, just like in Spanish (as shown in Table 5.1)? Chierchia argues that a principle of economy, Avoid Structure, will force choosing generic bare plurals over generic definite plurals in English. Avoid Structure says that when a bare nominal and a determiner+nominal have the same meaning (e.g., when both *tigers* and *the tigers* can potentially denote the kind ‘tigers’), the simpler structure will be chosen. In English, the generic reading of definite plurals is blocked by the bare plural, as summarized in Table 5.1.

In contrast, Romance languages are [-arg, +pred]. In this type of language, nouns always denote predicates and bare nouns are prohibited in argument position; therefore, the absence of the bare plurals allows the definite plurals to include both the generic and the definite reading, as presented in the Spanish example in (15) and Table 5.1.

Chierchia (1998: 392) suggests that Romance definite plurals readily admit generic reading in that definite articles will have a universal reading when appearing with plural or mass nouns in generic sentences or sentences with kind-level predicates.

In Chierchia's theory, Mandarin is a [+arg, -pred] language, allowing bare nouns to appear freely in argument positions. He suggests that all Mandarin nouns start out as kind denoting (1998: 401). This easily accounts for the fact that Mandarin bare nominals have generic reference. The type-shifting operation and blocking principle can account for Mandarin bare nominal's capability of obtaining various readings, as discussed in Chapter 2.

The crucial point is that, regarding generic expressions in the subject position, English bare plurals can receive a generic reading, but never an existential definite reading, while Mandarin bare nominals, as well as Spanish definite plurals, can be interpreted as generic or existential definite. As for definite expressions, English definite plurals and their closest counterpart in Mandarin, demonstrative plurals, only have existential definite readings (see Table 5.1). This chapter will examine nominal interpretations by studying Mandarin bare nominals in contrast to demonstrative nominals.

5.3 Acquisition Background

In the previous section I discussed how languages differ in the form-meaning mapping of generic expressions. It is interesting to examine how bare, definite, and demonstrative nominals are interpreted cross-linguistically. As far as I know, there are no previous acquisition studies about the comprehension of bare and demonstrative nominals in child Mandarin.⁵⁰ This section discusses some background of the production and comprehension of generics and definites. Section 5.3.1 presents an

⁵⁰ A study about genericity in Mandarin is Gelman and Tardif (1998), which collected free speech production data of English- and Mandarin-speaking adults and children, but only adult speech data are analyzed in the paper.

overview of the acquisition of generics and definites, followed in Section 5.3.2 by an interpretation study of bare and definite nominals in English (Gelman and Raman 2003). Section 5.3.3 discusses a cross-linguistic study about the interpretation of bare, definite, and demonstrative nominals in English and Spanish (Pérez-Leroux et al. 2004). The methodology used in that study will be adopted in the current study. Acquisition and theoretical studies that set the stage for the hypotheses and predictions in this study will be discussed in Section 5.4.

5.3.1 Acquisition of Generics and Definites

The acquisition of generic nominals is challenging (Prasada 2000). Besides the challenge posed by the complexity of generic syntax and semantics, there are also cognitive challenges. First, generic sentences are exception-tolerating — it is acceptable to say that *candies are sweet* although some special flavored candies are sour. Another cognitive challenge is that generic knowledge is acquired by experiencing only a limited number of cases, sometimes even just one case. A child may believe that *spiders are scary* even though he has encountered only one spider up to that point in his life.

Regardless of these challenges, various studies have found that children produce and comprehend utterances about natural kinds very early. Children's utterances about biological kinds show systematic beliefs about the nature of being a member of a kind (Keil 1989, Pappas and Gelman 1998, Gelman and Tardif 1998, Gelman and Raman 2003, to name a few). Examples below shows that a Mandarin-speaking two-year-old is able to combine a bare subject nominal and a characterizing statement to form generic sentences.

(18) Context: The adult and Didi were reading a book about animals. Didi talked about many animals according to his opinion. Some of the animals were not present in the book and some of his statements are not true. (Didi 2;10.23, Chang corpora 2002)

Didi: laohu gan youyong
tiger dare swim
'Tigers dare to swim.'

Adult: laohu gan youyong o
tiger dare swim oh
'Oh! Tigers dare to swim!'

Didi: dui xiao laohu gan youyong
right small tiger dare swim
'Right! Small tigers dare to swim.'

Adult: ni fan-guoqu kan xia yi-ge shi shenme hao-bu-hao
you turn-over read next one-CL be what good-not-good
'You turn to the next page to see what is the next, ok?'

Adult: maomaochong
caterpillar
'Caterpillars!'

Adult: maomaochong hui bian shenme
caterpillar will become what
'What will caterpillars become?'

Didi: maomaochong ye hui youyong
caterpillar also can swim
'Caterpillars can swim, too.'

Adult: maomaochong ye hui youyong a
caterpillar also can swim A
'Caterpillars can swim, really?'

Didi: hudie ye hui youyong
butterfly also can swim
'Butterflies can swim, too.'

Didi kept mentioning many animals that can or cannot swim, dare or not dare swim: snails, tigers, spiders, crocodiles, camels, mice, foxes, rhinos, flamingos. Then, he suddenly switched the topic to the animals that can walk.)

Didi: kangaroo ye shiyong zou de
 kangaroo ye be use walk DE
'Kangaroos walk, too.'⁵¹

Didi: daxiang ye shi yong zou de
 elephant also be use walk DE
'Elephants walk, too.'

Adult: deng yi-xia kangaroo bu shi yong zou de
 wait one-moment kangaroo not be use walk DE

 kangaroo yong tiao de .
 kangaroo use jump DE

'Wait a second. Kangaroos do not *walk*. Kangaroos *jump*.'

Didi: tiao de
 jump DE
'Jump'

Given that Mandarin subject bare nouns are ambiguous between generic and existential readings, it is the context that prompts the listener to interpret the sentences as generic. In (18), the adult and Didi both use generic statements to respond to each other.

As for definite noun phrases, children also produce nominals with definite articles at a relatively young age. Based on the Brown (1973) corpora of three young children, English-speaking children use definite articles around age 1;6 to 2;7, as shown in (19a-c). There are no explicit articles in Mandarin; however, demonstratives like *zhe* 'this' and *na* 'that' can be found as early as at age two, as shown in (19d-e).

- (19) The emergence of English articles: *the* (Brown corpora, 1973:271)
- a. Adam write the paper (Adam 2;3.4, line 2414)
 - b. read the puzzle (Eve 1;6, line 621)
 - c. in the house ? (Sarah 2;7.5, line 1135)

⁵¹ *Shi...de* is an emphatic construction. This is the first-mention of *kangaroo* and *zou* 'walk' in the discourse. Thus, it is infelicitous to use *ye* 'also' here.

The emergence of Mandarin demonstratives: *zhe* ‘this’, *na* ‘that’
(Chang corpora, 2002)

d. mama wo yao zhe yizi (Nana 2;2.21, line 1087)
mom I want this chair

e. zai na difang (Didi 2;10.23, line 768)
at that place

Despite the early emergence of definites, children do not use them in the adult way. At least two problems can be found in children’s use of definite articles. First, many studies find that preschool children overuse the definites in contexts that do not meet the uniqueness requirement of definites (Karmiloff-Smith 1979 (French definites), Maratsos 1976, Wexler 2003, to name a few). For instance, in one of Maratsos’s experiments where several identical objects are present, over half of the four-year-olds mistakenly use *the*, while *a* was expected. Second, children treat definites as non-referential or generic – comparable to definites in Spanish, as discussed in section 5.2 (de Villiers and Roeper 1995, Pérez-Leroux et al. 2004). It is as if children underspecify definites, which allows the generic interpretation as in English and Spanish. Or as de Villiers and Roeper (1995:73, 101) suggest – children start out with the default assumption that all nominals are NPs only, and D is absent in early child grammar due to Economy of Projection and thus they would treat definites as non-referential, like expletives.

In summary, generic and definite nominals appear at an early age in child language, but the full grammar of definites has not been acquired yet at this point. Also, definites may be misused in inappropriate contexts or misinterpreted as non-referential.

5.3.2 Gelman and Raman (2003)

In comparing bare nominals with definite nominals, Gelman and Raman (2003) conducted a study on the interpretation of English bare and definite plurals. They tested two- to four-year-olds on their ability to use determiners as a cue to nominal interpretation. The children would be shown a picture of two penguins and asked either a question with a bare subject (*Do birds fly?*) or a question with a definite subject (*Do the birds fly?*). The data reveal that the children appropriately differentiate between the generic and definite questions most of the time, interpreting bare plurals as generic and definite plurals as existential. However, in their experiments the children did exhibit errors, which occurred in both directions: definite questions received generic responses between 5% and 10% of the time, and generic questions received existential responses between 16% and 27% of the time. As stated by Gelman and Raman, the existential reading is the more salient one in this task, since it is supported by a picture of the specific referents. The saliency of the existential reading in comparison to the generic reading may explain why the children were more accurate answering definite questions than answering generic questions. Based on the results, they conclude that young children have begun to distinguish between generic (bare plurals) and non-generic phrases (definite plurals) from a very young age.

Although as treated in Gelman and Raman's study the generic responses for English definite plurals are errors, in the next section we will see that the generic reading for definite nominals is actually allowed in child English.

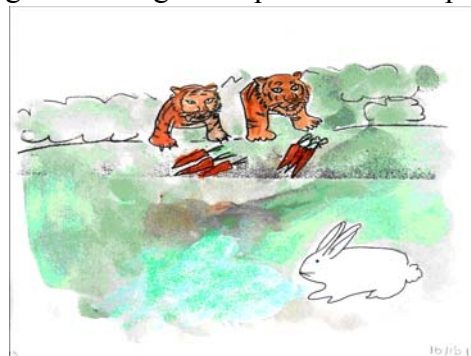
5.3.3 Pérez-Leroux, Munn, Schmitt, and DeIrish (2004)

Pérez-Leroux, Munn, Schmitt, and DeIrish (2004) designed a comprehension task to examine children's interpretation of definite plurals, in contrast to bare plurals, in English and Spanish. Based on Chierchia (1998), the definite article is semantically the same in both languages – definite plurals should be able to have generic or existential definite interpretations. However, in English the bare plural blocks the definite plural generic. Adopting Chierchia, they hypothesized that if English-speaking children have problems deciding which form can be interpreted as generic, they would allow definite plurals to have a generic interpretation. Also, the percentage of definite plurals being interpreted as generic was expected to be lower for English-speaking children than for Spanish-speaking children, given that this is a grammatical interpretation in Spanish.

In the comprehension task designed by Pérez-Leroux et al., eight stories were created, each containing two atypical members of a kind (e.g. two vegetarian tigers) and one observer, member of another kind (e.g. a rabbit). A sample set of experiment item is illustrated below.

Figure 5.1 English experiment example

Fredi the tiger and Pepi the tiger only eat vegetables. Look, they're eating carrots. The rabbit is happy because he won't get eaten. Now let me ask you some questions



- | | | |
|---------------------------------------|-------------------------------------|---------------------|
| (20) a. Immediate question: | Do the tigers eat meat? | (Yes: Gen, No: Def) |
| b. Positive distracter: | Do the tigers have stripes? | |
| c. Negative distracter: | Do you have stripes? | |
| d. Delayed question: | Do tigers eat carrots? | (Yes: Def, No: Gen) |
| e. Demonstrative follow-up questions: | | |
| How about those tigers? | Do those tigers eat carrots? | (Yes: Def, No: Gen) |

Subjects were presented with yes/no questions about the atypical characters, as in (20). Each story was followed by four questions: one immediate question, two distracters (one positive, one negative) and a delayed question. The delayed question was introduced to evaluate whether a pragmatic factor (distance from a discourse antecedent) plays a role in nominal interpretation: after listening to a story that introduces two atypical animals, a higher percentage of existential definite readings for the immediate question was expected, compared to the delayed question. If a child assigned the incorrect generic interpretation to definites, then a demonstrative question was asked right after the delayed question. The demonstrative question was introduced to see if children choose the generic interpretation for definites because of linguistic or non-linguistic preference. If the error was due to a non-linguistic factor, children should also choose the incorrect generic reading for demonstratives. If children choose the interpretation based on the nominal types, it is expected that they will interpret definites and demonstratives differently and only assign the existential definite reading to demonstratives.

The answers to the questions indicate the subject's interpretation of the target noun phrase. Affirmative answers to questions about the canonical property of the kind (yes to '*tigers eat meat*' (20a)) and negative answers to questions about the non-canonical property (no to '*tigers eat carrots*' (20de)) would indicate acceptance of a generic reading of a noun phrase.⁵²

⁵² As mentioned in Section 5.2.2, generic 'type' reading is possible for English and Spanish demonstratives, but not for Mandarin demonstratives in general. If the generic 'type' reading of English and Spanish demonstratives was counted in the experiment results, will it result in high percentage of generic reading of demonstratives? Since the 'type' reading was not counted in the result; therefore, it is not a concern here. In the English and Spanish studies of Pérez-Leroux et al. and the current Mandarin study, the

The result of the first experiment of Pérez-Leroux et al. (2004), in which all sentences were in the present tense, shows that English bare plurals received high rates of generic readings in all age groups including children and adults (~80-95%) (bare plurals are ungrammatical in Spanish and were not tested in the study). On the contrary, the result for definite determiners is much more unexpected: English-speaking children assigned generic reading to definites ~70% and to demonstratives 13% of the time, while adults provided almost no generic reading for definites (and hence, no demonstrative follow-up questions for adults). The different semantics assigned to definites and demonstratives by English-speaking children illustrate that they interpreted the two determiners differently based on their linguistic knowledge about them. Spanish-speaking children assigned generic reading to definites ~80-95% (older and younger groups) and to demonstratives ~17-20% (older group) and ~18-42% (younger group) of the time. There is no adult data available. There was no effect of presentation order (immediate- or delayed mention) or a non-linguistic property (canonicity '*tigers eat meat*' or non-canonicity '*tigers eat carrots*') found in the English results. In the Spanish

generic reading is defined as accepting a canonical sentence like (i) and rejecting a non-canonical sentence like (ii), as presented in bold fonts below. Therefore, the 'type' reading is excluded.

- | | |
|-----------------------------------|--|
| (i) Do these tigers eat meat? | <p>Yes: generic (true to the kind 'tiger' in the real world)</p> <p>No: definite (true to the tigers in the context)
generic (true to the special subtype of tigers in the context)</p> |
| (ii) Do these tigers eat carrots? | <p>Yes: definite (true to the tigers in the context) or generic (true to the special subtype of tigers in the context)</p> <p>No: generic (true to the kind 'tiger' in the real world)</p> |

results, there is also no effect of canonicity, but there is significant effect of presentation order.

As discussed earlier, tense plays a role in genericity. Since English present tense may force a generic/habitual interpretation and Spanish present tense is ambiguous between a generic/habitual and a progressive reading, Pérez-Leroux et al. conducted another experiment to examine the effect of tense, either present or past, on definite noun phrase interpretation. The results show overall lower proportion of generic responses for definites cross-linguistically than that in the present-tense only experiment. There is no significant effect of past vs. present tense on the interpretation of definite noun phrases in English. Regardless of tense, English-speaking children assigned generic reading to definites 40-50% of the time, whereas adults did not interpret definites as generic. On the contrary, in the Spanish study, there is a major reduction in children's generic reading of the definite plurals in past tense sentences (40-50%) in comparison to that in present tense sentences (60-70%). The effect of tense is significant in interpreting Spanish definite plurals.

