

道可道非常道

Dào kě dào fēicháng dào

'The truth which can be spoken is not eternal truth'.

It was a teacher, friend, and confidant in China, who first opened my eyes to the irrelevance of what is said to a pursuit of what is meant and what is true. It was he who provided me with the first insight into the significance of silence in a social context where any words might cost lives.

This thesis is dedicated to him, and to all those, in China and elsewhere, who keep truth and faith alive in silence.

Unless otherwise acknowledged this thesis is the original work of the author.

A. Curtis.

Contents


A SILENCE MORE ELOQUENT: NP ELLIPSIS IN MANDARIN DISCOURSE

**A DISSERTATION SUBMITTED
TO THE FACULTY OF THE ARTS**

BY A. HELEN CHARTERS

**in partial fulfilment of the requirements for the degree of
MASTER OF ARTS**

*DEPARTMENT OF LINGUISTICS
THE AUSTRALIAN NATIONAL UNIVERSITY
CANBERRA, ACT
February 1996*



Contents

CHAPTER ONE INTRODUCTION	I
1. ELLIPSIS IN MANDARIN	1
1.1 Interpretation	2
2. ORGANISATION OF THE THESIS.....	4
2.1 Three Parameters for Study: Reference, Distribution and Interpretation.....	4
2.2 Overview of Each Chapter.....	5
CHAPTER TWO DISCOURSE DATA	11
1. INTRODUCTION: INTERPRETING ELLIPSIS IN CONTEXT.....	11
1.1 Some Preliminary Considerations	11
1.1.1 Recognition of Ellipsis	11
1.1.2 Valency: the Optional Expression of Semantic Roles	13
1.2 A Definition	14
1.3 Sub-types of Ellipsis	14
2. METHODOLOGY	15
2.1 The Corpus:.....	15
2.2 Linguistic Analysis	15
2.3 Statistical Analysis.....	16
3. DESCRIPTIVE STATISTICS.....	16
3.1 Frequencies of Ellipsis in Different Clause Types	16
4. LINGUISTIC ANALYSIS: DISTRIBUTION AND REFERENCE	17
4.1 Syntagmatic Variation:.....	17
4.2 Paradigmatic Variation:.....	19
4.2.1 Obligatory Ellipsis:.....	19
4.2.2 Optional Ellipsis:.....	20
4.2.3 Contrastive Ellipsis.....	20
4.3 Not Just an Empty Pronoun.....	22
4.4 Referential Variation	23
4.4.1 'Variable' (i.e. contextually sensitive) Reference:.....	23
4.4.2 Bound Co-reference.....	24
4.5 Interaction of Parameters	25
5. STRUCTURAL CONTEXTS: WHERE DOES ELLIPSIS HAPPEN?.....	25
5.1 Structures Involving OPTIONAL Ellipsis:.....	29
5.2 Structures Involving CONTRASTIVE Ellipsis:.....	29
5.3 Structures Involving OBLIGATORY Ellipsis:.....	29
6. MORE STATISTICAL MEASURES	30
6.1 Form and Grammatical Function	30
6.2 Paradigmatic and Referential Sub-types and Grammatical Function.....	30
7. CHAPTER SUMMARY: VARIATIONS IN ELLIPSIS.....	33
CHAPTER THREE ELLIPSIS IN DISCOURSE THEORY	35
1. INTRODUCTION	35
2. REFERENTIAL SUB-TYPES	36
2.1 Undifferentiated Ø	36
3. DISTRIBUTION.....	37
3.1 Conjoinability and Continuity	37
3.2 Topicality	41
3.3 Structural Conjunction	42

4. INTERPRETATION	44
4.1 'Free' Co-reference.....	44
4.2 Neo-Gricean Pragmatics	48
4.2.1 C-command and Governing Categories.....	49
4.2.2 'Preferred' Interpretations in 'Neutral' Contexts.....	53
4.2.3 Lack of Contrast between \emptyset and Pronoun	54
4.2.4 Objects in Sentential Complements	55
4.2.5 Arbitrary Reference when an Antecedent is Available	56
5. SUMMARY.....	57
5.1 Referential Patterns.....	57
5.2 Distribution.....	58
5.3 Interpretation	58

CHAPTER FOUR ELLIPSIS IN SYNTACTIC THEORY61

1. INTRODUCTION	61
2. DEFINITE AND INDEFINITE REFERENCE:.....	62
3. GOVERNMENT AND BINDING THEORY (GB)	65
3.1 Referential Categories.....	65
3.1.1 Anaphors and Pronominals: Definitions	65
3.1.2 PRO and pro	66
3.1.3 Traces	67
3.2 Distribution of GB's Empty Categories	67
3.2.1 Licensing Constraints.....	67
3.3 Interpretation: Mechanisms of Binding and Control.....	68
3.3.1 Binding	68
3.3.2 Control.....	69
3.3.3 Antecedent Binding of Wh-traces.....	72
3.4 GB and Mandarin	72
3.4.1 PRO in Mandarin.....	72
3.4.2 Relative Clauses	76
3.4.3 Generalised Control Theory	79
3.5 Summary of GB.....	87
4. LEXICAL FUNCTIONAL GRAMMAR (LFG).....	88
4.1 Referential Categories.....	88
4.2 Distribution.....	89
4.2.1 PRED 'pro'.....	90
4.3 Interpretation	90
4.3.1 F-command	90
4.3.2 Functional Control	91
4.3.3 Anaphoric Control.....	92
4.3.4 Constituent Control.....	94
4.4 Summary of LFG:	95
5. COMPARISON OF APPROACHES	96
5.1 Referential Patterns.....	96
5.2 Distribution.....	97
5.3 Interpretation and Mechanisms of Control.....	97
5.3.1 What Control Mechanism in what Structure?	97
5.4 Raising in Mandarin.....	98
5.5 Identification of Categories	102
5.6 Reference beyond the Sentence	102
6. CONCLUSION	103
7. MOVING FORWARD:	104

CHAPTER FIVE INTERPRETATION OF DEPENDENT CLAUSES.....107

1. INTRODUCTION: SYNTACTIC AND PRAGMATIC CONTROL.....	107
---	-----

2. CRITERIA FOR DISTINGUISHING SYNTACTIC AND PRAGMATIC CONTROL	108
2.1 Optional Ellipsis and Pragmatic Control	109
2.2 Obligatory Ellipsis and Syntactic Control	109
2.3 Control MECHANISMS for Subjects of Dependent Clauses: The Standard LFG Account.....	110
2.3.1 Tests for Anaphoric vs Functional Control.	110
3. EFFICIENCY OF EXPLANATIONS	115
3.1 The LFG Account: Inconsistent Categories and Additional Principles.....	115
3.2 Alternative Proposal.....	116
3.2.1 Representational Efficiency.....	117
3.2.2 Two Problems: Obviation Effects and Wh-cleft.....	118
3.3 Control Mechanisms Reviewed	119
4. VERB-COMPLEMENT TYPES IN MANDARIN	120
4.1 SCOMPS	121
4.1.1 Apposed Clauses not Complements:	121
4.1.2 Variation within SCOMPS	122
4.2 Absence of Obviation.....	125
4.2.1 Pragmatic Constraints on SCOMP Subject ellipsis	127
5. ANAPHORIC CONTROL RECONSIDERED	128
6. A SEMANTIC CLASSIFICATION OF MANDARIN VERBS	130
6.1 SCOMP Only - Unspecified Agency over Dispreferred Event.	131
6.2 XCOMP Only - Control of Agency	131
6.3 Transitive Verbs.....	132
6.3.1 XCOMP <SUBJ> - Retained Agency	132
6.3.2 XCOMP<OBJ>- Transferred Agency	133
6.3.3 'Dual control' Verbs: {XCOMP <OBJ> / XCOMP<SUBJ>}	133
6.4 Bi-valent Verbs	134
6.4.1 {SCOMP XCOMP} - Unspecified Agency Over Desired Event	134
6.4.2 {SCOMP XCOMP<SUBJ>}	134
6.5 Reduction of Ambiguity in Dual Control Structures.....	136
6.6 Conclusions: Obviation, Semantics and Argument Type.....	138
6.7 Results of WH-cleft Test in Mandarin	139
7. COMPLEMENT STRUCTURES IN THE CORPUS.....	141
7.1 Directional Complements as Compounds	141
7.2 Semantic Sub-Types of Serial Verb Constructions.....	142
7.2.1 Modals.....	142
7.2.2 Cognition and Affect.....	142
7.2.3 The Emphatic Copula.....	144
7.2.4 Motion Verbs	144
7.2.5 Transferred Agency.....	145
7.2.6 Presentatives.....	146
7.3 Functional Complements	147
8. CONCLUSION: SYNTACTIC CONTROL OF DEPENDENT SUBJECTS:	149
9. CHAPTER SUMMARY: STRUCTURES, ARGUMENTS AND CONTROL	149
CHAPTER SIX BOUND REFERENCE.....	153
1. INTRODUCTION: STRUCTURE AND PRAGMATIC CHOICES	153
2. SUBORDINATE STRUCTURES.....	154
2.1 Relativised Constituents	154
2.2 Equative Copulas.....	154
2.3 Subjects of VPs in NP positions	155
2.3.1 VPs in Subject Position	155
2.3.2 VPs in Object and Copula-Complement Positions	156
3. CO-ORDINATE STRUCTURES: CONJUNCTION AND APPPOSITION	158
3.1 Structural Positions of Conjunctions	158
3.1.1 Clausal Conjunctions:.....	158
3.1.2 Paired 'Correlative' Conjunctions	159

3.2 Constituent Conjunctions and Ellipsis Type.....	159
3.2.1 Obligatory Ellipsis: VP Conjunction	160
3.3 Apposition and the Position of the Subject	161
3.4 Contrastive Ellipsis - Conditionals, Cause and Correlation	163
3.4.1 Apposition.....	163
3.4.2 Overt Conjunctions.....	164
3.4.3 Paired Correlative Conjunctions.....	165
3.4.4 Aspectual Subordination	165
3.5 Optional Ellipsis - Concessives, Sequence and Paraphrase	168
3.5.1 Apposition.....	168
3.5.2 Overt Conjunctions:.....	168
3.6 Semantic Motivation for Structure Choice	168
3.6.1 <i>Jiù</i> Semantically Unspecified.....	169
4. BOUND OBJECT ELLIPSIS	170
4.1 Ellipsed Objects of SCOMPs - Contrastive?.....	170
4.2 Optional Object Ellipsis.....	170
4.3 Obligatory Object Ellipsis.....	171
4.3.1 Topic-Comment Sentences.....	171
4.4 Focus-Binding of Objects.....	174
4.4.1 Presentatives and Focus: "There is a lot I don't understand."	174
4.4.2 Topic and Focus.....	174
4.4.3 Conclusions.....	177
5. CHAPTER SUMMARY: THE BOUNDARIES OF SYNTACTIC CONTROL	177
5.1 Alternation of Structures	178
5.2 Frequency of Syntactic vs Pragmatic Control.....	178

CHAPTER SEVEN PRAGMATIC CONSTRAINTS ON OPTIONAL VARIABLE ELLIPSIS.181

1. INTRODUCTION: PRAGMATICS, DISCOURSE STRUCTURE AND INTERPRETATION	181
2. ZERO PRONOMINALS AND ACCESSIBILITY	181
3. OPTIONAL SUBJECT ELLIPSIS OF VARIABLE REFERENCE.....	182
3.1 Subject-Subject Co-reference.....	182
3.2 Exceptions to Subject-Subject Co-reference	184
3.2.1 Foregrounding of Antecedent.....	185
3.3 Backgrounding of Intervening Topics by Ellipsis	187
3.3.1 Ellipsis of Human Subjects.....	187
4. OPTIONAL OBJECT ELLIPSIS	191
4.1 Indefinite Ellipsis	191
4.2 Object-Object Co-reference.....	192
4.3 Exceptions to Object-Object Co-reference	193
4.3.1 Object-Subject Co-reference, No Suitable Object.....	193
4.3.2 <i>Bǎ</i> as a Topicalisation Strategy.....	197
4.3.3 No Structural Antecedent: Semantic cues.....	198
5. CHAPTER SUMMARY: SAME GF CO-REFERENCE AS A COROLLARY OF PRAGMATIC CONTROL	200

CHAPTER EIGHT THE ROLES AND FUNCTIONS OF ELLIPSIS203

1. INTRODUCTION: EXPRESSING RELATIONSHIP.....	203
1.1 The Function of Ellipsis.....	203
1.1.1 Ellipsis as a Marker of Interdependence	203
1.1.2 Levels of Relationship	204
2. THE PRAGMATICS OF REDUCED STRUCTURES	205
2.1 Agency.....	205
2.1.1 Dependent Clauses.....	205
2.1.2 Conjoined Clauses: Contrastive Reference and Switch Agency.....	205
2.1.3 Correlative Conjunctions: Focal Scope and Agency	206
2.2 Purpose and Causal Relationships.....	210

2.3 Similarity of Predications	211
2.4 Summary	212
3. SITUATIONS VACANT	212
3.1 The Distribution of Overt Pronouns	213
3.1.1 Backgrounding to Avoid Ambiguity.....	213
3.1.2 Avoidance of Agent-only sentences.....	213
3.1.3 Lexical Pronouns and Topic Closure.....	214
3.1.4 Position of the Overt Subject in Adjunctive Constructions.....	214
4. CHAPTER SUMMARY: PRAGMATIC FUNCTIONS OF REGULAR DEFINITE ELLIPSIS	216
SUMMARY AND MAJOR CONCLUSIONS.....	217
1. PROBLEMS OF PREVIOUS ACCOUNTS	217
2. AN ALTERNATIVE MODEL	218
2.1 Syntactic Control.....	218
2.2 Pragmatic Control.....	219
2.2.1 Same GF: a Corollary of Topicality and Focus.....	219
2.2.2 Hierarchy of Focus	220
2.2.3 Highlighting Strategies	220
3. APPLICATIONS OF THE MODEL TO COMMON CONSTRUCTIONS.....	220
3.1 Syntactic Control.....	220
3.1.1 Relativised Constituents.....	220
3.1.2 Reduced Structures.....	221
3.1.3 Classification of Mandarin Verbs.....	222
3.2 Pragmatic Control.....	222
3.2.1 Pragmatics and Structural Alternatives	223
3.2.2 Ellipsis and the Chain of Reference	223

LIST OF TABLES

TABLE 1: INTERACTION OF PARAMETERS.....	5
TABLE 2: PATTERNS OF ELLIPSIS IN NATIVE SPEAKER TEXTS	17
TABLE 3: INTERACTION OF PARAMETERS.....	25
TABLE 4: VARIABLE REFERENCE = OPTIONAL ELLIPSIS	25
TABLE 5: BOUND ELLIPSIS IS ALWAYS DEFINITE.....	25
TABLE 6: QUASI-ELLIPSIS: IDIOMATIC NON-ASSIGNMENT OF SUBJECT ROLE	27
TABLE 7: VARIABLE REFERENCE.....	27
TABLE 8: BOUND REFERENCE.....	28
TABLE 9: SUMMARY OF ELLIPSIS TYPES	29
TABLE 10: FREQUENCY OF FORM BY GRAMMATICAL FUNCTION.....	30
TABLE 11: FREQUENCY OF DEFINITE ELLIPSIS FOR SUBJECTS VS OBJECTS	30
TABLE 12: FREQUENCY OF MAJOR SUB-TYPES OF ELLIPSIS IN CORPUS	31
TABLE 13: FREQUENCY OF SUBJECT AND OBJECT ELLIPSIS	32
TABLE 14: FREQUENCY OF VARIABLE VS BOUND ELLIPSIS	33
TABLE 15: CORRELATIVES AND OVERT NPs.....	42
TABLE 16: REFERENTIAL PATTERNS IN 8 SYNTACTIC STRUCTURES.....	52
TABLE 17: GB'S EMPTY CATEGORIES	65
TABLE 18: JUDGEMENTS ON Ø SCOMP OBJECTS:.....	82
TABLE 19: INDIVIDUAL SYSTEMS OF REFERENCE FOR Ø OBJECTS:.....	82
TABLE 20: CONTROL MECHANISMS IN GB AND LFG	98
TABLE 21: COLLOCATIONS OF MANDARIN VERB TYPES.....	121
TABLE 22: VALENCY OPTIONS OF MANDARIN VERB TYPES.....	130
TABLE 23: PROPOSED CONTROL MECHANISMS: MANDARIN AND ENGLISH	139
TABLE 24: SEMANTIC SUB-TYPES IN MANDARIN SERIAL VERB CONSTRUCTIONS	143
TABLE 25: FREQUENCY OF PRAGMATIC AND SYNTACTIC CONTROL IN NP ELLIPSIS	179
TABLE 26: REFERENCE OF OPTIONALLY ELLIPSED SUBJECTS	183
TABLE 27: GF OF ANTECEDENTS FOR SUBJECTS AND OBJECTS.....	183
TABLE 28: GF OF ANTECEDENT FOR HUMAN VS INANIMATE Ø SUBJECTS	184

ABBREVIATIONS USED IN THIS THESIS

AGR	Agreement (as per GB)
ASP (in glosses)	Aspect marker
ASP (in text)	Antecedent Search Procedure (Huang Yan, 1994)
ATT	Attributive Marker
AUX	Auxiliary
CL	Nominal classifier
COMP	Complement
CONJ	Conjunction
COP	Copula
DRP	Disjoint Reference Preferment
EC	Empty Category
GB	Government and Binding Theory
GC	Governing Category (in GB)
GCR	Generalised Control Rule
GF	Grammatical Function
INFL	Inflection (as per GB)
IP	Inflectional Phrase, i.e. Sentence
IRH	Inter-clausal Relations Hierarchy
LFG	Lexical Functional Grammar
MSC	Modern Standard Chinese
NEG	Negator
NP	Noun Phrase
NP _{rel}	Relativised Constituent
NSM	Natural Semantic Meta-language
pl	plural
POSS	Possessive Marker
PP	Prepositional Phrase
pred'PRO'	LFG terminology for a morphologically unexpressed pronoun
pro	'little pro' GB terminology for a \emptyset pronominal
PRO	'big PRO' GB terminology for a \emptyset element that is both pronominal and anaphoric
sg	singular
S	Sentence
SCOMP	Sentential complement
SVO	Subject Verb Object word order
TOP	Topic
VCOMP	Dependent clause
VP	Verb Phrase
VSO	Verb - Subject - Object word order
XCOMP	Dependent clause where no overt Subject is permissible
XCOMP<SUBJ>	As above, where the covert Subject is controlled by the matrix Subject
XCOMP<OBJ>	As above, where the covert Subject is controlled by the matrix Object
{..... /	Paradigmatic alternatives ; valency options
ϕ features	Person, number, gender

This thesis was created with Word for Windows 6. The transliteration of Chinese is represented by the Pinyin system, using SILIPA fonts. The Chinese characters were produced using NJSTAR version 3 (professional) and the gb24 utility to create a pcx graphics file.

Acknowledgements

I want first to acknowledge the support I took from the memory of my Father, John Leslie: he taught me, wordlessly, to seek for myself after what is true (and to work all day Saturday and Sunday doing so); and also my Mother, Betty, who is very much a present and constant inspiration of courage and strength: yielding without and strong within.

Jock Hoe, my first Chinese teacher, really was 热情; his example convinced me that an adult can become a competent speaker, reader and writer of Chinese; his warm nature, and his deep concern for China made me want to meet that challenge.

I want to thank all my friends in NZ, Canberra and Beijing for supporting me in every way: Allanah for her companionship, intellectual stimulation and emotional and gustatory support over the two years I was studying Chinese in NZ; Fan Xiao Hong, Huang Wei-Min, and Paul Buckley, for Chinese conversations, especially in those difficult first months. Thanks to Paul, Bruce and Walt, for being Part Of The Chinese Experience, and my *didi* Geoffrey, for shouting me a trip home when I really needed one; Anne, Maxine, and Christine for walks along the Manawatu River; and all my friends and relatives at home, for keeping me warm in their hearts, and my possessions in their houses, during my absences in China and Australia.

My never-ending gratitude to Shen Zhi Jun, Cui Xi-Liang, Song Ying, Ma Jingwei, Daniel, Zang Qing, Xiao Fei, Long-Long, my room-mate Marita, and all the other Chinese friends who took this *lǎowài* into their hearts and lives in China, and made my time there easier, and unforgettable. Thanks also to Kate, Kazuko, and Kirsten for walks up and around Mt. Ainslie, and for telling me over and over that I'm 'a born academic' - I hope not, but it was sustaining. Kirsten's germanic perfectionism helped me hone my presentation in several areas.

Special thanks to Ian for thinking it was obvious that I'd want to put Chinese characters in my thesis, and use an IBM, and doing so much to make that possible, and to Chen Chun, Chen Yong, Yang Cong-Rong, Wang Da-Gang and other friends for assistance as informants. Chen Chun in particular was a tireless analyst and critic always willing to debate Mandarin syntax and meaning with me, and proof-read my work. Thanks to Chen Chun and Chen Yong also for sharing their correspondence with me, without which this work would not have been possible.

Special thanks too, to Verna for almost daily meetings at the 'electronic cafe'. You made me laugh, helped me keep my strengths in sight, and taught me new ones: the orange flashes in the Persian rug. And thanks to Tim, for Sundays at Tilley's, reminding me that other languages with morphology exist, and for acting like I know stuff.