Based on the findings, Pérez-Leroux et al. propose that the kind-referring reading is one of the basic readings for definites across languages in the given context, and English-speaking children, like Spanish-speaking children, have a grammar of definites that allows the possibility of generic readings.

5.4 Research Questions and Hypotheses

As discussed previously (Chapter 1 and Section 5.2.1), due to the lack of in/definite article and plural morphology, Mandarin bare nominals are ambiguous between generic and existential definite (singular or plural) readings in the subject

position. Therefore, it is reasonable to contrast the interpretation of bare nominals with that of another type of nominal that allows only one interpretation. Mandarin demonstrative nominals, which allow only the existential definite reading, will be examined in the study in contrast to bare nominals. This section presents the research questions, hypotheses, and predictions.

The main research questions of this chapter are as follows.

- (21) What is Mandarin-speaking children's interpretation of bare and demonstrative nominals?
- a. Do children, like adults, have both generic and existential definite readings for subject bare nominals but never interpret demonstrative nominals as generic?
 - b. Do they distinguish between bare and demonstrative nominals by assigning different interpretations to them?
 - c. Which interpretation do they prefer for bare and demonstrative nominals in particular contexts – generic or existential definite?

In addition, if children's interpretation of bare and demonstrative nominals in subject position is different from that of adults, what are the variables influencing their interpretation?

- (22) What are the variables influencing Mandarin-speaking children's interpretation of bare and demonstrative nominals?
- d. Will a pragmatic variable, presentation order of the nominal (immediate- or delayed-mention, i.e. distance from a discourse antecedent), influence the interpretation?
 - e. Will a non-linguistic property of the sentence (canonicity '*tigers eat meat*' and non-canonicity '*tigers eat carrots*') influence the interpretation?

To answer these questions, I propose four hypotheses and predictions.

First, the null hypothesis is that, with respect to general language acquisition, if Mandarin-speaking children's grammar is adult-like, they will have both generic and

existential definite interpretations for bare nominal subjects but only existential definite interpretation for demonstrative nominals. Assuming type-shifting is automatic and innate as Chierchia (1998) suggests, it should be possible for children to assign either a generic or existential definite reading to bare nominals according to the context.

The second hypothesis is based on Chierchia (1998) and other relevant studies – de Villiers and Roeper (1995), Pérex-Leroux et al. (2004), and Vergnaud and Zubizarreta (1992). Assuming Chierchia's Nominal Mapping Parameter, I hypothesize that all Mandarin nouns start out as mass and kind-denoting, can function as bare arguments, and can freely shift between indefinite and definite readings, as shown in Chapter 2. The most economical way for Mandarin-speaking children to interpret bare nominals seems to be referring to kinds, since kind-denoting is the default interpretation and does not require paying attention to the discourse. For this reason, the kind reading should be preferred whenever possible. Accordingly, younger Mandarin-speaking children, like English-speaking children, will prefer the generic reading for bare nominals with characterizing statements. Older Mandarin-speaking children's generic interpretation for bare nominals will decrease and diverge from that of English-speaking children. English-speaking children should always allow the generic reading for bare nominals since this is the only acceptable reading given a characterizing statement. In brief, I predict that children will at first allow generic readings of bare nominals in Mandarin as much as what is evidenced in English.

de Villiers and Roeper (1995:73, 101) claim that children initially assume NP as the maximal projection for complex nominals until a DP projection is justified. Accordingly, it is anticipated that children may project NP for Mandarin demonstrative

nominals and assign the generic reading to them, as if the demonstrative is transparent to them and the NP receives a generic interpretation by default (Chierchia 1998).

Vergnaud and Zubizarreta (1992) argue that the kind-referring reading is one of the possible readings for definites across languages, which is attested in the acquisition of definite nominals of English and Spanish in Pérez-Leroux et al. (2004). If Mandarin-speaking children treat the demonstrative the same as English and Spanish-speaking children treat definite articles, they may accept the generic reading as part of the semantics of demonstratives. Based on de Villiers and Roeper, Vergnaud and Zubizarreta, and Pérez-Leroux et al., I predict that children will allow generic readings of demonstrative nominals in Mandarin.

The third hypothesis is based on Crain et al. (1994) and two other studies – Maratsos (1976) and Karmiloff-Smith (1979). The Semantic Subset Principle in Crain et al. argues for the following (p. 455).

- (23) ‘Semantic Subset Principle: If the interpretative component of UG makes two interpretations, A and B, available for a sentence, S, and if interpretation A makes S true in a narrower range of circumstances than interpretation B does, then interpretation A is hypothesized before B in the course of language development.’

In other words, children initially hypothesize semantic representations that make a sentence true in the smallest set of circumstances. For example, sentence (24) has two possible interpretations, (24a) and (24b). ((24) is Crain et al.'s (17).)

- (24) The dinosaur is only painting a house.
a. The only thing the dinosaur is doing is painting a house. (subset)
b. The only thing the dinosaur is painting is a house. (superset)

The interpretation in (24a) is a subset of that in (24b) because (24a) is only true when the dinosaur is painting a house without doing anything else. When the dinosaur is painting a house, eating an apple, and flying a kite at the same time, only (24b) is true, but not

(24a). Crain et al.'s study finds that only the subset interpretation is accepted by children. Accordingly, they argue that the subset interpretation must be the initial interpretation available to children.

In the current study, the range making the existential definite interpretation true is narrower than the range making the generic interpretation true. The circumstances that make the existential definite interpretation true (*The tigers in the picture eat carrots*) are a subset of the circumstances that make the generic interpretation true (*All tigers in the world eat carrots*). Hence, the existential definite reading for bare nominals should emerge earlier than the generic reading in the acquisition of Mandarin. Despite the ambiguous readings of Mandarin bare nominals, children may initially over-assign the existential definite reading to bare nominals because it is the more restricted reading of Mandarin bare nominals (the semantic subset) and is acquired at an earlier age.

The idea that the existential definite reading for Mandarin bare arguments is preferred is consistent with the studies of Maratsos (1976:63) and Karmiloff-Smith (1979:71-72). Both studies bring forth the egocentric theory to account for the overuse of English and French definites in places where indefinites are expected.⁵³ In Maratsos' Cave story experiment, children were told about a groups of boys and girls stuck in a cave because of rain, and eventually one went outside to see whether it was still raining. When asked 'who do you think went outside?' many children responded with an incorrect egocentric answer 'the girl went out,' even though no particular girl had been introduced to the listener, and 'a girl' should be used. Based on the egocentric theory, it is possible that Mandarin-speaking children, while at the egocentric stage, will have

⁵³ Maratsos' study tested children of age three and four. In Karmiloff-Smith's study of French, children did not use the definites correctly until age six.

difficulties considering the listener's viewpoints. Therefore, in the Mandarin story of carrot-eating tigers, when asked 'do tigers eat meat?' they will answer 'no, (the) tigers do not eat meat' interpreting the bare nominal as existential definite, even though meat-eating tigers are not introduced in the context and children should consult their real world or linguistic knowledge to interpret the ambiguous Mandarin bare nominals.

Mandarin demonstrative nominals only have an existential definite interpretation. In accordance with the Semantic Subset Principle and the egocentric theory, it is predicted that children will interpret demonstrative nominals as existential definite. According to these studies, it is predicted that children will prefer existential definite readings of bare and demonstrative nominals in Mandarin.

The fourth hypothesis and prediction concern the variables that may affect interpretation. Pérez-Leroux et al. (2004) examined two variables that may influence children's interpretation of nominals – pragmatics (presentation order) and non-linguistic properties (canonicity of the sentence). Although no significant effects were found in the English and Spanish data, except the pragmatic variable in the Spanish data, there are good reasons to assume these two variables may have a significant effect in Mandarin. Mandarin lacks in/definite articles and number morphology and relies on the context to make proper referential and number interpretation. Therefore, it is argued to be a discourse-oriented language (Huang 1984:549-551, Tsao 1977). The context, such as pragmatic or non-linguistic variables, should consequently have more effects on Mandarin than it has on languages with articles, such as English and Spanish. Accordingly, I predict that these two variables will have effects on children's interpretation of nominals.

Predictions about children's interpretation of Mandarin nominals are summarized below.

- (25) a. Prediction based on the availability of an adult-like grammar of children (null hypothesis): Young children will have both generic and existential definite interpretations for bare nominal subjects but only existential definite interpretation for demonstrative nominals.
- b. Prediction based on the Nominal Mapping Hypothesis and others: Young children will have generic readings of bare and demonstrative nominals.
- c. Prediction based on the Semantic Subset Principle and others: Young children will have existential definite readings of bare and demonstrative nominals.
- d. Prediction based on the discourse-oriented nature of Mandarin: The pragmatic (immediate- or delayed-mention) and non-linguistic (canonicity or non-canonicity) variables examined in this study will have effects on the interpretation of nominals.

5.5 A Comprehension Experiment

The experiment in this study was adapted, with some changes, from its English version created by Pérez-Leroux et. al (2004). Due to the lack of obligatory number morphology in Mandarin, bare nominals can have either singular or plural interpretations. To avoid this ambiguity, in the current study, all Mandarin bare nominals have plural interpretation in the context and may be interpreted as generic or existential definite, and all demonstrative nominals appear with the plural demonstrative *zhexie* ‘these.’ All Mandarin test sentences appear without any aspect morphology and with characterizing statements only. Therefore, the aspect-less test sentences are anticipated to trigger more generic interpretation for bare nominals, while demonstrative nominals should be interpreted as existential definite only. The English/Spanish (Pérez-Leroux et al. 2004)

and Mandarin experiments are basically identical in order to compare the results, except for the following changes.

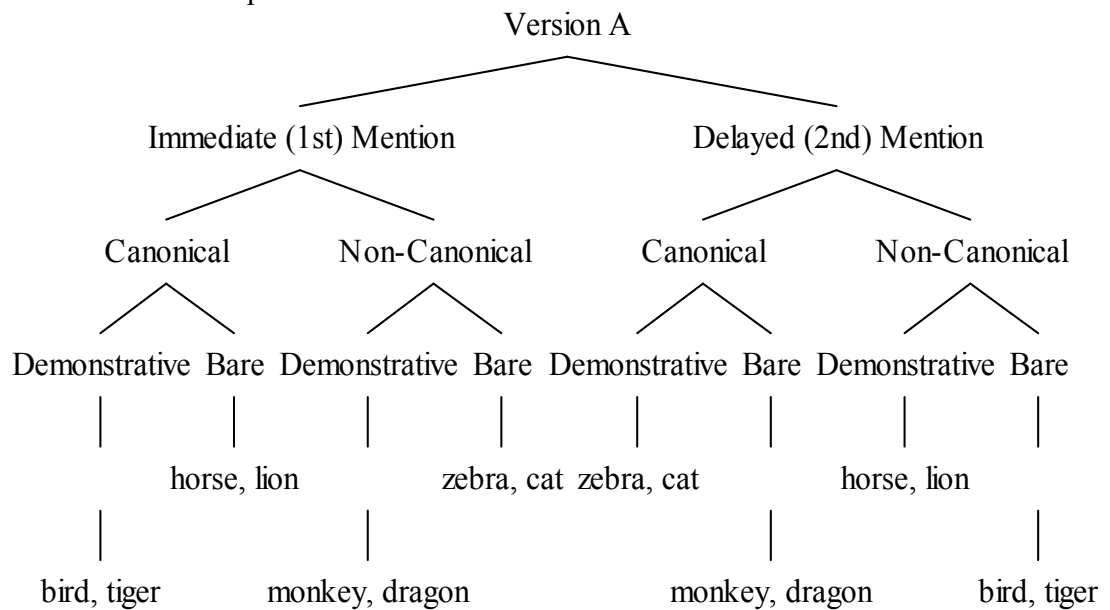
First, there is no definite article in Mandarin, and the demonstrative *zhexie* ‘these’ is used. Second, instead of only one observing animals in the pictures used in the experiment of Pérez-Leroux et al., there are two observing animals of the same kind in the pictures in the Mandarin task (all visual stimuli were made to have plural reading). Third, in Pérez-Leroux et al., the animals in the present tense study have proper names, but the animals in the present vs. past tense study do not have proper names. They find that when the animals have names, it is more natural to use a pronoun. This could have explained the higher number of generic responses for definite nominals because children expect the experimenter to continue calling them by name, and not by using definite nominals. The Mandarin study introduces the stories without naming the animals.

5.5.1 Stimuli and Procedure

There were eight stories, and each had two atypical members of a kind and two observers. The eight target stories were about spotted zebras (no stripes), monkeys that eat grass (not bananas), cats that love to take a bath (not to stay dry), tigers that eat carrots (not meat), three-legged horses (not four-legged), birds that live in caves (not in nests), dragons that breathe bubbles (not fire), and lions that live on boats (not on savannah). Each story was told when a laptop screen was showing a full-sized picture with the two atypical animals and two observers. The picture was presented all the time until the participants finished answering all questions. Child participants were tested individually by the experimenter, a native Mandarin speaker, where adult controls were tested in groups. They watched pictures in a projected slide show and circled the answers

on an answer sheet. Four yes/no questions were asked in each story – two target questions and two distraction questions. (One of the two distraction questions asked for positive response, and the other, negative). There were a total of 32 questions – 16 targets and 16 distracters. Additionally, 24 filler questions were tested — either yes/no questions or questions that have two answer options to choose from. There were two versions of the experiments, A and B, and the test questions were counterbalanced across stories in the presentation order (immediate- or delayed-mention), canonicity (typical or atypical properties), and determiners (bare or demonstrative). Within each version, half of the stories started with the canonical question, and the other half started with the non-canonical question. The counterbalancing schema of the study is presented as follows.

Figure 5.2 The counterbalancing design in version A and B of the comprehension experiment



```

graph TD
    VB[Version B] --> IM[Immediate 1st Mention]
    VB --> DM[Delayed 2nd Mention]
    IM --> IC[Canonical]
    IM --> IN[Non-Canonical]
    DM --> DC[Canonical]
    DM --> DN[Non-Canonical]
    IC --> IDC[Demonstrative]
    IC --> ICB[Bare]
    IN --> IND[Demonstrative]
    IN --> INB[Bare]
    DC --> DDC[monkeys, dragons]
    DN --> DDC
    DN --> DDB[birds, tigers]
    DDB --> DDB1[zebras, cats]
    DDB --> DDB2[horses, lions]
    IND --> IDB1[zebras, cats]
    IND --> IDB2[horses, lions]
    IDB1 --> IDB1_1[zebras, cats]
    IDB1 --> IDB1_2[horses, lions]
    IDB2 --> IDB2_1[zebras, cats]
    IDB2 --> IDB2_2[horses, lions]
    ICB --> ICB1[birds, tigers]
    ICB1 --> ICB1_1[horses, lions]
    ICB1 --> ICB1_2[zebras, cats]
    INB --> INB1[monkeys, dragons]
    INB1 --> INB1_1[zebras, cats]
    INB1 --> INB1_2[horses, lions]
  
```

The diagram illustrates the classification of mentions in Version B. It starts with 'Version B' at the top, which branches into 'Immediate (1st) Mention' and 'Delayed (2nd) Mention'. Each of these branches further into 'Canonical' and 'Non-Canonical'. Under 'Immediate (1st) Mention', 'Canonical' branches into 'Demonstrative' and 'Bare', leading to 'monkeys, dragons' and 'birds, tigers' respectively. 'Non-Canonical' branches into 'Demonstrative' and 'Bare', leading to 'zebras, cats' and 'horses, lions' respectively. Under 'Delayed (2nd) Mention', 'Canonical' branches into 'Demonstrative' and 'Bare', leading to 'monkeys, dragons' and 'birds, tigers' respectively. 'Non-Canonical' branches into 'Demonstrative' and 'Bare', leading to 'zebras, cats' and 'horses, lions' respectively. The examples are further grouped into 'zebras, cats' and 'horses, lions' for each category.