I'd like to express my thanks also to staff in the Linguistics Department of ANU, who are all supportive. Especially, I'd like to thank Harold Koch for being so available to give advice, this made me feel looked out for from my first week and helped me make a difficult transition; Phil Rose, Andy Kirkpatrick, and my supervisors Tony Liddicoat and Avery Andrews, have all spent time listening to me slowly clarify my thinking, about linguistics in general and my thesis topic in particular. Avery often asked questions whose relevance I only discovered weeks later, and picked me up on several subtle, and not so subtle, errors of interpretation (of syntax, not Chinese); Tony got excited about the discourse aspects of my work, which helped me feel o.k. about my determination to straddle fences. Thanks also to my advisor Yang Tie-Jun, for checking my Chinese, and thanks to Phil Rose for proof-reading all my tones, a task that few could do and none could cherish (any remaining errors are my own!), and Angela for proof-reading,

especially co-indexes!! I would also like to acknowledge the support of the ANU in awarding me a scholarship which allowed me to do this work.

More than anyone, Kate Hayes, as an avid but ironically, non-linguistic listener, helped me to clarify my ideas. Discussing my work with her forced me to trim it to its transparent essence, and her enthusiasm made me see it again as fresh and interesting, carrying me through many an analytical block.

Finally, no work on Chinese is complete without a reference to *Kongzi* (Confucius). It was a couplet of his that spurred me to return to university:

学而不思则罔
思而不学则殆

xué ér bù sī zé wǎng sī ér bù xué zé dài

‘Study, but don’t think, and you’ll be misled; think but don’t study, and you are lost.’

ABSTRACT

In Mandarin, ellipsis can occur wherever the discourse context makes interpretation possible. This thesis addresses the central question: what are the specific features of such a discourse context; what syntactic and pragmatic factors facilitate or control the interpretation of ellipsis in Mandarin.

Beginning from a description of the frequency and distribution of ellipsis in a corpus of informal native discourse, a classification of ellipsis is developed that accounts for its interpretation in all cases. Ellipsis is shown to be primarily binary in nature: syntactically controlled ellipsis results in a bound interpretation, where a specific syntactic relationship holds between an anaphor and its antecedent; pragmatically controlled ellipsis results in a contextually sensitive interpretation, more readily described in terms of pragmatic values of topicality and focus.

The major contributor to bound ellipsis is the Subject position of dependent clauses. A classification of Mandarin verbs is proposed, based on their ability to take a dependent clause with an overt Subject and/or one where no overt Subject is acceptable. These valency options are shown to be related to lexical semantics. In other words, the syntactic constraints imposed by a specific valency option appear to reflect grammaticalisation of underlying semantic variation.

Zero elements under pragmatic control are essentially determined by a correspondence of topicality and focal value between the anaphor and an antecedent: \emptyset elements in focal positions (typically Objects) tend to refer to the last focal entity along a hierarchy: *bǎ* object > sentence Topic > nearest previous Object > nearest previous Subject.

Zero elements of established topicality, typically Subjects, tend to refer to the nearest preceding non-focal topic. These essentially pragmatic patterns combine with norms for the expression of focus and topicality to produce a tendency towards Subject-Subject co-reference and Object-Topic co-reference.

Chapter One

Introduction

老婆被人骗到四川去。他去找了一趟，没找着。
挤火车回家时，因火车严重超载，在十堰被警察赶下车了。
于是就起了报复社会之心。

*Lǎopo bèi rén piàn dào Sìchuān. Tā qù zhǎo le yītàng, méi zhǎo-zháo.
Jǐ huǒchē húi jiā shí, yīn huǒchē yánzhòng chāozài, zài Shíyàn bèi jǐngchá
gǎn xià chē le. Yúshì jiù qǐ le bàofù shèhuì zhī xīn.*

Ø's Old Lady, was bundled off to Sichuan by someone. He went off Ø to look for Ø, but Ø didn't find Ø. Ø Jammed on the train Ø coming home, because the train was seriously over-crowded, Ø was forced by the guard Ø to alight at Shiyan. Because of this, Ø then came up with the idea of Ø taking revenge on society.

1. ELLIPSIS IN MANDARIN

Mandarin is one of the official languages of the United Nations. It is spoken by an estimated 760 million people in mainland China and Taiwan, about 70% of all Chinese speakers there (Smith et al, 1993). In 1991, Australia had an estimated 235,000 speakers of Chinese languages (Smith et al 1993:125) and a large proportion of these are Mandarin speakers, though precise figures are unavailable.

One of the most immediately obvious characteristics of Mandarin discourse, and one of its most enchanting, is the brevity of expression made possible by widespread ellipsis. The extract at the start of this Chapter is from a letter written by a young man in Mainland China, reporting on some hot local gossip (the stabbing of nine people) to a Chinese friend in Australia. The extract refers to four people, each one mentioned explicitly only once. There are also three explicit mentions of inanimate entities; apart from these, in four sentences with nine verbs, providing a total of 18 possible NP positions, there are no other overt references to people or things: 11 places are unfilled;

11 semantic roles potentially unspecified. Two of the sentences have no lexical Subject, and one has no lexical Object; five clauses involve reference to participants which are represented only in other sentences, beyond the reach of syntactic ties.

So how is it that even non-speakers of Chinese with no previous exposure to Mandarin discourse patterns can easily understand the patterns of co-reference? This, essentially, is the question which this thesis sets out to answer. It aims to provide an overview of both syntactic and discourse factors influencing the interpretation of ellipsis in Mandarin discourse, and to demonstrate that ellipsis is used in Mandarin in such a way as to balance efficiency of expression with perfect clarity, adding to, rather than detracting from expressive versatility. In exploring this topic, I have always kept the task of the listener and the learner at the forefront, because I want to explore how they might analyse the system in order to understand the communicative intention.

1.1 Interpretation

Interpretation depends on a relationship between the ellipsed element and another known element in the context. This relationship could be semantic, pragmatic, or syntactic and each of these has its defenders in the literature on Mandarin (see La Polla 1993, 1994; Huang Yan, 1994, for arguments for pragmatic control; Huang C.T. James (1984, 1989, 1993), Battistella (1985), Hou and Kitagawa (1987), Paul Chen (1990), Tan (1991) for arguments for syntactic control).

The main pragmatic or discourse accounts of ellipsis in Mandarin treat ellipsis as a unitary phenomenon, and so have failed to take account of the **systematic** nature of variation in the reference of ellipsis. This has resulted in overly strong claims about the irrelevance of syntactic relationships to the issue.

Li and Thompson's (1979) and Chen Ping's (1984) analyses of Mandarin discourse, as well as La Polla's (1994) discussion of 'coreferent deletion' in Mandarin have all concluded that co-reference in Mandarin discourse is 'free', by which they mean that neither its occurrence nor its interpretation can be described or predicted by syntactic rules. Instead they say it is subject to the influence of various pragmatic considerations.

Counter-evidence to this position is readily available in the form of semantically or pragmatically absurd sentences, such as:

- 1) 我洗了衣服晾出去。
wǒ xǐ-le yīfu liàng-chū-qù
1sg wash-ASP clothes air-out-go
'After I'd washed the clothes [I] went out to air.'
* 'After I'd washed the clothes [they] went out to air'
*'After I'd washed the clothes [I] put [them] out to air.'

If pragmatics were the only determining factor, it would not be possible for syntactic structure to **dictate** an absurd interpretation as in 1).

In fact, observable variations of reference can be readily shown to be a principled variation, dependent on the syntactic structure in which ellipsis occurs, and on the Grammatical Function (GF) associated with the unexpressed element. So, syntax clearly does have a role to play in the interpretation of ellipsis. Nonetheless, syntactic theory has not been completely successful in accounting for variations in patterns of co-reference in terms of a systematic variance in syntactic structure. Nor can it account for ellipsis which has reference beyond the sentence.

The two main syntactic theories relevant to ellipsis are Chomsky's (1980, 1981) Government and Binding theories (GB) (see for example Brody, 1985, James Huang, 1982, 1984, 1987, 1989, 1993; Bouchard, 1989) and Lexical Functional Grammar (LFG) developed by Bresnan and others (see Bresnan, 1982). Both theories postulate different sub-categories of ellipsis with different rules for the interpretation of each. Since zero elements, by definition provide no morphological basis for sub-classification, they are most naturally classified on the basis of structural position, but it should be remembered that their distribution is not directly observable. This means that recognising and classifying instances and sub-types of ellipsis is problematic, and a poorly defined sub-class of ellipsis is left poorly accounted for.

Huang Yan (1994) in his neo-Gricean account, takes a somewhat intermediate position, stating that a pragmatic theory

of course, does not deny the existence of distinct syntactic, semantic and pragmatic levels and modes of explanation in the study of anaphora. On the contrary, it presumes the independence, or at least partial independence, of an irreducible grammaticalised stratum for pragmatically motivated constraints on anaphora. ' (Huang Yan, 1994:148).

Though I disagree with some details of Huang Yan's analysis, I believe that his insight is essentially correct: there exist both grammaticalised and pragmatically motivated constraints on anaphora. However, I argue that only one type of constraint operates on any given element of a given syntactic structure. Analysis of the frequency of syntactic structures in a corpus of informal written texts shows that interpretation is constrained by syntactic relationships in the majority of cases.

This thesis presents the results of that analysis, and describes and attempts to explain the distribution and reference of ellipsis in Mandarin discourse, on the basis of both syntactic and pragmatic factors. Syntactic constraints on reference are seen as operating in a rigid and principled way, but only across certain structural relationships; where these relationships do not hold, semantic and pragmatic factors are deemed to be in force. Thus pragmatics does not over-ride syntax, but is simply effective outside the domain controlled by syntax.

2. ORGANISATION OF THE THESIS

2.1 Three Parameters for Study: Reference, Distribution and Interpretation

Three main issues emerge in a discussion of ellipsis: The existence or otherwise of sub-types of ellipsis; the distribution of ellipsis (and any sub-types) through a text; and the process or processes of interpretation. These three aspects are used as a framework around which to organise the discussion of empirical findings and theoretical approaches in this thesis. Sub-types are established on the basis of referential effect i.e. whether ellipsed arguments are constrained to a single interpretation, or can have different interpretations according to context. Distribution is considered from both syntagmatic and paradigmatic perspectives; and interpretation is explained via a range of syntactic and /or pragmatic links between an anaphoric device and its antecedent.

2.2 Overview of Each Chapter

This introductory chapter explains how data and discussion are organised in the thesis and introduces the methodology and the main theoretical frameworks that are drawn on in what follows.

Chapter Two presents empirical data on the distribution and referential effect of ellipsis in a corpus of native speaker texts.

The genre which is under study is that of informal written texts. Letters were collected from native speakers. They were analysed and a number of straightforward frequency measures performed. Informal written discourse was chosen as the object of study for many reasons: firstly, native-speaker data of this kind is fairly readily available in the form of personal letters; secondly, this work is preliminary to a study of the acquisition of ellipsis by non-native learners of Mandarin (see Charters, to appear). It was felt that informal written texts would provide the most comparable material for such an investigation. The corpus is cited throughout the thesis; sentences from these texts are annotated with a reference number NS1-8 followed by the clause numbers in the text.

Chapter Two continues with a classification of ellipsis in Mandarin based on referential effect, with two categories: Variable and Bound ellipsis. This basic two-way structural division gives rise to an apparent three-way split in terms of paradigmatic alternatives; in addition to obligatorily ellipsed elements with no alternative means of expression, and optionally ellipsed elements, which could be replaced by an overt pronoun with the same interpretation, there is, in some cases, a referential contrast between two alternative forms: an overt NP and \emptyset . I call this contrastive ellipsis. The relationship between these categories is shown in Table 1.

Table 1 : Interaction of parameters

	Bound	Variable
Optional		
Contrastive		
Obligatory		

The next part of the thesis looks first at current theories of ellipsis, both pragmatic, in Chapter Three, and syntactic, in Chapter Four. Chapter Three reviews the work of Li

and Thompson (1979), Chen Ping (1984), La Polla (1993, 1994), and the neo-Gricean account of Huang Yan (1994). The former are all found to overstate the case for free co-reference and pragmatic control in Mandarin. The neo-Gricean account is judged to be cumbersome and insufficiently constrained: it proposes only tendencies, which make it difficult to account for the rigidity of referential mapping that can give rise to forced absurd interpretations. It is also weakened by the existence of a great many exceptions to its basic premises, and the principles invoked to explain these exceptions are found to be insufficient and not generalisable in many cases.

Chapter Four positions the present study in relation to Shopen's (1972) classification of nominal ellipsis into definite and indefinite ellipsis, then discusses the classifications of ellipsis and mechanisms of interpretation in Government and Binding theory (Chomsky, 1981; Haegeman, 1991; James Huang, 1984, 1989), and Lexical Functional Grammar (Horrocks, 1987; Bresnan, 1982).

The GB analysis is criticised for inconsistencies in the establishment of the empty categories themselves: identification of sites for PRO and pro both rely ultimately on verbal morphology of a sort entirely lacking in Mandarin; there is inconsistency of referential effect, and structural distribution within the category PRO, making identification on the basis of distribution also problematic.

Ellipsis in Mandarin often refers outside the sentence and therefore fails to satisfy the licensing constraints requiring local identification for the GB empty category (EC) little pro. James Huang's (1984, 1989, 1993) attempts to re-work the GB analysis to fit the Mandarin facts are also discussed. One major problem with these is that they postulate for every sentence, an empty topic position whose own reference still remains to be accounted for.

LFG offers a promising alternative, with the possibility of direct lexically controlled mapping of Grammatical Functions to the argument of another verb (functional control). Language users know an argument has been ellipsed simply because they know the semantic specifications of each lexical item in their vocabulary. This

means ellipsis need not be structurally represented, at least at c-structure level, and so pre-empts the problems of position and distribution.

However, the reference of LFG's null pronoun, PRED 'pro' is constrained to arguments which f-command it and must therefore be contained within the boundaries of a sentence. This then still fails to account for ellipsis which has a specific (as opposed to arbitrary or indefinite) reference **outside** the sentence.

The empty categories proposed in either GB or LFG are therefore insufficient to accommodate all instances of ellipsis in Mandarin. Neither theory claims to be able to account for the interpretation of ellipsis which refers beyond the sentence.

Chapters Five and Six discuss the syntactic processes available within GB and LFG that can account efficiently for the interpretation of obligatory and contrastive ellipsis in different syntactic contexts. Chapter Five discusses dependent verb complements; Chapter Six looks at VPs in NP positions; relative clauses; and Topic comment structures with no post-verbal Object. Ellipsis in these structures is not sensitive to changes in context beyond the sentence.

It is argued that interpretations which **are** sensitive to such a change in context are clearly not syntactically constrained, since syntax is itself constrained to intra-sentence relationships. Instead these instances of ellipsis are interpreted by pragmatic means.

On the other hand, where a syntactic relationship which is easily recognisable and easily characterised, holds consistently between \emptyset anaphor and antecedent within a single sentence, syntactic processes of control are more plausible. This is the case for structures discussed in these Chapters. Of course, that is not to say that syntactic processes may not have their roots in lexical semantics.

Contrastive ellipsis arises only in conjoined structures and is therefore attributed to the existence of two different underlying syntactic structures, one conjoined at sentence level, with two overt Subjects, and one conjoined at VP level, with only one. Where an argument position is removed by way of adjunction, conjunction, or complementation argument sharing results. In complement structures, this is a consequence of functional

control, which is itself semantically motivated; in other conjoined structures it is a consequence of the normal syntactic processes which map Grammatical Functions to arguments in a specific structural relationship to the verb. The relationships are identical for both partners in the conjunction. A similar analysis is possible for some, but not all instances of optional ellipsis. A contrast in reference is explained as a pragmatic choice of structure by the speaker.

Syntactic mechanisms of control are found to account for roughly two thirds of all cases of ellipsis in the corpus; pragmatic mechanisms are found to be similar for ellipsed Subjects and ellipsed Objects, but the syntactic control of Objects operates in quite different structures from those involving syntactic control of Subject ellipsis.

Subject ellipsis is significantly more frequent than Object ellipsis (corrected $\chi^2 = 53$, $df = 1$ $p > .0001$, see Tables 10 and 11, Chapter Two).

Pragmatics plays an important role in the selection of a specific structure from among alternatives, as well as interpretation of one third of all instances of ellipsis in the corpus.

Chapter Seven discusses the functioning of topicality, foregrounding and backgrounding in pragmatically determined interpretation of ellipsis in apposed and overtly conjoined clauses; clausal adjuncts; correlative constructions; and independent clauses and SCOMPs.

Where ellipsis arises because a structural position is unfilled, the reference of the ellipsed argument is determined pragmatically. This is achieved by reference to pragmatic norms about the representation of new, established, topical, and focal information. Empty positions associated with established referents (typically Subjects, see Du Bois, 1987) will be interpreted as referring to the last-‘mentioned’ topical but **backgrounded** participant (a Subject); empty positions associated with focal or new referents (typically Objects) will be interpreted as referring to the last-‘mentioned’ topical and **highlighted** participant (a topicalised Object, a sentence Topic, the nearest previous Object, or an overt Subject, in that order). Semantic plausibility plays a marginal role only.

Mandarin speakers have choices about the lexicalisation of NPs that are not available in languages that do not have regular definite ellipsis. Chapter Eight discusses this more creative aspect of ellipsis in Mandarin discourse: its cohesive, pragmatic and semantic functions and effects. Writers manage the discourse context in such a way that the intended interpretation is generally unmistakable.

Chapter Two

Discourse Data

1. *INTRODUCTION: INTERPRETING ELLIPSIS IN CONTEXT*

This chapter first discusses some preliminary considerations of the nature of ellipsis. It then presents statistics on the syntactic frames, referential patterns and frequency of different types of ellipsis within a corpus of Mandarin texts. The object under study is informal written discourse of native speakers.

On the basis of this corpus, two classifications of elliptic reference are proposed. It is classified firstly as 'variable' or 'bound' on the basis of whether reference can be affected by context beyond the sentence, and secondly, as obligatory, contrastive or optional on the basis of the acceptability of an overt expression of the same reference in a given syntactic structure.

1.1 Some Preliminary Considerations

1.1.1 Recognition of Ellipsis

The question of how ellipsis is interpreted pre-supposes the ability to recognise that a reference to some entity is intended, where no word appears to represent that entity. Some theoretical approaches extend this basic assumption to a stand on the structural 'presence' of a 'null category', or element whose role in the language must be acquired (by native and non-native learners) like any other.

Haegeman, speaking in the context of Government and Binding Theory (GB) puts it very clearly (1991:402-3):

If we assume that null elements are an actual component of the grammar of natural languages, we must assume that the language learner has the ability to posit such null elements in the representations he assigns to sentences. He needs to have arguments for positing these categories and ways of identifying them...The learner must know and the grammar must specify (i) in what conditions these elements can occur and (ii) how they can be interpreted, i.e. how such null elements can be given semantic content.

A common supposition accompanying a characterisation of ellipsis as 'deletion' or the occurrence of a null element in a given structural position, is that two sentences, one with and one without a lexical pronoun, are equivalent. In fact, they are neither pragmatically nor referentially equivalent. In the following examples, a \emptyset Subject must refer within its own sentence, while the overt pronoun in b) can refer beyond its own sentence:

- 2a) The dog_i was searching for drugs. The man_i opened the door and \emptyset _i smelled them.
b) The dog_i was searching for drugs. The man_i opened the door and he_j smelled them.

As shown by the indexing, in 2a) the \emptyset Subject can only co-refer with the NP 'the man'. By contrast, in 2b), the overt Subject 'he' could refer to any of a number of preceding Subjects in the discourse, even in previous sentences. It may be therefore preferable to suppose that 2a) and b) have different syntactic structures, and that either type of sentence may be independently generated. The idea of ellipsis as an unfilled structural position is by no means pre-ordained.

An alternative viewpoint is that a learner, or language-user knows what semantic roles are assigned by verbs as a direct consequence of lexical acquisition. If certain roles are known to be assigned, but there is no overt representation of the bearer of that role, they seek to identify that role-bearer from amongst the entities that are overtly realised elsewhere in the discourse or the extra-linguistic context. There is not necessarily a fixed position in which a null element 'occurs' in the surface structure. This is the view of ellipsis in both Lexical Functional Grammar, and in pragmatic and discourse functional accounts. It is also the viewpoint preferred in this thesis.

It is important also to distinguish between covert definite reference, i.e. the unique specification of a role-bearer in the absence of an overt representation, and indefinite or arbitrary reference, i.e. the non-specification of a role-bearer. In the case of the latter, a listener may assume or guess the identity of some unspecified role-bearer, but this is only an implicature, it can be contradicted without accountability, and sometimes with humorous effect as in the following:

- 3) A: "Did that guy give you back the money he owed you?"
B: "Let's just say, money exchanged hands."

A: "So we can go out for dinner?"

B: "No. He made me feel such a jerk, I gave him every cent I had."

Shopen's (1972) distinction between definite and indefinite ellipsis (discussed on pg. 62 below) provides a theoretical framework within which such a distinction can be made.

The non-realisation of semantic arguments of a verb is recognisable at the level of the clause, so a consideration of clause types and inter-clausal relationships is a major consideration in the treatment of ellipsis; the question of alternative valency options is also important.