Figure 5.3 Mandarin experiment example

(These two tigers only eat vegetables. See, they're eating carrots. The rabbits are happy because They won't get eaten. Now let me ask you.)



(25) Version A:

Immediate, Canonical, Demonstrative

- a. **zhexie laohu** chi rou ma?
these tiger eat meat MA
'Do these tigers eat meat?'

Filler questions

- b. **zhexie laohu** you tiaowen ma?
these tiger have stripe MA
'Do these tigers have stripes?'

- c. **ni** you tiaowen ma?
you have stripe MA
'Do you have stripes?'

Delayed, Non-canonical, Bare

- d. **laohu** chi hongluobo ma?
tiger eat carrot MA
'Do tigers eat carrots?'

(26) Version B:

Immediate, Canonical, Bare

- a. **laohu** chi rou ma?
tiger eat meat MA
'Do tigers/the tigers eat meat?'

Filler questions

- b. **laohu** you tiaowen ma?
tiger have stripe MA
'Do tigers/the tigers have stripes?'

- c. **ni** you tiaowen ma?
you have stripe MA
'Do you have stripes?'

Delayed, Non-canonical, Demonstrative

- d. **zhexie laohu** chi hongluobo ma?
these tiger eat carrot MA
'Do these tigers eat carrots?'

5.5.2 Subjects

The participants in the comprehension experiment, including version A and B, were sixty-one preschoolers and forty seven undergraduates, as listed below.

Table 5.2 Subjects of the Experiment in Version A and B

Version	Age Group	Subjects	Mean Age	Standard Deviation
A	Three	8	3;6	0;2
	Four	14	4;6	0;3
	Five	9	5;8	0;3
	Adult	21	Undergraduates	
B	Three	9	3;9	0;3
	Four	9	4;5	0;3
	Five	14	5;5	0;3
	Adult	26	Undergraduates	

The child participants were recruited from local preschools in Taichung, Taiwan. All of the adult controls were undergraduates at National Tsing Hua University in Taiwan.

5.5.3 Results

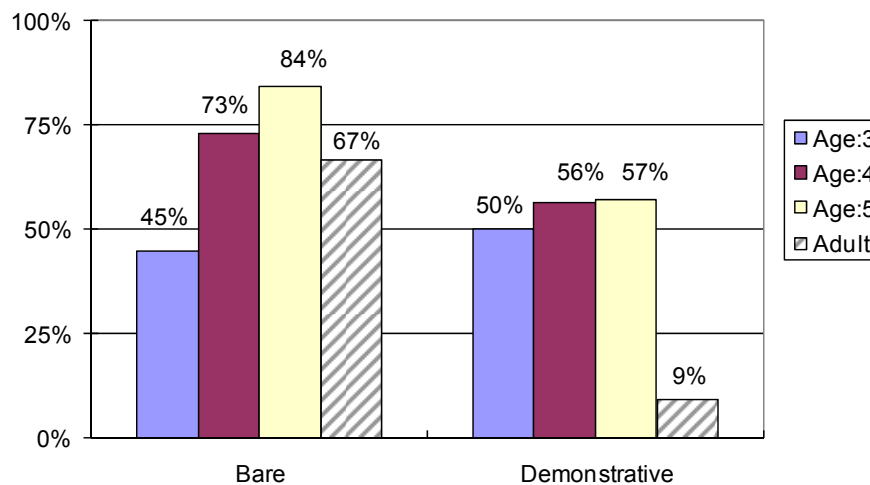
Since no significant effect was found in the variable of version (mixed design ANOVA, $F(1,94) = .047, p = .829$), the results from the two versions were combined. The results presented here are from a total of 110 subjects, divided into four different age groups.

Table 5.3 Subjects of the Experiment

Age Group	Subjects	Mean Age	Standard Deviation
Three	17	3;8	0;3
Four	23	4;5	0;3
Five	23	5;6	0;4
Adult	47	Undergraduates	

Figure 5.4 shows the mean proportion of generic responses for bare and demonstrative nominals among all four age groups. Based on a mixed design ANOVA, there is a significant effect of nominal types by age group interaction ($F(3,94) = 9.992, p < .001$). Among the children subjects, there are also significant effects of nominal types ($F(1,51) = 4.873, p = .03$) and age groups ($F(2,51) = 5.515, p = .007$).

Figure 5.4 Mean proportion of generic responses



As shown in Figure 5.4, the subjects generally prefer assigning the generic interpretation to bare nominals in comparison to demonstratives. The interpretation of bare nominals

shows an A-shaped tendency: the tendency of interpreting bare nominals as generic increases from three- to five-year-olds and then decreases in the adult group. The three-year-olds assign interpretations to bare and demonstrative nominals at chance level (exact binomial p (two-tailed) $> .14$). It is apparent that older children and adults prefer the generic reading for bare nominals more than for demonstrative nominals. There exists a considerable amount of generic responses for demonstrative nominals, particularly in four- and five-year-olds (56-57%). Even adults gave a small portion of non-target generic reading (9%). The results show that children, like adults, have both generic and existential definite readings for subject bare nominals. Four- and five-year-olds provide 73% and 84% generic reading respectively, which means the existential definite reading only accounts for 27% and 16% respectively. Children distinguish between bare and demonstrative nominals as there is a significant difference between the interpretation of bare nominals and that of demonstrative nominals ($F(1,51) = 4.873, p = .03$). Four- and five-year-olds assign the generic reading to bare nominals 73% and 84% of the time respectively, while they do so to demonstrative nominals 56% and 57% of the time respectively. Expectedly, the significant distinction between bare and demonstrative nominals is mainly due to the older age group (five-year-olds), not the four-year-olds (Age Five: $t(44) = 5.03, p = .000$; Age Four: $t(44) = 1.12, p = .25$).

The experiment examined two variables that may influence the interpretation of nominals. The first variable is the presentation order. Table 5.4 presents the percentage of generic responses for nominals in the immediate- or a delayed-mention sentences.

Table 5.4 Percentage of generic responses: presentation order

Age	Three	Four	Five	Adult
Immediate-mention	43%	57%	69%	34%
Delayed-mention	53%	75%	70%	42%

As expected, nominals appearing in the delayed-mention sentences receive more generic readings than that appearing in the immediate-mention sentences. Presentation order has a significant effect on the nominal interpretation (mixed design ANOVA, $F(1,94) = 14.631, p = .000$). Age also has a significant effect ($F(1,94) = 17.259, p = .000$). The four-year-olds are significantly affected by the presentation order when deciding which reading to assign to the nominals ($t(44) = -2.88, p = .006$). (Age Three: $t(32) = -1.40, p = .17$; Age Five: $t(44) = -.173, p = .86$; Adult: $t(92) = -1.86, p = .07$.)

Examining this pragmatic effect more closely in different nominal groups, as shown in Figure 5.5 and 5.6, we can see that presentation order has a more consistent effect for demonstratives than for bare nominals across ages.

Figure 5.5 Percentage of generic responses for bare nominals: presentation order

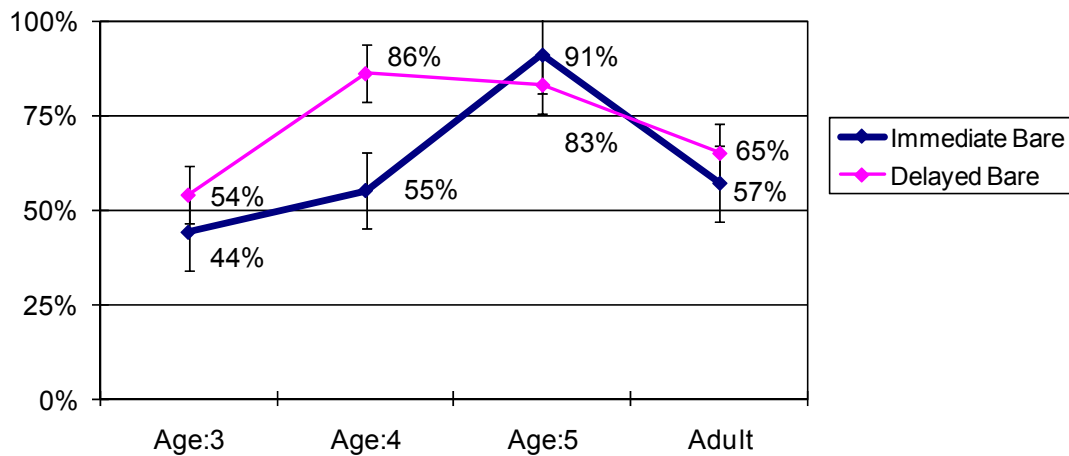
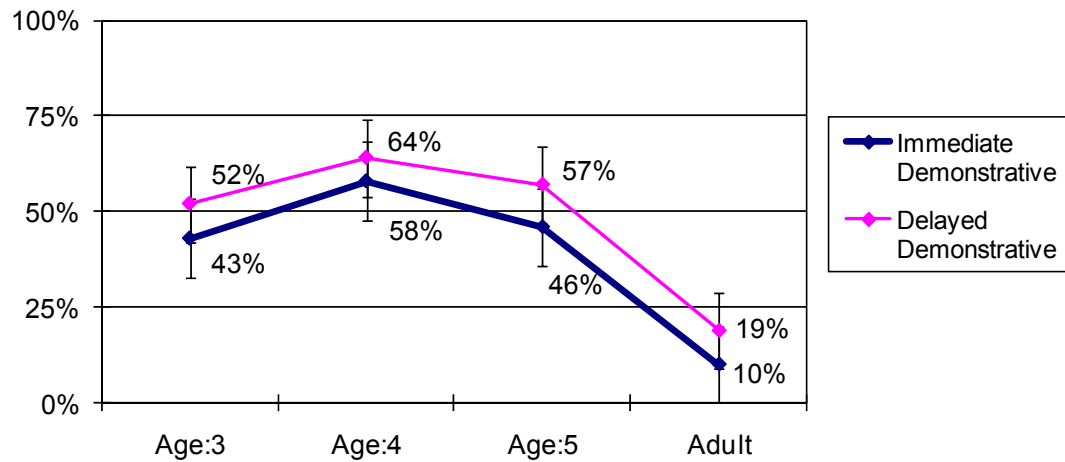


Figure 5.6 Percentage of generic responses for demonstrative nominals:
presentation order



In Figure 5.5, presentation order shows a significant effect on the reading assigned to bare nominals among all subjects ($t(218) = -2.24, p = .0026$). Four-year-olds are significantly affected by the presentation order when interpreting bare nominals (Age Four: $t(44) = -4.15, p = .0002$). (Age Three: $t(32) = -1.14, p = .26$; Age Five: $t(44) = 1.48, p = .15$; Adult: $t(92) = -1.12, p = .26$.)

Presentation order also shows a significant effect on the interpretation of demonstrative nominals among all subjects ($t(218) = -2.00, p = .0047$), as shown in Figure 5.6. However, this pragmatic variable does not have a significant effect within any age groups, although as predicted, nominals in delayed-mention do receive more generic responses (Age Three: $t(32) = -1.35, p = .19$; Age Four: $t(44) = .795, p = .43$; Age Five: $t(44) = -.173, p = .86$; Adult: $t(92) = -1.86, p = .067$).

The second variable is a non-linguistic property of the predicates. Table 5.5 presents the percentage of generic responses for nominals appearing in sentences with either the canonical (*tigers eat meat*) or non-canonical (*tigers eat carrots*) property.

Table 5.5 Percentage of generic responses: property of predicates

	Age	Three	Four	Five	Adult
Canonical		67%	66%	66%	45%
Non-canonical		29%	66%	72%	30%

As expected, nominals appearing with the canonical property generally receive more generic interpretation than that appearing with the non-canonical property. A mixed design ANOVA shows that there are significant effects of this property of predicates ($F(1,102) = 7.660, p = .007$) and of age groups ($F(3,102) = 17.452, p = .001$). Significant effects are found in the group of the three-year-olds and the adults, who make a significant distinction between the reading for nominals with canonical property and those with non-canonical property (Age Three: $t(32) = 4.9, p = .0001$; Adult: $t(92) = 3.27, p = .001$). (Age Four: $t(44) = 0, p = 1$; Age Five: $t(44) = .977, p = .33$.)

Figure 5.7 and Figure 5.8 present the effect of canonicity on the interpretation of bare and demonstrative nominals respectively.

Figure 5.7 Percentage of generic responses for bare nominals: property of predicates

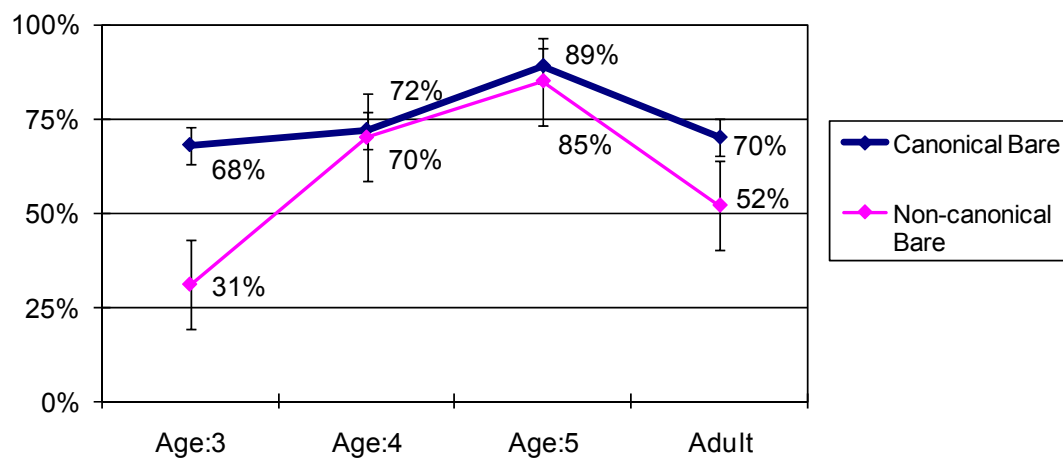
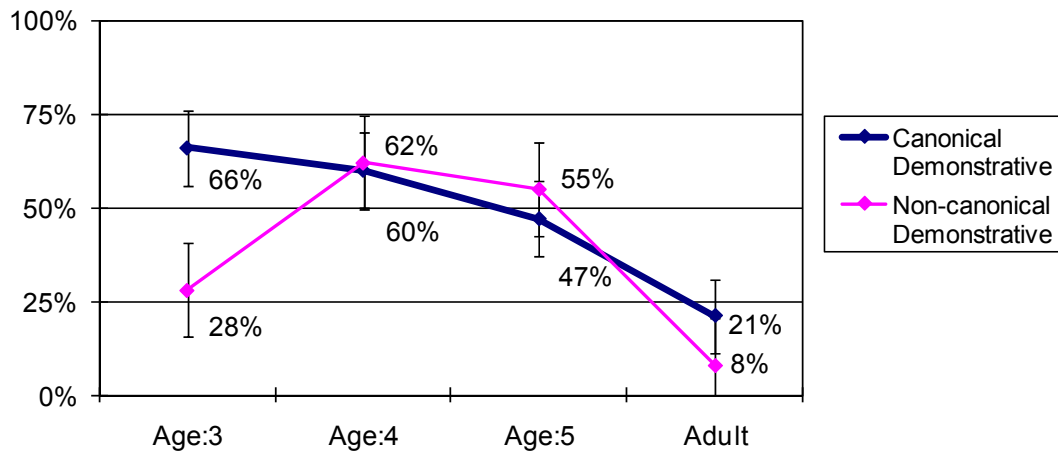


Figure 5.8 Percentage of generic responses for demonstrative nominals: property of predicates



In Figure 5.7, canonicity of the sentences is found to have a significant effect on the reading assigned to bare nominals among all subjects ($t(218) = 2.95, p = .0035$). Three-year-olds and adults are significantly affected by the canonicity when interpreting bare nominals (Age Three: $t(32) = 3.64, p = .0009$; Adult: $t(92) = 2.42, p = .017$). (Age Four: $t(44) = 307, p = .76$; Age Five: $t(44) = -.780, p = .44$.)