1.1.2 Valency: the Optional Expression of Semantic Roles.

Conventionally established variation in the number of arguments taken by a single lexical form is commonly referred to as valency (Mosel, 1991). Many verbs which do not have conventionally reduced valencies may nonetheless be presented with reduced valency for certain pragmatic effects:

4) 'After teasing the normally placid Golden Labrador for hours, it finally bit.'

while some verbs are more restricted; consider:

5) 'After losing all her life, Gertrude finally won / ?found').

The capacity of individual verbs to appear in reduced valency frames clearly varies within a single language, and near semantic equivalents within or across languages may have different valency options.

In English, intransitive valency options generally involve omission of an unspecified NP in patient role, other languages, including Mandarin, have intransitive verbs where it is the Agent role that is omitted. Shopen (1973) suggests that, while languages may vary as to the acceptability or not of definite ellipsis, in those languages where it does occur, it will be 'regular' not lexically governed (the so-called 'pro-drop' languages), that is all verbs will allow ellipsis of contextually retrievable arguments. On the other hand, he predicts that 'indefinite ellipsis' will be lexically governed in all languages. Though there is some evidence (Enfield 1994) that this is not the case for Lao, it does appear to hold for Mandarin, as will be illustrated in the chapters that follow.

Unlike in English, unspecified Objects in Mandarin are generally represented by a generic or incorporated NP; an ellipsed Object is more likely to have definite rather than indefinite reference. Nonetheless, for any language, there is, by and large, native speaker agreement as to the normality and acceptability of reduced valencies in given contexts. This means that any definition of ellipsis must make reference to context.

1.2 A Definition

The definition employed in this thesis is this: ellipsis pertains in any syntactic structure in which there is not a 1-1 correspondence between lexical forms and the core arguments of a verb. Core arguments are the sub-categorisers of a verb, i.e. those whose ellipsis results in ungrammaticality, unless they are contextually retrievable.

1.3 Sub-types of Ellipsis

Empirical observations (see Chapter Four) show that ellipsis occurs in a variety of syntactic structures, is sometimes obligatory and sometimes optional, and produces a variety of referential effects, that is, in different strings and texts there may be different structural relationships between a covert reference to an entity and some overt reference to the same entity. In some cases a covert reference may allow of only one interpretation, in others it may be ambiguous, and in others it may vary with changes to context beyond the sentence in which it 'occurs'.

In this thesis ellipsis is held to represent a number of independent phenomena and may or may not involve empty structural positions. Where it is important to differentiate, PRO will be used to refer to instances where the semantic argument of a verb is understood to be uniquely constrained, but no lexical representation is permissible. Where the distinction is not relevant to the issue under discussion, 'Ø' will be used, or nothing at all.

2. METHODOLOGY

2.1 The Corpus:

An investigation was made of sentence structures used in letters written by 8 native speakers of Mandarin. The native speakers were aged between 20 and 50. There were four women and four men, all residents of the People's Republic of China, seven of whom have never been outside Mainland China. They live or grew up in areas where the local dialect is Mandarin (Beijing, Henan, Hubei, Dongbei). Four of the native-speakers were university educated, the others have graduated from high school. One speaks English well enough to study at an English-speaking university, the others have, practically speaking, little or no English.

The texts were extracts from letters to friends and relatives in Australia¹. Each extract consists of at least ten sentences of continuous text. In all cases the extract began at the point where the writer begins to introduce a new anecdote or piece of news, for most, immediately after the salutation.

2.2 Linguistic Analysis

Each text was analysed in terms of the structure of each sentence, the number of possible arguments and the number and form (pronoun vs noun) of realised arguments. For ellipsed arguments, both their Grammatical Function within their own clause, and their structural relationship or position relative to their antecedent was noted.

In addition, each ellipsed item was assessed in terms of its ability, in that structural position, to be replaced by an overt NP, and the effects of such a substitution on the interpretation of reference. The effects of placing the same syntactic structure into different semantic contexts were also investigated.

For each text the number of lexicalised and ellipsed Subject, Object and (for Copula constructions) Complement NPs was counted. Lexicalised NPs were divided into

¹ My thanks to Chen Yong and Chen Chun for generously sharing their personal correspondence with me. Thanks also to their friends and mine, the original authors of the texts. Where the texts are cited, names have been altered. Unless otherwise indicated, additional examples are the invention of the author and their grammaticality has been verified by native-speakers.

pronominal, nominal and DE Phrases². The number of possible arguments was assessed by considering the maximum number of semantic roles assigned to core (i.e. definite) arguments of a specific verb **in the sentence structure as it occurred in the text**. This means that the same verb may be assessed in different sentences as having a) different valency options; b) different expressions of the same underlying valency, for instance when lexicalisation is constrained in a given construction.

2.3 Statistical Analysis

A χ^2 test was used to determine whether there were a) significant differences in the use of overt forms vs ellipsis for Subject and Object arguments, and b) for ellipsed arguments, whether there were significant differences in the proportion of Subjects and Objects falling into various sub-types of ellipsis. These sub-types are introduced in the next section, and the results of the statistical analysis follows. 95% probability was taken as the level of significance.

Data for Complements and DE constructions were collected too, but the numbers were too small to be included in the χ^2 tests.

3. DESCRIPTIVE STATISTICS

3.1 Frequencies of Ellipsis in Different Clause Types

There are a total of 251 clauses in the native text sample (see Table 2), 124 clauses with a single-place predicate, 111 transitives (including also structures with two arguments and a complement), 3 di-transitives and 13 Copula plus Complement constructions.

Various patterns of ellipsis are possible within these clause types. Of the single-place predicates, 64 clauses are fully realised and 60 have ellipsed Subjects. Of the transitive clauses 22 are fully realised, 63 have ellipsed Subjects and realised Objects, 8

²The DE Phrase is a nominal modifier construction. The head of this construction can be ellipsed, in which case the DE Phrase is often considered to have a pronominal function and DE is often glossed as a 'Nominaliser'. Because the DE phrase is itself an overt nominal element, even when its head is ellipsed, it was counted separately.

have ellipsed Objects and realised Subjects, and 18 have no realised NP at all. In the three di-transitive sentences, two have ellipsed Subjects and overt semantic 'themes' as *bǎ* Objects, and one has a realised agent and an (ellipsed) Obj₂ 'theme', but all of them have an unmarked expressed Object (two locatives and a recipient).

Table 2. Patterns of Ellipsis in Native Speaker Texts

Verb type	Total Clause type	all NPs overt	Subject overt, Obj/Comp ellipsed	Subject ellipsed, Obj/Comp overt	all NPs ellipsed
Intransitive	124	64			60
Cop-Comp	13	9	2	2	Ø
Transitive	111	22	8	63	18
Di-transitive	3	Ø	Ø	3	Ø
Totals	251	95	10	68	78

4. LINGUISTIC ANALYSIS: DISTRIBUTION AND REFERENCE

Initial reading of texts and previous studies (e.g. Li and Thompson, 1979, Chen, 1984, La Polla, 1994) suggested that ellipsis might be usefully classified in terms of distribution, and in terms of referential effect.

Distribution was considered from the viewpoint of syntagmatic position (GF) and paradigmatic alternatives. The latter is taken to be an indication of meaningful variation in syntactic terms. That is to say, if Ø can substitute for different sets of items in different frames, this is evidence that the frames belong to different structural types, and it is this structural variation which licenses one set of alternatives in one instance, and a different set in another. It was considered important to investigate both types of distribution, since GFs can be involved in a variety of higher-level structures e.g. relative clauses. Subjects (or Objects, or any other GF) will not necessarily exhibit the same characteristics in all positions throughout a discourse.

4.1 Syntagmatic Variation:

Ellipsed NPs may bear a variety of Grammatical Functions e.g.

6) i) Subject:

于是就起了报复社会之心。

Yúshì Ø *jiù* *qǐ-le* *bàofù* *shèhuì* *zhī* *xīn*.

therefore Ø then rise-ASP revenge society POSS mind

'And so Ø came up with the notion of taking revenge against society.'

NS8:15

7) ii) Object:

他去找了一趟。

Tā qù zhǎo-le Ø yītàng,

3sg go seek-ASP Ø one.trip

'He went on a trip to look for Ø.'

NS8:12

8) iii) Obj 2

以后回国也可以教我。

yǐhòu húi guó yě kěyǐ jiāo wǒ Ø

After return country also can teach me Ø

'After [you] return home [you] can teach me [Ø].'

NS5:23-4

There is a question as to whether oblique NPs can be ellipsed or not. Clearly there are circumstances in which the qualification or modification conveyed by an oblique may be inferred in context, but it is questionable whether this phenomenon is the same as ellipsis of core NPs. Firstly, unless there is a stranded preposition³, ellipsis of Oblique arguments would be PP ellipsis and not NP ellipsis, the focus of this thesis; secondly, since obliques are not subcategorised for, there can be no **direct** reference from the verbal semantics to definite NPs in an oblique function. At most, verbal semantics may imply the generic involvement of a certain class of entity (e.g. an instrument or location).

³ This may happen in Mandarin with indefinite reference, producing a kind of passive reading:

? wo gei Ø zhuangle che

1sg for/to crash.ASP car

Some-one's smashed the car for me

but no such examples occurred in the texts. Note that it is the overt NP which is interpreted as the Oblique (maleficiary), despite its presence in Subject position, and the ellipsed NP is interpreted as the Agent, even though, by analogy the position of the Ø element is as Object of the preposition, usually the maleficiary. In other words, the mapping of semantic role to syntactic position is reversed with respect to a fully overt sentence:

Wo gei ta zhuangle che.

Ag Mal

1sg to/for3sg crash car

'I crashed the car for him.'

Wo gei Ø zhuangle che

Mal Ag

1sg to/for Ø crash car

'The car was crashed for me.'

The explanation for this probably lies in the preposition 'gei's history as a verb, which like all Mandarin verbs, can have a passive reading, so the two versions are, historically: 'I was given (Ø= by agent) smashed car', vs '*I gave him_{mal} smashed car' cf '*I gave (Ø=mal) smashed car'. So the ellipsed argument is still interpreted as a core, not an oblique **semantic** argument.

4.2 Paradigmatic Variation:

There is also variation with regard, firstly, to the acceptability of inserting an overt NP into a structure, and secondly, to the effect which such insertion has on co-reference.

Paradigmatic sub-classes include:

- a) obligatory
- b) optional
- c) contrastive

4.2.1 Obligatory Ellipsis:

Obligatory ellipsis refers to structures where overt representation of an argument in a specific position is ungrammatical. For example, the Subjects of dependent verbs in serial verb constructions are obligatorily ellipsed. In the first example the matrix is an imperative, the second is a simple declarative (\emptyset 's in the following examples are limited to those relevant to the point under discussion):

9) **VCOMP Subject**, transitive matrix verb

你写信 (*你) 告我。

*nǐ xiě xìn { \emptyset /*nǐ} gào wǒ*

2sg write letter { \emptyset /*2sg} tell me

'Write and \emptyset /*you tell me.'

after NS5:15

10) **VCOMP Subject**, intransitive (modal) matrix verb with adjunct PP

我真想 (*我) 跟你聊聊。

*wǒ zhēn xiǎng { \emptyset /*wǒ} gēn nǐ liáo-liáo*

1sg really want { \emptyset /*1sg} with you chat

'I really want /*I to chat with you.'

NS1:1

Obligatory ellipsis also occurs in Mandarin in relative clauses, and topic-comment structures, example 11) contains both in a single structure.

11) **Relative Clause and Object in 'Topic-comment'**

不能在洗碗机里洗的就用手洗。

[_{top} *bù néng zài xǐwǎnjīlǐ xǐ \emptyset de] \emptyset , jiù yòng shǒu xǐ \emptyset*

NEG able in dishwasher wash \emptyset DE \emptyset , then use hands wash \emptyset

'Those \emptyset which \emptyset can't be washed in the dish-washer, [I] wash \emptyset by hand.'

NS3: 24c

The relevant Ø's in 11) are in bold, the other is an optionally ellipsed head. In the Chinese text, the first Ø represents the Object position of the relative clause, which is the relativised constituent and is bound to refer to the head. This is precisely parallel to the situation in the English except English has a 'relative pronoun'. The second Ø in the Mandarin text is the head. The third is the Object position of the main verb, *xǐ* 'wash', English allows a re-iterative (stressed) pronoun here, but Mandarin does not.

4.2.2 Optional Ellipsis:

In Optional ellipsis the zero element can alternate paradigmatically with an overt lexeme **of the same reference**. These are illustrated by examples 12) and 13). These sentences are the same as at 9) and 10), but we are considering the Subject position of the matrix verb, not the dependent verbs.

12) IMPERATIVE

(你 / Ø) 写信告我。
 {nǐ / Ø} xiě xìn gào wǒ
 {you / Ø} write letter tell me
 '(You) write and tell me.'
 NS5:15

13) 1st PERSON

(我) 真想跟你聊聊。
 {Ø / wǒ} zhēn xiǎng gēn nǐ liáo-liáo
 {Ø / 1sg} really want with you chat
 '(I) Really want to chat with you.'
 NS1:1

(cf. 'Ø Nice to see you'; 'Ø Pleased to meet you')

4.2.3 Contrastive Ellipsis

In Contrastive ellipsis, Ø alternates (paradigmatically) with an overt lexeme of a **different reference only**. Contrastive ellipsis is a characteristic of paired clauses, and is associated with semantic subordination: conditionals, cause, result, temporal interdependence etc. Overt pronouns in these constructions are essentially switch-Subject markers. In examples 14) - 16) the a) version is the original and the b) version has been constructed to illustrate the effect of lexicalisation:

14a) With overt marker of sub-ordination:

他挤火车回家时，

ta₁ jǐ huǒchē húi jiā shí,

3sg cram train return home time

'Ø cramming on the train to come home,

Ø 在十堰被警察赶下车了。

Ø₁ zài Shíyàn bèi jǐngchá gǎn xià chē-le.

Ø at Shiyán by guard drive descend train-ASP

he was put off by the guard at Shiyán.'

after NS8:14a

b) 他挤火车回家时，他在十堰被警察赶下车了

ta₁ jǐ huǒchē húi jiā shí, ta₂ zài shíyàn bèi jǐngchágǎn xià chē-le.

3sg cram train return home time 3sg at Shiyán by guard drive descend train-ASP

'When he₁ was cramming on the train to come home, he₂ was put off by the guard at S.'

15a) No overt marker:

我住在这里要帮助做家务。

wǒ zhù zài zhèlǐ Ø yào bāngzhù zuò jiāwù

1sg live at here-LOC Ø must help do housework.

'Ø To live here, I must help with the house-work.'

NS3:16

b) 我住在这里他要帮助做家务。

wǒ zhù zài zhèlǐ ta yào bāngzhù zuò jiāwù

1sg live at here-LOC 3sg must help do housework.

'[If] I [am to] live here, he must help do the house-work.'

Note that in English the b) sentences are ambiguous unless the pronouns themselves differ in person, number or gender.

There are some examples, which I have called nominal control, where it is the possessor of a noun (generally one which refers to a psychological state) which controls an ellipsed Subject. The same phenomenon occurs in English, with similar semantic effects.

16) Nominal control:

Ø 还无生孩子的打算。

Ø hái wú shēng háizi de dǎsuàn

Ø still without bear child DE intention

'[She] still has no intention of Ø having a child.'

NS1: 12

- 17) 还无她生孩子的打算。
 ? \emptyset hái wú tā shēng hái'zi de dǎsuàn
 \emptyset still without 3sg bear child DE intention
 ? '[He] has no intention of her having a child.'

Certain nouns clearly control Subject positions of their arguments just like matrix verbs do.

4.3 Not just an Empty Pronoun

The shift in the interpretation of reference when a zero element is replaced by an overt one is evidence that these ellipsed arguments are not simply phonologically empty versions of an overt pronoun, in either semantic or structural terms (see Shopen, 1972 for discussion of this point). Either there is a structural difference between the two sentences, i.e. conjoined VPs vs conjoined sentences, or they have the same underlying structure but the two forms, \emptyset vs an overt pronoun, have different referential components of meaning, or are attributed different levels of topicality. In other words, if we want to argue for a morphologically empty form present in structure, that \emptyset must actually be something other than just a morphologically unexpressed equivalent of an overt pronoun or it could not produce a different interpretation: it must have a different semantic structure. The important distinction would be one of relative salience in the attention of the interlocutors, and therefore retrievability. However, other instances of \emptyset produce the same interpretation as a pronoun and therefore must have the same meaning or pragmatic status as overt pronouns. So proposing structural presence here would require a polysemous \emptyset element - not a very satisfactory solution as differentiation would be problematic.

To recap then, in terms of distribution, syntagmatically, we have Subject, Object and Obj 2 ellipsis, and paradigmatically we have obligatory, optional and contrastive ellipsis. The existence of contrastive ellipsis shows that \emptyset elements cannot simply be 'morphologically unexpressed pronouns'.

4.4 Referential Variation

Practically speaking, in a given stretch of discourse, most instances of ellipsed arguments have only one likely interpretation. This norm of retrievability in context masks potential variability at the level of syntactic structure.

Examples 18)- 21) demonstrate the difference in referential effect which I have labelled bound reference and variable reference, though it should be remembered that 'variable' refers to an underlying potential, not to variability of interpretation in a single context.

4.4.1 'Variable' (i.e. Contextually Sensitive) Reference:

Substitution of one lexical item for another of the same class, can alter the semantics of a context without altering syntactic relationships. In some cases, where different lexical items impose different semantic constraints on their arguments, this results in an alteration of interpretations of co-reference. In example 18) the version with a) as its second sentence is the original from the corpus. It contains the verb *huài* with a perfective marker *le*. This combination corresponds to the English 'broke down' and demands an inanimate Subject. In b) the verb has been replaced with *fán* 'fed-up', and this demands a human Subject. The relevant argument is the final Subject (bold), it and its coreferent elements are under-lined:

18) 山东有个农民老人买了一台电视机，

Shāndōng yǒu ge nóngmín lǎorén₁ mǎile yītái diànshìjī₂.

Shandong had a country-folk old-person buy-ASP one-CL TV.

There was an old peasant from Shandong [who] bought a tv.

回家看了两天就坏了。

a) \emptyset_1 húi jiā \emptyset_1 kànle \emptyset_2 liǎng tiān \emptyset_2 jiù huàile

\emptyset Return home \emptyset watch-ASP \emptyset two days \emptyset then bad-ASP.

[He]went home and \emptyset watched \emptyset for two days, then [it] broke down.

after NS2:3-10

回家看了两天就烦了。

b) \emptyset_1 húi jiā \emptyset_1 kànle \emptyset_2 liǎng tiān \emptyset_1 jiù fánle

\emptyset Return home \emptyset watch-ASP \emptyset two days \emptyset then fed.up-ASP.

[He] went home and \emptyset watched \emptyset for two days, then \emptyset was fed up

The Subject of the final predicate is interpreted as referring to the old man, or to the TV dependent entirely on the semantics of that predicate. This means that there is no syntactic constraint imposing a **specific** antecedent-anaphor relationship. In a) there is co-reference between a \emptyset Subject and a sentence **external** overt Object NP; in b) there is co-reference between two \emptyset Subjects in the same sentence.

4.4.2 Bound Co-reference

The term 'bound' co-reference refers to cases where the unrealised argument is constrained to co-reference with an item **in a specific structural relationship** to the verb of which it is an argument. Variable co-reference (exemplified in 18) above) arises only where ellipsis is optional. The sub-types of bound co-reference are illustrated in examples 19) to 21).

4.4.2.1 Examples of Bound Reference for Each Paradigmatic Type:

19) IN COPULA CONSTRUCTION (OPTIONAL ELLIPSIS)

我的论文是关于普通话异读字的。

wǒde lùnwén shì guānyú pǔtōnghuà yìdúzì de \emptyset

1sgPOSS paper COP about MSC homographs DE \emptyset

My paper was one \emptyset about homographs in Modern Standard Chinese.

NS1: 9

20) ASPECTUAL SUBORDINATION (CONTRASTIVE ELLIPSIS)

好多同学都是

hǎo duō tóngxué; dōu shì

good many students all are

\emptyset 分别了三四年后见过面。

\emptyset_i fēnbié-le sān-sì nián hòu jiàn-guò-miàn

\emptyset_i separate-ASP three-four years after see-ASP-face

'A good many students were meeting after \emptyset having been separated for three or four years.'

NS6: 6-6a

21) OBJECT CONTROL (OBLIGATORY ELLIPSIS)

我读到一条消息说...

wǒ; dúdào yītiáo xiāoxi; \emptyset_j shuō "....."

1sg read.ASP one.item news \emptyset say

'I_i read a news item_i \emptyset_{vj} saying "...'. '

NS2:1

4.5 Interaction of Parameters

In general terms, all instances of obligatory ellipsis (VCOMPs, Relativised Constituent, topicalised Object) are also instances of bound co-reference, but not all bound co-reference results from obligatory ellipsis. Table 3 shows the observed interaction of referential effect and paradigmatic type.

Table 3 : Interaction of parameters

	Bound	Variable
Optional		
Contrastive		
Obligatory		

5. STRUCTURAL CONTEXTS: WHERE DOES ELLIPSIS HAPPEN?

Types of ellipsis are related to specific syntactic structures. Particularly important is the distinction between ellipsis of variable reference, and that of bound reference. Where reference is variable, ellipsis is always optional, and may be definite, or indefinite (see Table 4); where reference is bound, ellipsis may be optional, contrastive or obligatory, but it is always definite (see Table 5).