Similarly, in Figure 5.8 significant effects are found in the reading assigned to the demonstrative nominals appearing in either canonical or non-canonical sentences among all subjects ($t(218) = 1.99, p = .0047$) and among three-year-olds and adults (Age Three: $t(32) = 4.8, p = .0001$; Adult: $t(92) = 2.95, p = .0004$). (Age Four: $t(44) = -.22, p = .83$; Age Five: $t(44) = .851, p = .40$.)

In sum, bare nominals significantly obtain a higher proportion of generic reading than demonstrative nominals. Demonstrative nominals obtain a high proportion of generic reading only in children, but not in adults. Significant interpretation differences are also found across age groups. Mandarin-speaking children, like adults, have both

generic and existential definite readings for subject bare nominals. The four- and five-year-olds significantly distinguish between bare and demonstrative nominals.

Both variables examined in the study significantly affect the interpretation of bare and demonstrative nominals. First, a pragmatic variable – presentation order (immediate- or delayed-mention) – significantly influences the interpretation. A significant effect is found in the four-year-old group. Nominals appearing in delayed-mention sentences receive more generic readings. Second, a non-linguistic property of the predicates (canonicity or non-canonicity) also significantly influences the interpretation. Significant effects are found in the group of three-year-olds and adults. Nominals appearing in canonical sentences receive more generic readings.

5.5.4 Individual Analyses and Discussion

In this study, I found two types of possible interpretation biases. First, the generic bias – always choosing the generic reading regardless of bare or demonstrative nominals – involves selecting the non-target generic interpretation for demonstratives. The generic bias is only found in children – 2 four-year-olds and 3 five-year-olds (5 out of 63). Second, the existential definite bias – always choosing the existential definite reading regardless of bare and demonstrative nominals – is a grammatically plausible bias because bare nominals can be interpreted as either generic or existential definite. The existential definite bias is only found in five adults (found in 5 of 47 adults). No existential definite bias is found in children.

Before examining the individual responses more closely, given the overall high percentage of generic responses for both types of nominals, I hypothesize that those who

provide more generic responses for demonstratives will also have more generic responses for bare nominals because they may be biased by the story content and therefore prefer generic reading regardless of the nominal types.⁵⁴ The predication is that there will be a positive correlation between the generic reading for demonstratives and the generic reading for bare nominals – especially among the subjects who prefer generic reading for demonstratives – but not vice versa.

As shown in Table 5.6, a positive correlation is found between adults' generic interpretation of demonstratives and that of bare nominals ($r(47) = .23, p = .06$). As predicted, the adults who assigned generic reading for demonstratives over half of the time (the gray area) also preferred generic reading for bare nominals and this positive correlation is significant ($r(5) = .80, p = .05$).

⁵⁴ The experiment of this study was revised from its English version to the Mandarin version for the purpose of comparing. To discover whether there are any cultural differences about the knowledge of the target animals, a knowledge test was conducted. The knowledge test included 16 target questions – exactly the 8 pairs of animate bare nominal subjects in the comprehension task (the questions for the knowledge test are listed in Appendix IV). One pair of target questions is provided below.

- (i) a. shizi zhu zai chuan shang ma?
 lion live PROG boat on MA
 'Do lions live on a boat?'
 b. shizi zhu zai caoyuan li ma?
 lion live PROG savannah in MA
 'Do lions live in the savannah?'

There were 26 fillers, including 8 pairs of inanimate bare nominal subjects and 10 distracters. The participants were 13 Mandarin native-speaking adults. Their responses show that the average percentage of correct answers is 86%. *Changjinglu* 'zebra' and *mao* 'cat' have lower percentages, 70% and 66% respectively. Excluding the responses for these two animals, the average percentage of correct answers is 93%. Therefore in the individual analysis, the responses for 'zebra' and 'cat' were not included. Only the other six animals were examined.

Table 5.6 Listing of individual responses of adults

	↓						↓						↓							
Generic responses for Demonstratives	0	0	0	0	0	0	0	1	1	1	1	1	1	2	2	2	4	5	5	
Generic responses for Bare Nominals	0	1	2	3	4	5	6	0	1	2	4	5	6	1	3	5	4	5	6	
<i>Number of subjects</i>	5	1	2	3	3	7	5	2	1	2	3	4	1	1	1	1	1	2	2	47
Correlation ⁵⁵							-.06						.80				.23			

On the contrary, for those who prefer a generic reading for bare nominals, their interpretation decisions only reflect the given stories, and their interpretation of bare nominals does not predict their interpretation of demonstratives. As indicated by the arrows on the top of Table 5.6, the adults who assigned generic readings for bare nominals all the time (6 out of 6) did not necessarily prefer generic reading for demonstratives.

In Table 5.6, most adult participants (89%, 42 out of 47) assign the non-target generic readings to demonstrative nominals between zero and two times. Adults seldom assign generic readings to demonstratives. The reading they assign to bare nominals depends on their decision on each question (this accords with the fact that bare nominals are ambiguous between generic and existential definite reading in the given context).

In the three-year-olds' data, there is a significant positive correlation between the generic interpretation of demonstrative and of bare nominals overall ($r(17) = .49, p = .02$), as shown in Table 5.7. In the prefer-generic-reading-for-demonstratives group (gray area), a positive correlation also exists ($r(6) = .40, p = .22$).

⁵⁵ Positive correlation (+); negative correlation (-). The degree of relationship is measured by the numerical value of the correlation (0 to 1). Correlation does not mean causation.

Table 5.7 Listing of individual responses of three-year-olds

														↓
Generic responses for Dem.	1	1	2	2	2	3	3	3	4	4	4	4	5	
Generic responses for Bare	1	2	2	3	4	2	1	3	2	3	4	6	5	
<i>Number of subjects</i>	1	1	1	2	1	1	1	3	2	1	1	1	1	17
Correlation	.22								.40				.49	

The four-year-olds' data are closest to the adult's data: significant positive correlations are found both in the whole age group ($r(23) = .63, p = .00$) and in the prefer-generic-reading-for-demonstratives group (gray area) ($r(15) = .53, p = .02$). As what is found in the adult group, the prefer-generic-reading-for-bare group provided various amount of generic reading for demonstratives (indicated by the arrows on the top of Table 5.8).

Table 5.8 Listing of individual responses of four-year-olds

						↓						↓			↓			↓	
Generic responses for Dem.	0	1	1	2	2	2	3	3	4	4	4	4	5	5	6	6			
Generic responses for Bare	1	2	5	4	5	6	3	4	3	4	5	6	5	6	5	6			
<i>Number of subjects</i>	1	1	1	1	1	1	1	1	1	2	3	1	3	2	1	2	23		
Correlation	0								.53								.63		

In the result (Figure 5.2) of the mean percentage of generic reading for each nominal type and age group, the five-year-olds' data are most puzzling. They perform the highest percentage, 57%, of generic reading for demonstratives. After examining five-year-olds' individual data in Table 5.7, I find that their behavior is actually similar to that of other age groups with an overall significant positive correlation ($r(23) = .44, p = .01$) between their generic interpretation of demonstratives and that of bare nominals. Although the negative correlation in the prefer-generic-reading-for-demonstratives group (gray area) is different from the positive correlation found in this sub-group of other age groups ($r(10) = -.31, p = .19$), this correlation is not significant. This simply suggests that individual preferences do exist.

Table 5.9 Listing of individual responses of five-year-olds

				↓	↓	↓		↓		↓	
Generic responses for Dem.	0	0	2	2	3	4	5	5	6	6	
Generic responses for Bare	4	5	5	6	6	6	5	6	5	6	
<i>Number of subjects</i>	2	2	1	4	4	2	1	2	2	3	23
Correlation				.32				-.31			.44

Comparing the scatter charts of each adult's and each child's responses in Figure 5.9 and 5.10, I find that the pattern in children's data is similar to that in adult's data. No data appear in the lower right area in either Figures. This supports the prediction that there is a positive correlation between the generic reading assigned to demonstratives and that assigned to bare nominals. On the contrary, the preference of assigning generic responses to bare nominals does not predict the preferred reading of demonstratives.

Figure 5.9 Individual adult's generic reading
(n=47, each dot represents one person)

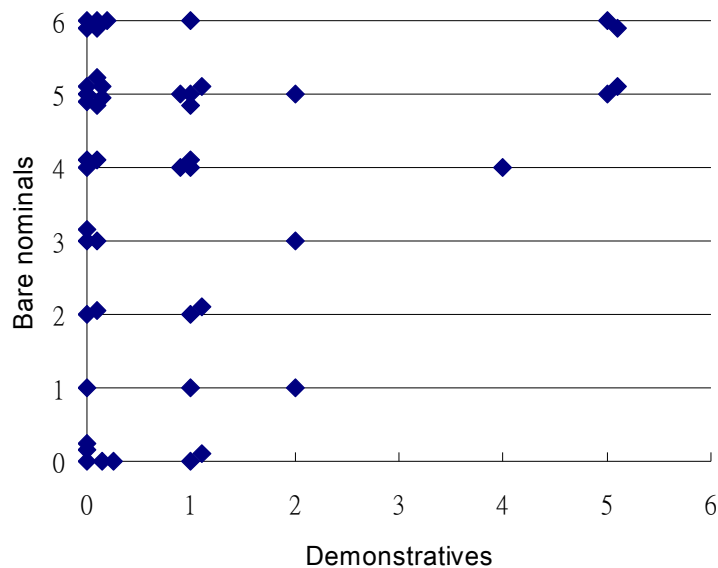
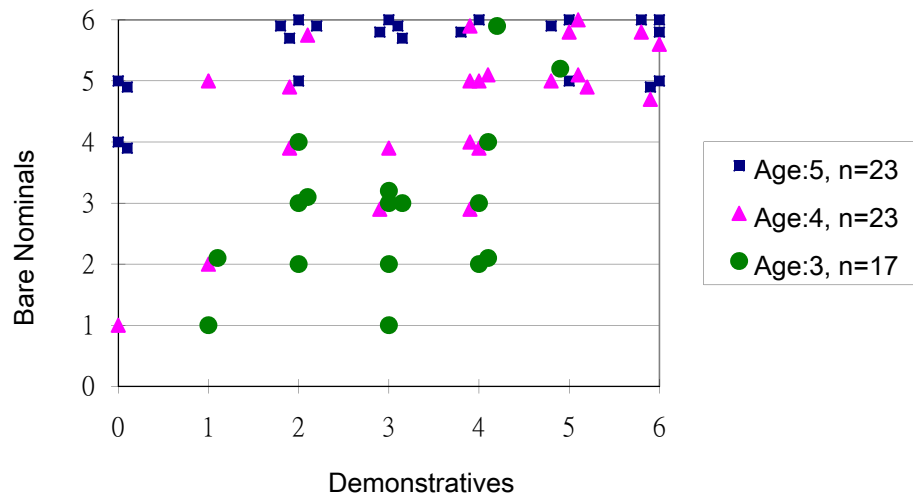


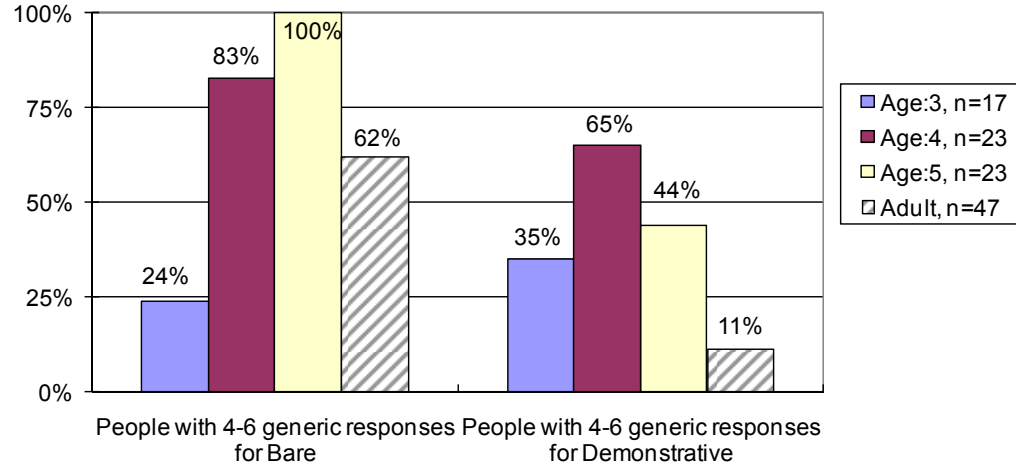
Figure 5.10 Individual child's generic reading
(n=63, each dot represents one child)



These two scatter plots reveal that the four-year-olds' pattern resembles more to that of the adults than other age groups do. In the individual analysis, the four-year-olds behave more adult-like than the five-year-olds; therefore, the puzzle of the high generic responses for demonstratives (and also for bare nominals) in the five-year-olds remains unsolved. (A possible explanation will be provided in the general discussion section.)

The individual analysis of those providing generic reading more than half of the time helps clarify the data, especially the unexpected non-target generic reading for demonstratives assigned by five-year-olds and adults, averaging 57% and 9% respectively. Figure 5.11 illustrates that the number of subjects who prefer the non-target generic reading for demonstratives apparently is lower in both the five-year-old and adult groups.

Figure 5.11 Percentage of people who prefer generic readings



In the adult group, 11% of them gave generic responses to demonstratives more than half of the time (4-6 responses). Clearly, these people are the main cause for the unexpected generic responses for demonstratives in the adult group, when the majority of others assigned existential definite readings to demonstratives. The situation is similar with the five-year-olds. The high mean of generic responses of demonstratives is a skewed result from 44% of the five-year-olds who have preference for generic readings. A total of 24% and 35% three-year-olds prefer generic readings for bare and demonstrative nominals respectively. The numbers are much lower than three-year-olds' mean percentages of generic responses, 45% for bare and 50% for demonstratives. This indicates that, unexpectedly, not many three-year-olds show a preference for generic readings in their overall responses. As for the number of people preferring generic reading for bare nominals, a surprisingly high percentage in both the four- and five-year-old groups needs an explanation. (The discussion section will suggest a possible explanation based on the connection between aspect and genericity.)

5.5.5 Summary

The study finds that Mandarin-speaking children assign generic interpretations to bare nominals more than they assign existential definite reading to bare nominals. Demonstrative nominals also receive generic interpretation about half of the time. Mandarin-speaking children, like adults, have both generic and existential definite readings for subject bare nominals. Children, especially four- and five-year-olds, significantly distinguish between bare and demonstrative nominals – bare nominals receive more generic readings than demonstratives do. At least two variables significantly affect Mandarin speakers' interpretation of bare and demonstrative nominals: pragmatics (presentation order) and non-linguistic properties of the predicates (in this case, part of the knowledge of the world). Nominals appearing in delayed-mention sentences and in sentences with canonical property receive more generic readings.

5.6 General Discussion and Concluding Remarks

The goal of this chapter was to answer the main research questions: (i) How do Mandarin-speaking children interpret bare and demonstrative nominals in Mandarin? (ii) What are the variables influencing their interpretation of bare and demonstrative nominals?

The findings are summarized below, excluding the three-year-olds, who seem to perform at chance level:

(27) **Nominal interpretation:**

What is Mandarin-speaking children's interpretation of bare and demonstrative nominals?

Finding: For bare nominals: the generic reading is preferred; 73% and 84% in age four and five respectively.

For demonstrative nominals: a considerable amount of generic responses; 56% and 57% in age four and five respectively.

- a. Do children, like adults, have both generic and existential definite readings for subject bare nominals but never interpret demonstrative nominals as generic?

Finding: Like adults, children assign both generic and existential definite readings to bare nominals. But the generic responses assigned by four- and five-year-olds (73-84%) are higher than that by adults (67%).

Although adults provide an unexpected 9% generic responses of demonstrative nominals, it is still significantly different from children's interpretation. Four- and five-year-olds assign generic reading to demonstratives 56-57% of the time.

- b. Do they distinguish between bare and demonstrative nominals by assigning different interpretations?

Finding: Yes, there is a significant difference between the interpretation of bare nominals and that of demonstrative nominals. The four- and five-year-olds assign the generic reading to bare nominals 73-84% of the time, while assigning the generic reading to demonstrative nominals 56-57% of the time.

- c. Which interpretation do they prefer for bare and demonstrative nominals in particular contexts – generic or existential definite?

Finding: The four- and five-year-olds prefer the generic interpretation for both bare (73-84%) and demonstrative nominals (56-57%).

Variables affecting interpretation:

What are the variables influencing their interpretation of bare and demonstrative nominals?

- d. Will a pragmatic variable, presentation order of the nominals (immediate- or delayed-mention, i.e. distance from a discourse antecedent), influence the interpretation?

Finding: Yes, the pragmatic variable examined in the experiment significantly influences the Mandarin-speaking subjects' interpretation. Nominals appearing in the delayed-mention sentences receive more generic readings.

- e. Will a non-linguistic property of the sentence (canonicity '*tigers eat meat*' and non-canonicity '*tigers eat carrots*') influence the interpretation?

Finding: Yes, the non-linguistic property of the predicates examined in the experiment significantly influences the Mandarin-speaking subjects' interpretation. Nominals appearing in sentences with canonical property receive more generic readings.

5.6.1 Interpretation of Bare and Demonstrative Noun Phrases

The study finds that Mandarin-speaking children and adults make a statistically significant distinction between the interpretation of bare and demonstrative nominals. Across age groups, bare nominals tend to be generic, while demonstrative nominals tend to be existential definite (although individual analysis shows that some child subjects and a few adults allow generic readings to both types of nominals).

In adult Mandarin grammar, bare nominal subjects have the chance to be interpreted as generic about half of the time because of their ambiguity between the generic and existential definite interpretations. The results show the generic reading for bare nominals in adult subjects is higher than fifty percent, 67%. Such a result is expected because the given context (animate objects and lack of aspect makers) supposedly should trigger more generic reading, as found in the Spanish and English studies by Pérez-Leroux et al. (2004). The result shows that the four- and five-year-olds have a strong preference toward generic reading in this study.

There is an unexpected high rate of the non-target generic reading for demonstrative nominals. The findings for demonstrative nominals deviate from our

common belief that demonstrative nominals, such as the ones used in our experiment following Mandarin *zhexie* ‘these,’ should never have a generic reading referring to the natural kinds in the real world.

The high proportion of the generic interpretation of bare nominals in child Mandarin supports the prediction, based on Chierchia (1998), that young children will allow generic readings of bare and demonstrative nominals in Mandarin. The result, which presents a generic preference for bare nominals and a considerable amount of generic readings for demonstratives, does not support the prediction based on Semantic Subset Principle of Crain et al. (1994), which predicts that children will assign the definite reading for both bare and demonstrative nominals at the initial stage. Following Chierchia’s theory, it would be economical for children to interpret bare nominals as kind-referring (i.e. generic) since this is the default reading. To interpret bare nominals as existential definite, like adults are capable of, requires referring to the discourse and is more demanding.

5.6.2 Variables Affecting the Interpretation

Studying the interpretation of nominals requires examining both the linguistic and the non-linguistic factors. The study shows that a pragmatic variable (the presentation order) and a non-linguistic property (the canonicity and non-canonicity of the sentence) both have statistically significant effects on Mandarin nominal interpretation. Presentation order plays a significant role in four-year-olds' nominal interpretation. It suggests that they are more pragmatically aware. The non-linguistic property of predicates has a significant effect on the three-year-olds. This suggests that they rely on a

non-linguistic property of predicates (in this case, part of their knowledge of the world) more than on the linguistic knowledge when interpreting sentences.

The findings of the significant effects of the pragmatic variable and the non-linguistic property are different from the findings for English and Spanish in Pérez-Leroux et al. (2004). In the English study, both variables are non-significant.

Presentation order is significant in the Spanish study. Only in Mandarin does the variable of non-linguistic property play a significant role. The fact that the significance of both variables is found in Mandarin but not in other two languages supports my prediction, given that Mandarin is a discourse-oriented language and relies more on pragmatics and non-linguistic properties during comprehension than structure-oriented languages, such as English or Spanish, do (Tsao 1977, Huang 1984:549-51).

5.6.3 Non-target Interpretation of Demonstrative Noun Phrases

This research has discovered a non-target generic interpretation of Mandarin demonstrative nominals assigned by both children and adults. The analysis of individual differences provides some insight about this.

First, there is a generic bias found in five children, who always selected the generic reading regardless of nominals types, but no adults had such a bias. Five adults selected the existential definite reading for both nominal types all of the time, but no children had such a bias. Why the children show a generic bias but not an existential definite bias may be that they did not pay close attention to the experiment, resulting in answers strictly based on their real-world knowledge.

Second, there exists a positive correlation between the generic reading for demonstratives and for bare nominals (the correlation is significant in all the child

groups). Participants who preferred the generic reading for demonstratives tended to assign generic readings to bare nominals, but not vice versa. This suggests that non-target generic reading of demonstratives probably results from personal preference.

Third, a closer look at the individual's responses also reveals that the high percentage of non-target generic responses of demonstratives is a skewed result caused by certain subjects with a strong preference for generic readings.

One possible account for assigning the generic reading to demonstratives is that cross-linguistically all children allow generic readings for definite expressions for a certain period of time. Mandarin-speaking children seem to underspecify Mandarin demonstratives and treat them like the Spanish or English definites, which allow generic readings. In other words, for Mandarin-speaking children, demonstrative *zhexie* 'these' is a 'less-specified' determiner head like Spanish *los* 'the' and allows generic interpretation as part of its grammar (Pérez-Leroux et al. 2004). However, after Mandarin-speaking children acquire the adult grammar of demonstratives, they should no longer interpret demonstrative as generic; thus the generic bias is not found in adult subjects. For the five adults unexpectedly interpreting demonstratives as generic over half of the time (shown in Table 5.4), further investigation is needed to find out a possible explanation.

The non-target reading of demonstratives in Mandarin brings us back to a fundamental question: when do children acquire demonstratives as determiners in Mandarin? In the section of acquisition background, we saw that articles are produced early in English, and it is the same for demonstratives in Mandarin. In Chang's corpora (2002), almost all the cases of demonstratives, *zhe* 'that' and *na* 'this' (157 and 39 tokens respectively), produced by Nana (2;0-2;6) were used as subject pronouns, except a possibly idiomatic use (28a) and a determiner use (28b).

- (28) a. zhe yangzi (possibly a fixed term for Nana at that time)
 this way/style (Nana 2;0)
 ‘this way’
- b. mama wo yao zhe yizi (Nana 2;3)
 Mom I want this chair
 ‘Mom, I want this chair.’

There are no plural demonstratives in all of Nana’s data. Some plural demonstrative cases, i.e. *zhexie* ‘these’ and *naxie* ‘those,’ are found in the data of an older child, Didi. Then again, there are only one *zhexie* and two *naxie* cases in Didi’s data (2;9-3;3) and all of them were used as subject pronouns. Neither the singular nor plural demonstratives are at first used as determiners by young children. Since Mandarin-speaking children do not initially acquire demonstratives as determiners, but as pronouns, it is not surprising that the child subjects in the study assigned considerable proportion of generic readings to *zhexie*, as if they treated *zhexie*-N as an unanalyzed unit and interpreted the whole unit as generic. Mandarin-speaking children may first project *zhexie*-N as NP, not DP (de Villier and Roeper 1995), and Mandarin NPs refer to kind by default (Chierchia 1998).

Although the reason for the considerable proportion of generic reading assigned to Mandarin demonstratives requires more scrutiny, it is possible that Mandarin demonstrative determiners are more similar to Spanish or English definite articles than expected – both are possible generic expressions. Partee (2006) mentions a common observation from C. Lyons (1999) and J. Lyons (1975) that definite articles usually derive from demonstratives. The finding seems to suggest that Mandarin demonstrative is becoming a definite article and sometimes displays the properties of definite articles, as has been argued in previous literature (Chen 2004, S. Huang 1999).

This study presents some novel and noteworthy findings about the issue among Mandarin-speaking children and adults. The findings of the non-target interpretation and

the generic and existential definite bias especially call for a more comprehensive theoretical explanation. I hope that more theoretical and acquisition studies using cross-linguistic data will shed more light on the issue, which will further our understanding of nominal interpretation and acquisition.

CHAPTER SIX

CONCLUSION

6.1 Goal and Hypotheses of the Dissertation

The goal of this research is to contribute to the understanding of the acquisition of Mandarin noun phrases. The research starts with two language-specific challenges that Mandarin poses to language acquisition as follows.

Challenge 1. Production

Since bare nominals are able to convey almost all kinds of interpretations in an appropriate context, what are the contexts that trigger the emergence of non-bare nominals?

Challenge 2. Interpretation

- a. How do Mandarin-speaking children interpret bare nominals given that bare nominals can have various interpretations?
- b. In contrast, what is the interpretation that children assign to non-bare nominals, such as demonstrative nominals?
- c. Do children have the same interpretation as adults do?

To answer these questions, a spontaneous speech study and an experimental study were conducted to examine the acquisition of noun phrases in Mandarin. The spontaneous speech study focused on children's production of nominals, while the experimental study focused on their interpretation of the nominals.

With respect to production, in adult Mandarin, the distribution of different nominal types is not completely without restrictions. Some nominal types can only appear in certain syntactic positions and with certain semantic interpretations. The hypothesis is that if children's speech production reflects the target language, according to the distinct property of Mandarin bare and non-bare nominals, it is expected that when acquiring nominals, Mandarin-speaking children will systematically produce nominals

with particular structures, in particular constructions, and with particular interpretations. Given that no known research has studied the emergence of non-bare nominals in Mandarin, predictions in the production study are mainly based on the universal nominal structure (Cinque 2005), the acquisition path of nominals proposed by Roeper (2006), and the grammar of Mandarin nominals (Huang et al. 2009). Most of the predications have been supported.

With respect to interpretation, Mandarin bare nominals are ambiguous between the generic and existential definite reading in the subject position, while Mandarin demonstrative nominals allow only the existential definite reading. Therefore, the hypothesis is that if Mandarin-speaking children's grammar is adult-like, they will have both generic and definite interpretations for subject bare nominals and only existential definite interpretation for demonstrative nominals. Given that no known research has studied the interpretation of bare and demonstrative nominals in Mandarin, predictions in the interpretation study are based on several previous studies about nominals in other languages. Based on these studies, two competing predictions can be made: (i) Chierchia (1998), de Villiers and Roeper (1995), and Pérex-Leroux et al. (2004) – young children allow generic reading for Mandarin nominals; (ii) Maratsos (1976), Karmiloff-Smith (1979), and Crain et al. (1994) – young children do not allow generic reading for Mandarin nominals. The first predication is supported by the study.

6.2 Summary and Concluding Remarks

The research presented in this dissertation collected empirical data on the production of Mandarin-speaking children from existent corpora (Chang 2002). It also

collected experimental data on the comprehension of bare and non-bare nominals by Mandarin-speaking adult and children. The findings of the production data will be summarized first, followed by that of the comprehension data.

The production data consist of spontaneous speech data produced by two Mandarin-speaking children: Nana (2;0-2;6) and Didi (2;10-3;3). The results show that four variables – MLU, Interpretation, Verb Type, and Aspect Marker – have significant effects on the variation of bare and non-bare nominals. The other three variables, Age, Syntactic Position, and Clause Force, do not have significant effects. Each variable (a.k.a. factor group) that had a significant effect was further scrutinized to understand which factor in that group triggers the production of non-bare nominals. The factors that are associated with the production of non-bare nominals are: longer MLU (over 3.0), nominals with existential interpretation, sentences with copular or resultative verbs, and bare verbs without aspect markers. The factors that are not associated with the production of non-bare nominals are: shorter MLU (under 3.0), nominals with generic or predicative interpretation, sentences with activity verbs, and sentence with aspect markers.

There are more findings about the nominal itself. (1) Adjective, which occupies the adjunct position of NP in a syntactic tree, is the first element children add to a Mandarin bare noun root (age 2;0). (2) Possessives nominals, one of the nominal defaults according to Roeper, emerge as early as age 2;1. (3) Although classifiers, numerals, and demonstratives all first appear at age 2;2, the frequency-order (from the most frequent to the least frequent) is – the general classifier *ge*, numerals, and demonstratives. (4) Although children produce non-bare nominals in various syntactic positions, non-bare nominals most frequently associate with the object position – a total of 44% non-bare nominals with an identifiable syntactic position appear as objects. (5) The interpretation

with which non-bare nominals are most frequently associated is the existential interpretation. A total of 74% non-bare nominals have existential interpretation.

These findings are generally consistent with my predictions and with the linguistic and acquisition assumptions this research is built on: Borer (2005), Chang (2007), Cinque (2005), Huang et al. (2009), and Roeper (2006).