Table 4: Variable Reference = Optional ellipsis

		Subjects	Objects	Other	Total
Variable	Indefinite	10			10
	Definite	4			4

Table 5: Bound Ellipsis is Always Definite

	Subject	Object	Other	Total
Optional	25		1	26
Contrastive	28			28
Obligatory	60	8		68
	113	8	1	122

A more detailed breakdown, by syntactic structure is presented in Tables 6 - 8.

Table 6 presents what I have called 'quasi-ellipsis' because it appears to form a special sub-class of indefinite ellipsis, where a generic NP of indefinite reference (e.g. *ré*n 'people') cannot appear in the 'missing' argument position. These are all idiomatic expressions and are readily explained as functioning like single lexical items with a reduced valency compared to the valency of their component parts.

Table 7 is in two parts ; 7i includes instances of indefinite reference where an overt generic NP could be substituted, and it can therefore be argued that a semantic role is assigned; reference is considered variable because it is not tied to a single entity in a specific syntactic relationship. Table 7ii presents contexts where NPs are ellipsed with definite but variable reference. Table 8 shows where NPs are ellipsed with definite and bound reference. As mentioned above, ellipsis of bound reference can be further sub-divided into Optional, Contrastive and Obligatory ellipsis; Table 8 is sub-divided accordingly. Table 9 presents a summary of the other tables, giving the overall frequencies for non-referential, variable and bound ellipsis. (In the tables, Ø indicates contextually sensitive and P indicates bound.)

Table	Non-referential	Variable	Bound	Total
Table 7i	1	0	0	1
Table 7ii	0	1	0	1
Table 8 (Optional)	0	0	1	1
Table 8 (Contrastive)	0	0	1	1
Table 8 (Obligatory)	0	0	1	1
Table 9	1	1	3	5
Total	1	1	5	7

Table 6: QUASI-ELLIPSIS: Idiomatic non-assignment of Subject role

(No reference implied, No overt NP possible)

PARADIGMATIC TYPE:	STRUCTURAL CONTEXT	Notes on Structure	S	O	Ø
(No overt NP possible)	Subjectless idioms	<i>zhēnshì, méi bànfǎ</i> etc. '[It]'s really' '[there]'s nothing to be done'	12		12
12	NON-REFERENTIAL		12		12

Table 7i): VARIABLE REFERENCE - Indefinite

PARADIGMATIC TYPE:	STRUCTURAL CONTEXT	Notes on Structure	S	O	Other	Ø
OPTIONAL (Overt indefinite NP possible)	Subject ellipsis	arb Subject, main verb arb Subj in R.C.	6 1			6 1
	Object ellipsis	arb Object, main verb		3		3
10	VARIABLE Optional indefinite		7	3		10

Table 7ii): VARIABLE REFERENCE - Definite

OPTIONAL	Simple Sentences					
	Subject ellipsis	Imperative	1			1
		Subj of main/indep	29			29
		Subj= <i>yǒu</i> obj	1			1
	Object ellipsis			9		9
	OBJ2 ellipsis	(<i>jiāo</i> 'teach')			1	1
	OPTIONAL SIMPLE		31	9	1	41
	Complex Sentences					
	1) Core GF =VP					
	Subject Ellipsis	VP in SUBJ position	1			1
	2) Co-ordination					
	a) overt conjunction					
	Object ellipsis			1		1
	b) Apposition					
	Object ellipsis			1		1
	3) embedding					
	Subject ellipsis	SCOMP Subj	3			3
		SCOMP Subj = <i>yǒu</i> obj	1			1
	OPTIONAL COMPLEX		5	2		7
48	VARIABLE Definite		36	11	1	48
10	VARIABLE Indefinite (from Table 7i)		7	3		10
58	TOTAL VARIABLE		43	14	1	58

Table 8) : BOUND REFERENCE - Definite

PARADIGMATIC TYPE:	STRUCTURAL CONTEXT	Notes on Structure	S	O	Oth	P
OPTIONAL	4) HEADS of RCs					
	COPULA sentences	r.c. = SUBJ (<i>yě, dàn, hái, hòulái, jiù</i>)	1			1
	a) Overt conj. Subject ellipsis	(and, but, still, later, then)	17			17
	b) Apposition Subject ellipsis	(NP in Clause >1)	7			7
		r.c.= COMP			1	1
26	BOUND OPTIONAL		25		1	26
CONTRASTIVE	5) Subordination ii)					
	a) overt Subject ellipsis	time : <i>zhe, le, shíhòu</i> cause : <i>zhīxià,</i> result: <i>jiù</i>	12			12
	b) apposition	purpose, time, cause	12			12
	6) Core GF =VP Subject Ellipsis	VP in OBJ position	2			2
	7) Nominal control		2			2
28	BOUND CONTRASTIVE		28			28
OBLIGATORY	8) Serial verbs	Matrix verb =				
		COPULA	2			2
		intrans / passive	30			30
		Trans subj control	13			13
		Trans Obj control	9			9
		Shared control	2			2
		subtotal VCOMPs	56			56
		9) RCIs: NP rel				
		Subject (of RC)	2			2
		Object (of RC)		3		3
		10) Topic Comment				
	Object Ellipsis	Obj = topic Obj = <i>yǒu</i> obj		4 1	4 1	
	11) Conjoined Verb		2		2	
	subtotal OTHER		4	8	12	
68	BOUND OBLIGATORY		60	8		68
122	TOTAL BOUND		113	8	1	122

Table 9: Summary of Ellipsis Types

		Sub	Obj	Other	Total
Non-Referential		12			12
Variable	Indefinite	7	3		10
	Definite	36	11	1	48
Bound (always definite)	Optional	25		1	26
	Contrastive	28			28
	Obligatory	60	8		68
Total Ellipsed NPs		168	22	2	192

Examination of the distribution of sub- types of definite ellipsis with respect to syntactic structures (as illustrated in Tables 6 to 8) reveals that contrastive and obligatory ellipsis occur in fewer syntactic contexts than optional ellipsis does. The following list summarises the data in the Tables above and makes this point clearer.

5.1 Structures Involving OPTIONAL Ellipsis:

- 1) Subjects and Objects in main clauses or one-clause sentences, including the 2nd person addressee of Imperatives (44);
- 2) Subjects and Objects in the second of two clauses joined by any one of a variety of overt conjunctions or by apposition (25);
- 3) Subjects of VPs which are themselves in an NP position (3);
- 4) Heads of Relative clauses (2);
- 5) Obj2 ellipsis (1)

5.2 Structures Involving CONTRASTIVE Ellipsis:

- 1) Subjects of sub-ordinate clauses introduced by overt conjunctions: *jiù* 'then', *-hòu* 'after' or in apposition (24)
- 2) Possessor-Subjects of verbs of cognition (2)
(i.e. Nominal control e.g. *tā wú shēng hái zi de dǎsuàn* 'she has no intention of having a child'.)

5.3 Structures Involving OBLIGATORY Ellipsis:

- 1) Subjects of VCOMPs (56);
- 2) Relativised constituent in Relative clauses (5) (Subj or Object of Rel clause);
- 3) Object In Topic-Comment sentence (5)

It should be noted that the contexts of occurrence are, for the most part, complementary, the only exception being in the area of conjoined clauses, where some exhibit Contrastive ellipsis, and others Optional ellipsis. This point will be explored further in the next chapter.

6. MORE STATISTICAL MEASURES

6.1 Form and Grammatical Function

Table 10 shows the frequencies with which different referential devices are used for core arguments in the texts.

Table 10: Frequency of Form by Grammatical Function

Grammatical Function	Form:			Total Ellipsed	N	prn	Total Overt	Total
	∅							
	Non-ref	Variable	Bound					
Subject	12	43	113	168	52	49	101	269
Object	3	11	8	22	8	85	93	115
Complement	∅	∅	1	1	∅	10	10	11
OBJ2		1		1	2		2	3
	15	55	122	192			206	398

A χ^2 test was performed on the data for Subjects and Objects only (other NP types were too few for inclusion) to determine whether there is a significant difference in the frequency of ellipsis generally with respect to GF (see Table 11). Objects were found to be ellipsed significantly less frequently than Subjects (corrected $\chi^2 = 58.102$, $df = 1$, $p > .0001$, E for ∅ Subjects = 122 for ∅ Objects = 53).

Table 11: Frequency of Definite Ellipsis for Subjects vs Objects

Grammatical Function	Ellipsed	Overt	TOTALS
Subject	156 E= 122.88	101 E= 135.12	257
Object	19 E=53.12	93 E=59.88	112
TOTALS	175	194	369

6.2 Paradigmatic and Referential Sub-types and Grammatical Function

Table 9 above (pg. 29) presents the division of sub-types of ellipsis according to GF. Since expected values were very low in some categories (i.e. indefinite, optionally

bound and contrastively ellipsed Objects) categories were collapsed for the purposes of a χ^2 test, into referential indefinite vs variable definite (Optionally Variable) vs bound (Optionally bound, Contrastive and Obligatorily ellipsed) reference; idiomatic \emptyset Subjects were excluded.

Looking at the overall frequency of these major sub-types of ellipsis in the texts (Table 12), the clear majority (90%) is accounted for by ellipsis of definite reference, either variable or bound).

Table 12: Frequency of Major sub-types of Ellipsis in Corpus

Indefinite	Variable definite	Bound
10	48	122
E=60	E=60	E=60

This is statistically highly significant (one-way $\chi^2 = 66.7$, $p < .0001$), which is not surprising, given that the texts are narratives about specific people and events, but it is also a reflection of the low frequency of indefinite Object ellipsis, i.e. intransitive valency options with semantically transitive verbs.

Note that indefinite ellipsis of Objects (i.e. intransitive valencies for inherently transitive verbs) is very infrequent (see Table 10). There were only two cases where the matrix verb was potentially transitive, but there was no overt Object. Of these, one is a Subject control verb, so the presence or absence of an Object makes no difference to co-reference.

- 22) 当时样子你可想而知了。
dāngshí yàngzi nǐ kě xiǎng ér zhī ∅ le
 at.the.time appearance 2sg can think ER know \emptyset ASP
 '[My] appearance at the time, you can imagine (lit. think and know).'

NS5:12

The other is a joint Subject -Object control verb (*bāngzhù*).

- 23) 我要帮助做家务。
 [Wǒ] yào \emptyset_i bāngzhù \emptyset_i zuò jiāwù.
 [I] require \emptyset help \emptyset do housework.
 '[I] must help do the housework.'

NS3: 5

In the actual text, the distance between this verb and the nearest potential referent for an ellipsed Object in 23) is so large that it seems much more likely to represent an intransitive valency (arbitrary Object).