The comprehension study examined data from 63 children (between age three and five) and 47 adults. The study finds that nominal types (bare or demonstrative) and age have significant effects on the nominal interpretation. The overall findings show that (1) Mandarin-speaking children, like adults, have both generic and existential definite readings for subject bare nominals. (2) They distinguish between bare and demonstrative nominals by interpreting them differently – bare nominals receive more generic interpretations. (3) They prefer generic readings for bare nominals as opposed to the existential definite reading. Demonstrative nominals also receive a considerable amount of generic interpretations. (4) At least two factors significantly affect how Mandarin-speaking children interpret bare and demonstrative nominals. The first factor is pragmatic – presentation order of the nominal (immediate- or delayed-mention). Sentences in delayed mention receive more generic readings. The other factor is a non-linguistic property of the predicates – canonicity or non-canonicity. Canonical sentences receive more generic readings.

The fact that there is a high percentage of generic interpretation of bare nominals in child Mandarin supports the prediction following Chierchia's theory (1998) that children will allow generic reading of bare nominals in Mandarin. The Semantic Subset Principle by Crain et al. (1994) is not supported by the data.

As for the non-target generic reading of demonstrative nominals, the individual analysis shows that those who assign generic responses to demonstratives also assign generic responses to bare nominals. A positive correlation exists between the generic reading for both nominal types among all adults and children in the study (the correlation is significant among all child groups). The non-target generic interpretation may also be accounted for semantically and syntactically: (1) based on Pérez-Leroux et al (2004), Mandarin-speaking children treat demonstratives as a less-specified determiner like the Spanish definite; (2) based on de Villiers and Roeper (1995), Mandarin-speaking children project demonstrative nominals as NP, not DP.

From the findings of the production and interpretation studies, it is suggested that children's production of demonstratives will precede the target interpretation of the demonstratives. While a Mandarin-speaking child may produce the demonstrative nominals in her own speech as early as at age 2;2, Mandarin-speaking three- to five-year-olds in general do not always associate the demonstrative nominals with the adult-like existential definite interpretation. Rather, in the given context the children associate the demonstrative nominals with the generic interpretation, which is a possible interpretation of English and Spanish definite nominals.

For future research, it is critical to examine the data of older Mandarin-speaking children in order to scrutinize the full acquisition path of the nominal structure. For example, the quantifiers produced by the two subjects, a two-year-old and a three-year-old, were very few. There are D-Num-C-N examples in the data but no Q-Num-C-N sequences. Therefore, more tokens are needed to understand the acquisition of determiners and quantifiers in child speech. Data produced by older children are also desired to examine whether there is a mass/count distinction in Mandarin nominals.

Cheng and Sybesma (1999) argue that the mass/count distinction is manifested at the classifier level – the classifiers for count nouns (count-classifiers) and those for mass nouns (massifiers) are different. Only limited count-classifiers and massifiers were found in the data. The production data of older children will be needed to scrutinize this argument.

In the interpretation study, a preference for the generic interpretation is found for bare nominals and the children also assigned the non-target generic interpretations to demonstrative nominals about half of the time. Pérez-Leroux et al. (2004) found that the percentage of generic interpretations is much smaller in the past tense study than in the present tense study. Mandarin is a tense-less language. It is important to investigate whether the presence of the perfective aspect marker will block the generic interpretation and result in the existential definite interpretation for both bare and demonstrative nominals (the only possible interpretation when the perfective marker is present).

APPENDICES

APPENDIX I

DATA TRANSCRIPTION AND CALCULATION OF MEAN LENGTH OF UTTERANCES

The study of noun phrase production in this dissertation used spontaneous speech data from the Chang corpora (2002). The following information regarding the transcription of the data and the calculation of mean length of utterances (MLU) is mainly adopted from Chang (2002).

For research that focuses on lexical and morpho-syntactic rather than phonological or phonetic issues, usually a broad phonemic (rather than narrow phonetic) transcription is used. For instance, Brown (1973) used the orthographic conventions of English for transcription. Chang (2002) used *Pinyin*, the official Mandarin Romanization system used in China, to transcribe the data along with the transcription format of the Codes for the Human Analysis of Transcripts (CHAT) used in the Child Language Data Exchange System (CHILDES) (MacWhinney 2000).

In a tonal language such as Mandarin, tonal information is obligatory to capture lexical or syntactic information.⁵⁶ The numbers 1, 2, 3, 4 are used immediately following each syllable to mark the level tone, raising tone, fall-raising tone, and falling tone, respectively. A syllable without tonal indication has a neutral tone. There is a phonological phenomenon called Tone Sandhi (Li and Thompson 1981) in Mandarin, which depicts the change of tones when syllables are put together. For instance, if two consecutive syllables both possess the fall-raising tone, the first one will switch to a

⁵⁶ The tone of a grammatical morpheme in Mandarin is usually neutral.

raising tone, such as *shui3jiao3* ‘water dumpling’ will become *shui2jiao3*. Tone Sandhi is not marked in the transcription. The original tone of the syllable is used instead.

In the Chang corpora, the mean length of utterances (MLU)^{57, 58} is counted basically in terms of morphemes, rather than words or Mandarin characters, which is one character for one syllable.⁵⁹ Chang followed the commonly accepted notion that a morpheme is ‘the smallest meaningful linguistic unit’ (Cairns 1996) while at the same time utilizing particular coding methods to better match her research focus. For instance, instead of assuming that children possess adult’s knowledge of morphology, she took a more conservative way and followed the suggestion from the CHAT Manual (p.52 and 161) (included in MacWhinney 2000) to treat some compounds, such as *bathing-suit*, *High-street*, or *Santa-Claus*, as one morpheme. Accordingly, the English word *Santa-*

⁵⁷ Erbaugh (1978:32) states that ‘MLU measurements for Mandarin are about as valid as those for English.’ She also points out that MLU measurements may underestimate the language development of Mandarin children because Mandarin lacks grammatical morphologies, such as tense, gender, and case. Her assertion is that Mandarin children may be mis-identified to belong to a lower stage of language development because of the seemingly shorter MLU. Although her argument seems reasonable, it is purely conceptual and is not based on any empirical evidence.

⁵⁸ Devescovi et al. (2005) studied two groups of children, 233 English-speakers and 233 Italian speakers, and asked their parents to report the three longest utterances produced by each child. They first made the corrected versions for each sentence, which they assume to be ‘the target’ utterances, and then use four different ways to code both the actual child utterances and the target sentences. Italian has a richer morphological system than English does because Italian marks gender, case, etc.; thus, it is not surprising that Italian children have longer MLUs on all measures. But, they found that the ratio of actual MLUs to target MLUs do not differ between Italian and English. This implies that although languages are different and there are different ways to code speech data, MLU remains valid in child language studies.

⁵⁹ Note that Nana and Didi were two and three year-olds respectively at the time of recording and they had almost no knowledge of the Mandarin writing system. In other words, the disyllabic word *pingguo* ‘apple’ has two characters in the writing system, but for toddlers and preschoolers, *pingguo* does not correspond to two characters in their brain instead it is simply a disyllabic word as the English word *apple*.

Claus in (1a) was treated as one morpheme because the child used it as an unanalyzed chunk. On the contrary, the Mandarin counterpart of *Santa-Claus*, *shengdan laogonggong* ‘Christmas old-man,’ was coded as two morphemes because the child was able to produce both *shengdan laogonggong* (1b) and *laogonggong* (1c) in the same recording. This implies her competence of separating *shengdan* ‘Christmas’ and *laogonggong* ‘old-man’ as two morphemes.

- (1) a. uhuh, Santa-Claus (Nana: 2;3.17)
1 (length of utterance: 1, line 813)
- b. Shengdan laogonggong
1 2 (length of utterance: 2, line 917)
Christmas old-man
'Santa Claus.'
- c. laogonggong zai shuijiao (length of utterance: 3, line 996)
1 2 3
old-man PROG sleep
'The old man is sleeping.'

Regarding the morpheme counts for the Resultative Verb Compounds, she assumes Li and Thompson's argument (1981:45-46) that although a compound is like a single word referring to a single object or action, it can be analyzed into two or more meaningful morphemes. In example (2a), the two elements of the Resultative Verb Compound *po-diao* (break-drop) 'break-off' can both serve as matrix verbs, *po* 'break' in (2b) and *diao* 'drop' in (2c); thus, *po-diao* is clearly a two-morpheme compound. Also, grammatical morphemes, such as the aspect marker *-le* and the nominal classifier *-ge* in (2a), are counted as single morphemes.

- (2) a. zhe-ge po-diao-le. (Nana: 2;1.24)
 1 2 3 4 5
 this-CL break-drop-PERF
 ‘This (thing) broke off.’
 (length of utterance: 5, line 2944)

- b. zhe-bian shi po-le.
 1 2 3 4 5 (length of utterance: 5, line 3978)
 this-side be break-PERF
 ‘It is true that this side has broken.’
 (The copula *shi* can be used to express ‘it is true that...’ (Li and Thompson 1981:151).)
- c. diao-le.
 1 2 (length of utterance: 2, line 4563)
 drop-PERF
 ‘(something) dropped.’

Besides, based on Brown’s rules for calculating MLU in the English data (1973:54), Chang applied some rules, as summarized below, when transcribing the Mandarin data to avoid inaccurate counts of MLU.

(3) Principles for calculating MLU

- a. To avoid stammering that makes the utterance longer, identical consecutive words are coded as one word by either following the CHAT program to code single word repetition (e.g. *ni3 ni3 ni3* ‘you you you’ is coded as *ni3(*3)* for three instances of ‘you’) or write a note in the comments (a separate line in the database) about how many times a multi-word string repeated (e.g. *na4-ge ren2 na4-ge ren2* ‘that-CL man that-CL man’ is coded as *na4-ge ren2* and the notes in the comment line indicates the times this phrase was produced).
- b. For cases where utterances are interrupted by another speaker, they were transcribed in the way that will be counted conservatively as separate utterances.
- c. For unintelligible words (usually because of the acoustics) a special mark, i.e., xxx, is used to indicate that the child said something unintelligible. The Computerized Language Analysis (CLAN) program of CHILDES will exclude xxx strings when calculating MLU.
- d. Discourse fillers are excluded, such as *eh*, *hom*, *uhuh*, *oh*, *mmm*, *yi2*, *ha2*, *a*, *ye*, *ya*, *hey*, *gosh*, etc. These fillers were transcribed with a ‘&’ sign and the MLU count in CLAN will exclude them.
- e. Songs and poems are transcribed as notes and are excluded from MLU. Numbers, such as one-hundred-eighty-four, are counted as one morpheme. (In these cases, consequently, the length of utterance will be shorter than that in reality.)

APPENDIX II

DATA EXTRACTING AND CRITERIA FOR ELIMINATING UNWANTED SENTENCES

In the traditional diary studies of child speech, researchers tended to be attracted by the interesting sentences and ignoring other utterances; therefore, Snyder (2007) suggests extracting the target sentences for study without reading the original scripts to avoid compromising the scientific integrity. At first, I created a list of nominals of a transcription file by skimming through the auto-generated word list and identified all the nominals by hand. Then I extracted all the utterances that include at least a noun in the noun list using a function of the Computerized Language Analysis (CLAN) program. These steps worked well for the utterances produced by Nana (2;0.8) in file nana01, in which she used only 129 different word types (273 tokens). However, these steps become more difficult to follow after Nana became older, e.g. she produced 279 different word types in file nana04 (2;1.24). When skimming through such a long list, errors become almost inevitable in judging which word is a noun and which is not. Given the difficulty of identifying the nominals without consulting the original transcript, I have to isolate the target sentences by directly going through the transcript, which is a method discouraged by Snyder.

In the process of coding, some sentences and noun tokens will be eliminated for the study according to the criteria listed below. Examples and exceptions are included under each criterion.

(1) Eliminated utterances:

a. Unclear sentences - Unintelligible words in the clauses

Nana02:1707

Nana: *bu2 shi4 &da4yu2*

not be ? fish

(It could be 'not fishing' or 'no big fish'.)

Nana02:2168

Investigator: *ni3-de shui4yi1*

your sleeping:wear

Nana: *xxx shui4yi1*

? sleeping:wear

(It is not sure whether this case is repetition or not.)

b. Immediate un-interrupted repetitions of the whole sentence of the same speaker will be counted only once.

c. Immediate repetitions of the whole or partial sentence of others (not interrupted by meaningful sentences)

- But creative use of already mentioned noun will be kept

Nana01:1276

Investigator: *xiang4 yi1-ge niao3-dedu4zi*

like one-CL bird-DE stomach

'like the stomach of a bird.'

Nana: *bu2 shi4 niao3*

not be bird

'not a bird'

d. Kinship terms⁶⁰

- But *agong* 'grandpa' and *ama* 'grandma' (these two terms were pronounced in Taiwanese, not Mandarin) were kept because the children used these two terms to refer to an American couple and they usually used these two terms with rich modifiers.
- Bare kinship terms are ambiguous between common nouns and proper names and thus are deleted. Modified kinship terms are kept because the kinship terms are clearly used as head nouns.

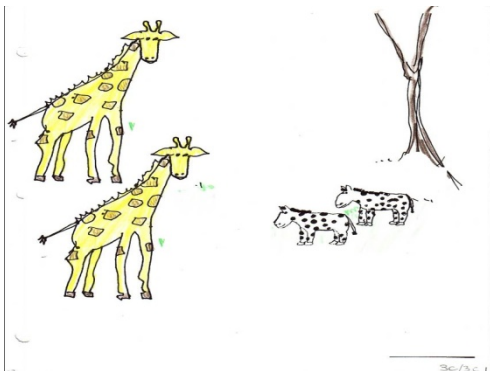
e. Proper names, Pronominals, Isolated English Alphabets


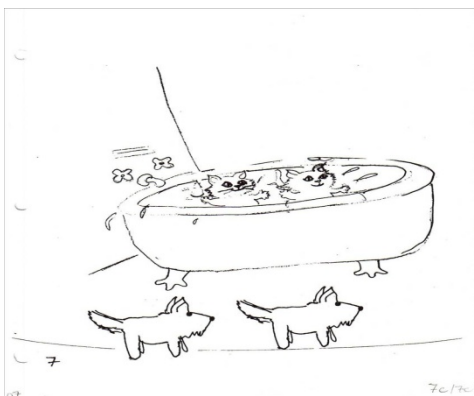
⁶⁰ Although kinship terms are common nominals and we may expect these terms appear in a complex noun phrase construction, such as *Mary's beautiful sister who lives in New York*; however, these terms are usually used as proper names in child language and are eliminated from the study. The eliminated kinship terms includes *ma/mama* 'mom,' *ba/baba* 'dad,' *jie/jiejie* 'older sister,' *mei/meimei* 'younger sister,' *ayi* 'aunt.'

Including the tokens and utterances that have any of the properties above will make the amount of noun tokens very large and results in skewed data – for example, the inclusion of proper names will result in large amount of bare nominals. Therefore, these tokens and utterances are not included in the analysis of the current study.

APPENDIX III


TARGET QUESTION SETS FOR THE COMPREHENSION EXPERIMENT

Version A	Version B
<p>(1) 你看這兩隻斑馬, 牠們有斑點, 長頸鹿在想為什麼牠們看起來不一樣, 現在讓我問你: Look at these two zebras. They are spotted. The giraffes wonder why they look different. Now let me ask you:</p> <p>斑馬有沒有斑點? Do zebras have spots?</p> <p>斑馬是黑白的嗎? Are zebras black and white?</p> <p>你有沒有斑點? Do you have spots?</p> <p>這些斑馬有沒有條紋? Do these zebras have stripes?</p>	<p>Figure AIII.1 Question Set 1</p>  <p>這些斑馬有沒有斑點? Do these zebras have spots?</p> <p>這些斑馬是黑白的嗎? Are these zebras black and white?</p> <p>你有沒有斑點? Do you have spots?</p> <p>斑馬有沒有條紋? Do zebras have stripes?</p>

Version A	Version B
<p>(2) 這裡有兩隻猴子,牠們喜歡吃草,這些小鳥覺得牠們很奇怪,因為牠們不喜歡香蕉,現在讓我問你: Here are two monkeys. They love eating grass! The birds think they are strange because they don't like bananas. Now let me ask you:</p> <p>這些猴子吃草嗎? Do these monkeys eat grass?</p> <p>這些猴子有尾巴嗎? Do these monkeys have tails?</p> <p>腳踏車有尾巴嗎? Does a bicycle have a tail?</p> <p>猴子吃香蕉嗎? Do monkeys eat bananas?</p>	<p>Figure AIII.2 Question Set 2</p>  <p>猴子吃草嗎? Do monkeys eat grass?</p> <p>猴子有尾巴嗎? Do monkeys have tails?</p> <p>腳踏車有尾巴嗎? Does a bicycle have a tail?</p> <p>這些猴子吃香蕉嗎? Do these monkeys eat bananas?</p>
<p>(3) 你看,這些在浴缸裡的貓咪,我猜這些狗狗覺得牠們很奇怪,因為牠們不怕弄濕,現在讓我問你: Look at the cats in the bathtub! I bet the dogs think they're weird because they don't mind getting wet. Now let me ask you:</p> <p>貓咪喜歡洗澡嗎 Do cats like taking baths?</p> <p>貓咪有天線嗎 Do cats have antennas?</p> <p>你喜歡洗澡嗎 Do you like taking baths?</p> <p>這些貓咪喜歡乾乾的嗎 Do these cats like staying dry?</p>	<p>Figure AIII.3 Question Set 3</p>  <p>這些貓咪喜歡洗澡嗎 Do these cats like taking baths?</p> <p>這些貓咪有天線嗎 Do these cats have antennas?</p> <p>你喜歡洗澡嗎 Do you like taking baths?</p> <p>貓咪喜歡乾乾的嗎 Do cats like staying dry?</p>

Version A	Version B
<p>(4) 這兩隻老虎只吃青菜, 你看牠們在吃紅蘿蔔, 這些兔子很高興, 因為老虎不會吃牠們, 現在讓我問你: These two tigers only eat vegetables. See, they're eating carrots. The rabbits are happy because They won't get eaten. Now let me ask you:</p> <p>這些老虎吃肉嗎? Do these tigers eat meat?</p> <p>這些老虎有條紋嗎? Do these tigers have stripes?</p> <p>你有條紋嗎? Do you have stripes?</p> <p>老虎吃紅蘿蔔嗎? Do tigers eat carrots?</p>	<p>Figure AIII.4 Question Set 4</p>  <p>老虎吃肉嗎? Do tigers eat meat?</p> <p>老虎有條紋嗎? Do tigers have stripes?</p> <p>你有條紋嗎? Do you have stripes?</p> <p>這些老虎吃紅蘿蔔嗎? Do these tigers eat carrots?</p>
<p>(5) 你看到的跟我看到的一樣嗎? 這兩隻馬少了一條腿, 我猜綿羊覺得牠們很可憐, 因為牠們沒有四條腿, 現在讓我問你: Do you see what I see? These two horses, are missing a leg! I bet the lambs feel sorry for them because they don't have four legs. Now let me ask you:</p> <p>馬有四條腿嗎? Do horses have four legs?</p> <p>馬在草地上嗎? Are horses in a pasture?</p> <p>你騎過馬嗎? Have you ridden a horse?</p> <p>這些馬有三條腿嗎? Do these horses have three legs?</p>	<p>Figure AIII.5 Question Set 5</p>  <p>這些馬有四條腿嗎? Do these horses have four legs?</p> <p>這些馬在草地上嗎? Are these horses in a pasture?</p> <p>你騎過馬嗎? Have you ridden a horse?</p> <p>馬有三條腿嗎? Do horses have three legs?</p>

Version A	Version B
<p>(6) 這兩隻小鳥住在洞裡面, 我猜貓頭鷹從牠們的巢看下來可能在想, 哇, 那個洞裡面一定很黑, 現在讓我問你: These two birds live in a cave. I bet the owls look down from their nests and think, “Wow! It must be dark in that cave!” Now let me ask you:</p> <p>這些小鳥住在鳥巢裡嗎? Do these birds live in a nest?</p> <p>這些小鳥有羽毛嗎? Do these birds have feathers?</p> <p>你的貓咪有沒有羽毛? Does your cat have feathers?</p> <p>小鳥住在洞裡嗎? Do birds live in a cave?</p>	<p>Figure AIII.6 Question Set 6</p>  <p>小鳥住在鳥巢裡嗎? Do birds live in a nest?</p> <p>小鳥有羽毛嗎? Do birds have feathers?</p> <p>你的貓咪有沒有羽毛? Does your cat have feathers?</p> <p>這些小鳥住在洞裡嗎? Do these birds live in a cave?</p>
<p>(7) 你看這兩隻龍, 這些青蛙覺得牠們不像別的龍會噴火, 反而噴出泡泡來, 很奇怪, 現在讓我問你: Look at these two dragons. The frogs think it's strange that they breathe bubbles out of their mouths instead of breathing fire like other dragons. Now let me ask you:</p> <p>這些龍會噴泡泡嗎? Do these dragons breathe bubbles?</p> <p>這些龍可不可怕? Are these dragons scary?</p> <p>小貝比可不可怕? Is a baby scary?</p> <p>龍會噴火嗎? Do dragons breathe fire?</p>	<p>Figure AIII.7 Question Set 7</p>  <p>龍會噴泡泡嗎? Do dragons breathe bubbles?</p> <p>龍可不可怕? Are dragons scary?</p> <p>小貝比可不可怕? Is a baby scary?</p> <p>這些龍會噴火嗎? Do these dragons breathe fire?</p>

Version A	Version B
<p>(8) 這兩隻獅子住在一艘大船上, 這些烏龜覺得很好笑, 因為牠們不住在草原裡, 現在讓我問你:</p> <p>These two lions live on a big boat. The turtles think it's funny that they don't live in the savannah. Now let me ask you:</p> <p>獅子住在草原裡嗎? Do lions live in the savannah?</p> <p>獅子有鬃毛嗎? Do lions have manes?</p> <p>我們說人有鬃毛還是人有頭髮? Do we say people have manes, or do we say people have hair?</p> <p>這些獅子住在船上嗎? Do these lions live on a boat?</p>	<p>Figure AIII.8 Question Set 8</p>  <p>這些獅子住在草原裡嗎? Do these lions live in the savannah?</p> <p>這些獅子有鬃毛嗎? Do these lions have manes?</p> <p>我們說人有鬃毛還是人有頭髮? Do we say people have manes, or do we say people have hair?</p> <p>獅子住在船上嗎? Do lions live on a boat?</p>

APPENDIX IV

TARGET QUESTIONS FOR THE KNOWLEDGE TEST

1. 斑馬有沒有斑點? Do zebras have spots?
2. 斑馬有沒有條紋? Do zebras have stripes?
3. 猴子吃草嗎? Do monkeys eat grass?
4. 猴子吃香蕉嗎? Do monkeys eat bananas?
5. 貓咪喜歡洗澡嗎? Do cats like taking baths?
6. 貓咪喜歡乾乾的嗎? Do cats like staying dry?
7. 老虎吃肉嗎? Do tigers eat meat?
8. 老虎吃紅蘿蔔嗎? Do tigers eat carrots?
9. 馬有四條腿嗎? Do horses have four legs?
10. 馬有三條腿嗎? Do horses have three legs?
11. 小鳥住在鳥巢裡嗎? Do robins live in a nest?
12. 小鳥住在洞裡嗎? Do robins live in a cave?
13. 這些龍會噴泡泡嗎? Do dragons breathe bubbles?
14. 龍會噴火嗎? Do dragons breathe fire?
15. 獅子住在草原裡嗎? Do lions live in the savannah?
16. 獅子住在船上嗎? Do lions live on a boat?

REFERENCES

REFERENCES

- Abbott, Barbara. 2004. Definiteness and Indefiniteness. *Handbook of Pragmatics*, ed. by Laurence R. Horn and Gregory Ward, 122-147, Oxford: Blackwell.
- Abbott, Barbara. 2010. *Reference*. Oxford: Oxford University Press.
- Abney, Steven. 1987. The English Noun Phrase in Its Sentential Aspect. Cambridge, MA: Massachusetts Institute of Technology doctoral dissertation.
- Alexiadou, Artemis, Liliane Haegeman, and Melita Stavrou. 2007. *Noun phrase in the generative perspective*. Berlin: Mouton de Gruyter.
- Bayley, Robert. 1994. Consonant cluster reduction in Tejano English. *Language Variation and Change* 6.303-323.
- Bayley, Robert, and Richard Young. (forthcoming). VarbRul: A special case of logistic regression. *Linguistic Data Computation*, ed. by Dennis R. Preston and Robert Bayley.
- Berwick, Robert. 1985. *The acquisition of syntactic knowledge*. Cambridge, MA: MIT Press.
- Blumenfeld, Ann. 1999. Acquisition of Compounds and Home. Haverford, ms.
- Borer, Hagit. 2005. *In name only*. Oxford: Oxford University Press.
- Brown, Roger. 1973. *A First Language*. Cambridge, MA: Harvard University Press.
- Cairns, Helen Smith. 1996. *The acquisition of language*. 2nd edition. Austin, TX: Pro-Ed inc.
- Carlson, Greg. 1977. *Reference to kinds in English*. Amherst, MA: University of Massachusetts at Amherst doctoral dissertation. Published 1980, by Garland Press, New York.
- Carlson, Greg, and Francis Jeffry Pelletier. (eds.) 1995. *The Generic Book*. Chicago: University of Chicago Press.
- Chang, Hsiang-Hua. 2002. Child acquisition of the aspect marker –le in Mandarin Chinese. East Lansing, MI: Michigan State University master's thesis.
- Chang, Hsiang-Hua. 2007. Mandarin determinerless classifier-nouns. Joint conference of the 15th Annual Meeting of the International Association of Chinese Linguistics and the 19th Annual North American Conference on Chinese Linguistics (IACL-15/NACCL-19), Columbia University.

- Chang, Hsiang-Hua, and Cristina Schmitt. 2008. Semantic Mappings of Bare and Demonstrative Nominals in Child Mandarin. The 20th Annual North American Conference on Chinese Linguistics (NACCL 20), Ohio State University.
- Chao, Yuan Ren. 1968. *A grammar of spoken Chinese*. Cambridge: Cambridge
- Chen, Ping. 2004. Identifiability and definiteness in Chinese. *Linguistics* 42.1129-1184.
- Cheng, Lai-Shen Lisa, and Rint Sybesma. 1998. Yi-wan tang, yi-ge tang: Classifiers and massifiers. *The Tsing Hua Journal of Chinese Studies* 28.385–412.
- Cheng, Lai-Shen Lisa, and Rint Sybesma. 1999. Bare and not-so-bare nominals and the structure of NP. *Linguistic Inquiry* 30.509-542.
- Chien, Yu-Chin, Barbara Lust, and Chi-Pang Chiang. 2003. Chinese Children's Comprehension of Count-Classifiers and Mass-Classifiers. *Journal of East Asian Linguistics* 12.91-120.
- Chien, Yu-Chin, and Ken Wexler. 1987a. Children's acquisition of reflexives and pronouns. *Papers and Reports on Child Language Development* 26.30-39.
- Chien, Yu-Chin, and Ken Wexler. 1987b. A comparison between Chinese-speaking and English-speaking children's acquisition of reflexives and pronouns. Paper presented at the 12th Annual Boston University Conference on Language Development (BUCLD), Boston University.
- Chien, Yu-Chin, and Ken Wexler. 1989. Children's knowledge of relative scope in Chinese. *Papers and Reports on Child Language Development* 28. 72-80.
- Chien, Yu-Chin, and Ken Wexler. 1990. Children's knowledge of locality conditions in binding as evidence for the modularity of syntax and pragmatics. *Language Acquisition* 1. 225-295.
- Chierchia, Gennaro. 1984. *Topics in the syntax and semantics of infinitives and gerunds*. Amherst, MA: University of Massachusetts at Amherst doctoral dissertation. Published 1988, Garland Press, New York.
- Chierchia, Gennaro. 1995. Individual-level Predicates as Inherent Generics. *The generic book*, ed. by Gregory Carlson and Francis Jeffry Pelletier, 176-223. Chicago: University of Chicago Press.
- Chierchia, Gennaro. 1998. Reference to kinds across languages. *Natural Language Semantics* 6.339-405.
- Chierchia, Gennaro, Maria T. Guasti, and Andrea Gualmini. 1999. Nouns and articles in child grammar and the syntax/semantics map. Paper presented at Generative Approach to Language Acquisition (GALA), Potsdam, Germany.

- Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
- Chomsky, Noam. 1986. *Knowledge of language: its nature, origin and use*. New York: Praeger.
- Chomsky, Noam. 1988. *Language and problems of knowledge*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: MIT Press.
- Cinque, Guglielmo. 2005. Deriving Greenberg's Universal 20 and its exceptions. *Linguistic Inquiry* 36.315-332.
- Contreras, Heles. 1986. Spanish bare NPs and the ECP. *Generative studies in Spanish*, ed. by Ivonne Bordelolís, Heles Contreras, and Karen Zagona, 25-49. The Netherlands: Foris Publications.
- Crain, Stephen, Weijia Ni, and Laura Conway. 1994. Learning, parsing and modularity. *Perspectives on sentence processing*, ed. by Charles Clifton, Jr., Lyn Frazier, and Keith Rayner, 443-467. Hillsdale, NJ: Lawrence Erlbaum Associates.
- De Cat, Cécile. 2002. French Dislocation. York, UK: University of York doctoral dissertation.
- de Villiers, Jill, and Tom Roeper. 1995. Barriers, binding and acquisition of the NP/DP distinction. *Language Acquisition* 4.76-104.
- Demuth, Katherine. 1996. Collecting Spontaneous Production Data. *Methods for Assessing Children's Syntax*, ed. by Dana McDaniel, Cecile McKee, and Helen Smith Cairns, 3-22. Cambridge, MA: MIT Press.
- Devescovi, Antonella, Maria Cristina Caselli, Daniela Marchione, Patrizio Pasqualetti, Judy Reilly, and Elizabeth Bates. 2005. A cross-linguistic study of the relationship between grammar and lexical development. *Journal of Child Language* 32.759-786.
- Diesing, Molly. 1992. *Indefinites*. Cambridge, MA: MIT Press.
- Erbaugh, Mary. 1978. Acquisition of temporal and aspectual distinctions in Mandarin. *Papers and reports on child language development* 15.30-37.
- Erbaugh, Mary. 1992. The acquisition of Mandarin. *The crosslinguistic study of language acquisition*, ed. by Dan Slobin, 373-455. Hillsdale, NJ: Lawrence Erlbaum.
- Fillmore, Charles. 1992. "Corpus linguistics" or "Computer-aided armchair linguistics"?

- Directions in Corpus Linguistics, Proceedings of Nobel Symposium 82, Stockholm, 4-8 August 1991*, ed. by Jan Svartvik, 35–66. Berlin: Mouton de Gruyter.
- Fodor, Janet D. 1998. Learning to parse? *Journal of Psycholinguistic Research* 27.285-318.
- Gelman, Susan, and Lakshmi Raman. 2003. Preschool children use linguistic form class and pragmatic cues to interpret generics. *Child Development* 74.308-325.
- Gelman, Susan, and Twila Tardif. 1998. A cross-linguistic comparison of generic noun phrases in English and Mandarin. *Cognition* 66.215-248.
- Golinkoff, Roberta, and Kathy Hirsh-Pasek. 2000. *How Babies Talk: The magic and mystery of language in the first three years*. New York: Penguin.
- Grice, H. Paul. 1975. Logic and Conversation. *Syntax and Semantics, Vol. 3, Speech Acts*, ed. by Peter Cole and Jerry L. Morgan, 41-58. New York: Academic Press.
- Guasti, Maria T. 2000. An excursion into interrogatives in Early English and Italian. *The acquisition of syntax: studies in comparative developmental linguistics*, ed. by Marc-Ariel Friedemann and Luigi Rizzi, 105-128. New York: Longman.
- Gundel, Jeanette, Nancy Hedberg and Ron Zacharski. 1993. Cognitive Status and the Form of Referring Expressions in Discourse. *Language* 69.274-307.
- Gundel, Jeanette, Nancy Hedberg and Ron Zacharski. 2001. Cognitive Status and Definite Descriptions in English: Why Accommodation is Unnecessary. *English Language and Linguistics* 5.273-295.
- Guy, Gregory. 1980. Variation in the group and the individual: The case of final stop deletion. *Locating language in time and space*, ed. by William Labov, 1-36. New York: Academic Press.
- Heim, Irene, and Angelika Kratzer. 1998. *Semantics and Generative Grammar*. Oxford: Blackwell.
- Hsieh, Miao-Ling. 2008. *The internal structure of noun phrases in Chinese*. Taiwan Journal of Linguistics Book Series in Chinese Linguistics. No. 2. Taipei, Taiwan: Crane Publishing Company.
- Hu, Qian. 1993. The acquisition of Chinese classifiers by young mandarin-speaking children. Boston: Boston University doctoral dissertation.
- Huang, C.-T. James. 1984. On the distribution and reference of empty pronouns. *Linguistics Inquiry* 15.531-574.

- Huang, C.-T. James, Y.-H. Audrey Li, and Yafei Li. 2009. *The Syntax of Chinese*. Cambridge, UK: Cambridge University Press.
- Huang, Shuanfan. 1999. The emergence of a grammatical category definite article in spoken Chinese. *Journal of Pragmatics* 31.77–94.
- Karmiloff-Smith, Annette. 1979. *A functional approach to child language: a study of determiners and reference*. London: Cambridge University Press.
- Keil, Frank. 1989. *Concepts, kinds, and cognitive development*. Cambridge, MA: Bradford book/MIT Press.
- Kratzer, Angelika. 1995. Stage-level and individual-level predicates. *The generic book*, ed. by Gregory Carlson and Francis Jeffry Pelletier, 125-175. Chicago: University of Chicago Press.
- Krifka, Manfred. 1995. Common Nouns: A Contrastive Analysis of English and Chinese, *The Generic Book*, ed. by Gregory Carlson and Francis Pelletier, 398-411. Chicago: Chicago University Press.
- Krifka, Manfred, Francis Pelletier, Gregory Carlson, Alice ter Meulen, Gennaro Chierchia, and Godehard Link. 1995. Genericity: An Introduction. *The generic book*, ed. by Gregory Carlson and Francis Pelletier, 1-124. Chicago: University of Chicago Press.
- Laurence, Stephen, and Margolis, Eric. 2001. The Poverty of the Stimulus Argument. *The British Journal for the Philosophy of Science* 52.217–276.
- Lee, Thomas H.-T. 1986a. Acquisition of Quantificational Scope in Mandarin Chinese. *Papers and Reports in Child Language Development* 25.57-64.
- Lee, Thomas H.-T. 1986b. Studies on quantification in Chinese. Los Angeles, CA: University of California, Los Angeles doctoral dissertation.
- Lee, Thomas H.-T. 1996. Theoretical issues in language development and Chinese child language. *New horizons in Chinese linguistics*, ed. by C. T. James Huang and Y. H. Audrey Li, 293–356. Dordrecht, the Netherlands: Kluwer Academic Publishers.
- Li, Charles, and Sandra Thompson. 1981. *Mandarin Chinese: A functional reference grammar*. Los Angeles, CA: University of California Press.
- Li, Peggy, David Barner, and Becky Huang. 2008. Classifiers as count syntax: Individuation and measurement in the acquisition of Mandarin Chinese. *Language, Learning and Development* 4.1-42.
- Li, Y.-H. Audrey. 1998. Argument determiner and number phrases. *Linguistic Inquiry* 29.693-702.

- Li, Y-H. Audrey. 1999. Plurality in a classifier language. *Journal of East Asian Linguistics* 8.75-99.
- Liao, Wei-wen Roger, and Yu-yun Iris. 2008. Multiple Classifiers Construction and Nominal Expressions in Chinese. Paper presented at the 39th Meeting of the North East Linguistic Society (NELS), Cornell University.
- Liu, Haiyong. 2008. A Case Study of the Acquisition of Mandarin Classifiers. *Language Research* 44.345–360.
- Liu, Haiyong. 2009. The Acquisition of Mandarin Aspects and Modals: Evidence from the acquisition of negation. *Language and Linguistics* 10.133-160.
- Longobardi, Giuseppe. 1994. Reference to kinds and proper names: a theory of N movement in syntax and logical form. *Linguistic Inquiry* 25.609-665.
- Longobardi, Giuseppe. 2001. How comparative is semantics? A unified parametric theory of bare nouns and proper names. *Natural Language Semantics* 9.335-369.
- Longobardi, Giuseppe. 2008. Reference to individuals, person, and the variety of mapping parameters. *Essays on Nominal Determination*, ed. by Henrik Müller and Alex Klinge, 189–211. Philadelphia/Amsterdam: John Benjamins.
- Lü, Shuxiang. 1980. *Xiandai Hanyu babai ci*. [800 Phrases of Modern Chinese]. Hong Kong: Shangwu yingshuguang.
- Lyons, Christopher. 1999. *Definiteness*. Cambridge: Cambridge University Press.
- Lyons, John. 1975. Deixis as the source of reference. *Formal Semantics of Natural Language*, ed. by Edward L. Keenan, 61-83. Cambridge: Cambridge University Press.
- MacWhinney, Brian. 2000. *The CHILDES Project: Tools for Analyzing Talk*. 3rd Edition. Mahwah, NJ: Lawrence Erlbaum Associates.
- Maratsos, Michael. 1976. *The use of definite and indefinite reference in young children*. Cambridge: Cambridge University Press.
- Marcus, Gary. 1993. Negative evidence in language acquisition. *Cognition* 46.53-85.
- Matthewson, Lisa, and Jeannette Schaeffer. 2000. Grammar and pragmatics in the acquisition of article systems. *UCLA Working Papers in Linguistics: Language Development and Breakdown* 1, 5.1-39.
- McNally, Louise. 2009. Properties, entity correlates of properties, and existentials. *Quantification, Definiteness, and Nominalization*, ed. by Anastasia Giannakidou and Monika Rathert, 163-187. Oxford: Oxford University Press.

- Miller, Karen. 2007. Variable input and the acquisition of plurality in two variables of Spanish. East Lansing: Michigan State University doctoral dissertation.
- Munn, Alan, Xiaofei Zhang, and Cristina Schmitt. 2009. The acquisition of plurality in a language without plurality. *Merging Features: Computation, Interpretation and Acquisition* ed. by José Brucart, Anna Gavarró and Jaume Solà, 310-328. Oxford: Oxford University Press.
- Munn, Alan, and Cristina Schmitt. 1999. Bare nouns and the morpho-syntax of number. *Current issues in Romance languages: selected papers from the 29th Linguistic Symposium on Romance Languages (LSRL)*, ed. by Teresa Satterfield, Christina Tortora, and Diana Cresti, 225-240. Amsterdam: John Benjamins.
- Paolillo, John. 2002. *Analyzing Linguistic Variation: Statistical Models and Methods*. Stanford, CA: CSLI Publications.
- Pappas, Athina, and Susan Gelman. 1998. Generic noun phrases in mother-child conversation. *Journal of Child Language* 25.19-33.
- Partee, Barbara. 1987. Noun Phrase Interpretation and Type-Shifting Principles. *Studies in Discourse Representation Theory and the Theory of Generalized Quantifiers*, ed. by Jeroen Groenendijk, Dick de Jongh, Martin Stokhof, and Martin Stokhof, 115-141. Dordrecht: Foris.
- Partee, Barbara. 2006. A note on Mandarin possessives, demonstratives, and definiteness. *Drawing the Boundaries of Meaning: Neo-Gricean Studies in Pragmatics and Semantics in Honor of Laurence R. Horn*, ed. by Betty Birner and Gregory Ward, 263-280. Amsterdam: John Benjamins.
- Pérez-Leroux, Ana Teresa, Alan Munn, Cristina Schmitt, and Michelle DeIrish. 2004. Learning definite determiners: Genericity and definiteness in English and Spanish. Supplementary Proceedings of the 27th Boston University Conference on Language Development (BUCLD).
- Prasada, Sandeep. 2000. Acquiring generic knowledge. *Trends in Cognitive Science* 4.66-72.
- Prince, Ellen. 1992. The ZPG letter: subjects, definiteness, and information-status. *Discourse description: diverse analyses of a fund-raising text*, ed. by William Mann and Sandra Thompson, 295-325. Amsterdam: John Benjamins.
- Pullum, Geoffrey K. and Barbara Scholz. 2002. Empirical assessment of stimulus poverty arguments. *The Linguistic Review* 19.9-50.
- Radford, Andrew. 1990. *Syntactic theory and the acquisition of English syntax: The nature of early child grammar of English*. London: Blackwell.

- Rizzi, Luigi. 1993/94. Some notes on linguistic theory and language development: The case of root infinitives. *Language Acquisition* 3.371-393.
- Rizzi, Luigi. 1994. Early null subjects and root null subjects. *Language acquisition studies in generative grammar*, ed. by Teun Hoekstra and Bonnie Schwartz, 151-176. Amsterdam: John Benjamins.
- Roberts, Julie. 1997. Acquisition of Variable Rules: A Study of (-t,d) Deletion in Preschool Children. *Journal of Child Language* 24.351-372.
- Robinson, J.S., H.R. Lawrence, and Sali A. Tagliamonte. 2001. GOLDVARB 2001: A multivariate analysis application for windows, users' manual.
- Roeper, Thomas. 1999. Universal Bilingualism. *Bilingualism: Language and Cognition* 2. 169-186.
- Roeper, Thomas. 2006. Watching Noun Phrases Emerge: Seeking Compositionality. *Semantics in Acquisition*, ed. by Veerle Van Geenhoven, 37-64. Dordrecht: Springer.
- Sachs, Jacqueline. 1983. Talking about the there and then: The emergence of displaced reference in parent-child discourse. *Children's language*, Vol. 4, ed. by K. E. Nelson, 1-28. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Sankoff, David. 1988. Variable rules. *Sociolinguistics: An International Handbook of the Science of Language and Society*, Vol. 2, ed. by Ulrich Ammon, Norbert Dittmar, and Klaus Mathheier, 984-997. Berlin: Walter de Gruyter.
- Sio, Joanna Ut-seong. 2006. Modification and reference in the Chinese nominal. Leiden, The Netherlands: Universiteit Leiden doctoral dissertation.
- Smith, Jennifer, Mercedes Durham, and Liane Fortune. 2006. Caregiver and child in the acquisition of (socio)linguistic norms in a Scottish dialect. *Proceedings of the 30th annual Boston University Conference on Language Development*, 572-583.
- Stromswold, Karin. 1996. Analyzing children's spontaneous speech. *Methods for assessing children's syntax*, ed. by. Dana McDaniel, Cecile McKee, and Helen Smith Cairns, 23-54. Cambridge, MA: The MIT Press.
- Su, Yi-ching. 2001. Scope and Specificity in Child Language: A Cross-Linguistic Study on English and Chinese. College Park, MD: University of Maryland doctoral dissertation.
- Synder, William. 2007. *Child language: the parametric approach*. Oxford/New York: Oxford University Press.
- Tagliamonte, Sali A. 2006. *Analysing Sociolinguistic Variation*. Cambridge: Cambridge University Press.

- Tao, Hongyin. 1999. The grammar of demonstratives in Mandarin conversational discourse: A case study. *Journal of Chinese Linguistics* 27.69-103.
- Tsai, Wei-Tien Dylan. 2001. On subject specificity and theory of syntax-semantics interface. *Journal of East Asian Linguistics* 10: 129-168.
- Tsao, Feng-fu. 1977. A functional study of topic in Chinese: the first step toward discourse analysis. Los Angeles, CA: University of Southern California.
- Vangsnes, Øystein. 2001. On noun phrase architecture, referentiality, and article systems. *Studia Linguistica* 55.3: 249-299.
- Vendler, Zeno. 1967. *Linguistics and Philosophy*. Ithaca, NY: Cornell University Press.
- Vergnaud, Jean-Roger and Maria Luisa Zubizarreta. 1992. The definite determiner and the inalienable constructions in French and in English. *Linguistic Inquiry* 23: 595-652
- Verkuyl, Henk. 1989. Aspectual Classes and Aspectual Composition, *Linguistics and Philosophy*, 12:1: 39-94, Dordrecht: Springer.
- Wang, Qi, Diane Lillo-Martin, and Catherine Best. 1992. Null subject versus null object: some evidence from the acquisition of Chinese and English. *Language Acquisition* 2: 221-254 .
- Wexler, Ken. 1994. Finiteness and head movement in early child grammar. *Verb Movement*, ed. by Norbert Horstein and David Lightfoot, 305-350. Cambridge: Cambridge University Press.
- Wexler, Ken. 2003. Cues don't Explain Learning: Maximal Trouble in the Determiner System. Ms. MIT.
- Wexler, Ken, and M. Rita Manzini. 1987. Parameter and learnability in binding theory. *Parameter setting*, ed. by Thomas Roeper and Edwin Williams, 41-76. Dordrecht: Kluwer Academic Publishers.
- Wijnen, Frank, and Maaïke Verrips. 1998. The acquisition of Dutch syntax. *The Acquisition of Dutch*, ed. by Steven Gillis and Annick De Houwer, 223-300. Amsterdam and Philadelphia: John Benjamins.
- Wu, Jianxin. 1999. Syntax and Semantics of Quantification in Chinese. College Park, MD: University of Maryland doctoral dissertation.
- Yang, Rong 2001. Common Nouns, Classifiers, and Quantification in Chinese. Newark: The State University of New Jersey doctoral dissertation.
- Young, Richard, and Robert Bayley. 1996. VARBRUL analysis for second language acquisition research. *Second Language Acquisition and Linguistic variation*, ed. by Robert Bayley and Dennis R. Preston, 253-306. Amsterdam and Philadelphia: John Benjamins.