Given that Mandarin tends to use incorporated generic nouns, rather than ellipsis to convey indefinite reference, a larger corpus might show the difference in the rate of ellipsis for indefinite Objects vs indefinite Subjects to reach a statistically significant level. Given the low frequency, the corpus for this study is too small to allow reliable comparisons.

Indefinite reference is where reference is not controlled (either syntactically or pragmatically) because there is no specifically identified antecedent in the text. Since the main concern here is to distinguish between syntactic and pragmatic control, subsequent measures refer to ellipsis of definite reference only.

Table 13 shows the relative frequency of bound vs variable ellipsis for Subjects and Objects. A two-way χ^2 (with Yates Correction factor) shows there is a statistically significant difference (at the level of 95%, though not at 99%) in the frequency of bound vs variable reference associated with the ellipsis of each GF ($\chi^2=5.951$ $p=.0147$). In other words, Objects are slightly more likely to be ellipsed with variable reference, and Subjects are slightly more likely to be ellipsed with bound reference.

Table 13: Frequency of Subject and Object Ellipsis for Bound vs Variable Definite Reference

GF	Variable	Bound	Total
Subject	43 E= 48.14	113 E=107.86	156
Object	11 E=5.86	8 E=13.14	19

Table 14 shows that there is a significant difference between the overall frequency of bound vs variable ellipsis in the texts generally (In other words, statistically speaking, an ellipsed NP is much more likely to be of bound reference than of 'variable' reference.

Table 14: Frequency of Variable vs Bound Ellipsis

Variable	Bound
54.5	120.5
E=87.5	E=87.5

(Corrected Observations, $\chi^2 = 24.89$ $p < .001$)

7. CHAPTER SUMMARY: VARIATIONS IN ELLIPSIS

This Chapter has presented data to show that ellipsis in Mandarin can be divided into sub-classes on the basis of referential effect and paradigmatic type whose occurrence correlates strongly with specific structural contexts.

The structures where variable ellipsis occurs are the same for Subjects and Objects, but the structures where contrastive and obligatory ellipsis occur are different.

Variable ellipsis accounts for just under one third of definite ellipsis in the corpus and is therefore an important sub-class of ellipsis in Mandarin discourse. In the light of these observations, the next chapter will discuss claims that co-reference in Mandarin is free (Li and Thompson, 1979, Chen Ping, 1984, and La Polla, 1993). Though these claims are clearly overstated, neither of the major syntactic theories, GB and LFG, can provide an account of the processes by which ellipsis of variable reference is interpreted. This will be discussed in Chapter Four.

Chapter Three

Ellipsis in Discourse Theory

"There seems to exist a class of language (such as Chinese, Japanese and Korean) where pragmatics appears to play a central role which in familiar European languages (such as English, French and German) is alleged to be played by grammar. In these 'pragmatic' languages, many of the constraints on the alleged grammatical processes such as intrasentential anaphora are, in fact, primarily due to principles of language use rather than rules of grammatical structure."

(Huang Yan 1994: 259)

1. INTRODUCTION

In this Chapter and the next, I will discuss how different theories of ellipsis result in different classifications of ellipsis. This Chapter discusses discourse and pragmatic theories. Though primarily a theory of pragmatics, Neo-Gricean theory makes reference to syntactic concepts of Binding and Control, and the concept of c-command, and these are introduced in this Chapter. The application of the Disjoint Reference Preference and Gricean principles of informativeness and quantity to an interpretation of reference is also discussed.

The main points to be made in this chapter are:

Reference of ellipsis

- 1) Ellipsis is not a unitary phenomenon as implied in the work of Li and Thompson (1979), Chen Ping (1984) and La Polla (1993), and claimed by Huang Yan (1994).
- 2) The referential possibilities of any instance of an unexpressed argument are best evaluated in a discourse context, as certain phenomena are never evident in de-contextualised sentences.

Distribution

- 3) Attempts to explain the distribution of ellipsis vs overt pronouns, on the basis of differences in thematic continuity or the 'tightness' of inter-clausal semantics have not been entirely successful.

Interpretation

- 4) Discourse- pragmatic models are too 'permissive': they tend to suggest that the interpretation of \emptyset is limited only by the number of semantically compatible antecedents, and that an arbitrary interpretation arises only in the absence of any semantically compatible antecedent, and this is not the case. Huang Yan's (1994) hierarchy of antecedent selection is not supported by empirical evidence.
- 5) Variability in reference usually results from structural variation and structural ambiguity. Contrary to La Polla's (1993) position, Subject is shown to control coreferent deletion in at least some cases.

2. REFERENTIAL SUB-TYPES

In discourse theory and descriptions, ellipsis has generally been viewed as a unitary and uniform phenomenon.

2.1 Undifferentiated \emptyset

Studies have been undertaken on the use of ellipsis, or 'zero anaphora' in the discourse of a number of languages (see for example Espino de Valdivia (1989) for Spanish and Japanese; Grima (1986) for Thai; Shopen (1972, 1973) and Fillmore (1969, 1986) for English; Besnier (1985) for Tuvaluan; Junghare (1988) for Marathi and Hindi). Many other studies look at reference tracking or 'anaphora' in general, (for example Lust, 1986, Tomlin and Pu, 1991), but these deal in large part with **overt** anaphors, i.e. pronouns, rather than with zero anaphora.

There are three main studies which specifically address the question of ellipsis in Mandarin from a discourse or pragmatic perspective. Li and Thompson's (1979) analysis of zero anaphora, Chen Ping's (1984) study of third person zero anaphora, and Huang Yan's (1994) book 'The Syntax and Pragmatics of Anaphora. A study with special reference to Chinese'; La Polla's (1993, 1994) works on grammatical categories also make reference to coreferent deletion in Mandarin.

None of these differentiate between sub-types of ellipsis; variations of referential effect are interpreted as evidence of free co-reference, and the possibility of systematic

variability is disregarded (see for example Li and Thompson 1979:312, Chen 1984:2; La Polla, (1993: 773).

This position is made explicit in Huang Yan's work; in his critique of GB he concludes (1992:27) that zero anaphors in Chinese are realised by 'syntactically undifferentiated gaps rather than by specific empty syntactic categories'.

In his later book Huang Yan (1994) takes a neo-Gricean approach. Neo-Gricean accounts focus on interpretation, with reference to Grice's (1975) principle that speakers say 'as little as necessary' to enable interpretation in a given context. Formal contrasts are assumed to be the major signalling device at the speaker's disposal, and form is therefore taken as the basic criteria for classification of anaphoric devices. This means that zero anaphora, having no overt realisation, always constitutes a single formal class of anaphoric device.

The same is true of Givon's work on discourse universals. Givon (1983) classifies anaphoric devices on the basis of phonetic and morphological form. He presents (1983:18) a 'scale of phonological size' where zero anaphora constitutes a single point on the scale:

↑ zero anaphora
| unstressed / bound pronoun (agreement)
| stressed/ independent pronouns
↓ full NP's.

However, Givon's use of a scalar system does not preclude a finer sub-division within each stage.

3. DISTRIBUTION

3.1 Conjoinability, and Continuity

In terms of distribution, Li and Thompson (1979) drew attention to the possibility of an association between conjoinability and ellipsis, but were unable to define or characterise conjoinability in such a way as to predict the occurrence of ellipsis in texts. They concluded that ellipsis is actually the norm in Mandarin, and lexical pronouns are

simply introduced at relatively even distances through a discourse for reasons that are essentially stylistic.

Chen (1984) takes up this idea and seeks to explain the distribution of overt lexical pronouns proposing that they are used to signal disjunction.

Givon (1983, 1984) makes a similar claim in general terms, suggesting that the use of anaphoric devices follows universal discourse principles, whereby the least marked anaphoric devices refer to the most accessible referents; \emptyset elements are, naturally, the least marked devices of all, and should therefore be restricted to contexts where their intended antecedent is obvious.

Givon (1984:131, note 27) states that "the difference between zero anaphora and unstressed pronouns in English is expressed not in terms of referential distance (both 1.00 on the average) and not so much in terms of potential ambiguity (both low values), but rather in terms of *thematic continuity*." In other words, the distribution of zero anaphora is a function of thematic continuity.

For Chen, disjunction, or the **absence** of semantic continuity between clauses, is reflected in factors such as: turning from background to foreground information, or from new to given; digressions; insertion of a temporal, locative, and adversive or other adverbial; pauses or hesitation when a theme has continued for a long time; paragraph boundaries; switches or turns in conversation, 'and so on'.

In fact, pauses, hesitations and paragraph breaks are better understood as symptoms, rather than causes of disjunction. Moreover, the relationship between temporal and locative adverbials and discontinuity is by no means certain. Givon (1983:8) states

thematic continuity is the overall matrix for all other continuities in the discourse. It is the hardest to specify, yet it is clearly and demonstrably there. Statistically it coincides with topic and action continuity to quite an extent within the thematic paragraph. ...However, topics/participants may change within the discourse without *necessarily* changing either action continuity or theme continuity (Givon, 1983:8).

In fact, in Mandarin, sub-ordinate clauses in adjunctive roles expressing minor diversions, or location of events in time or space are often precisely the context in which

ellipsis occurs. For example in the following tract, the first clause (underlined) is an adverbial of time, but the Subject is ellipsed, not lexicalised as Chen (1984) predicts:

24) 晚上回来把衣服收回来叠起来。

Ø wǎnshàng húilái Ø bǎ yīfu shòuhúilái Ø Ø dié-qǐ-lái
 evening return BA clothes take-back fold-up
 '[When I] come home in the evening, [I] take in the clothes and [I] fold [them] up.'
 NS3:19a-20

In addition, it is not necessarily the case that overt pronouns mark thematic disjunction in Mandarin. Consider the following extract (from a letter):

25) 我们带他们参加了一次去房县旅游。

wǒmen dài tāmen cānjiāle yīcì qù Fángxiàn lǚyóu
 1pl lead 3pl attend.ASP one-time go F. travel
 'We took them on a trip to Fangxian,...'

一次去武当山市里也去了。

yīcì qù wǔdāngshān shìlǐ Ø yě qù le,
 one-time go W. town Ø also go-ASP
 '...and one to Wudangshan, [we all] also went into town,...'

参加私人的生日宴会。

Ø hái cānjiā-le yīgè wǎnhuì
 Ø and attend-ASP one.CL party,
 '...and Ø went to a party,...'

也到陈涓的单位去。

Ø yě dào Chénjuān-de dānwèi qù-le
 Ø also to C.-POSS unit go-ASP
 '...and Ø went to C.'s work unit...'

参加私人的生日宴会。

cānjiā sīrén-de shēngrì yànhuì
 attend private-DE birthday party
 '...to attend a private birthday function....'

我们还请了一餐，

wǒmen hái qǐng-le Ø yīcān,
 1pl also invite-ASP Ø one-meal
 '...we also invited [them] for a meal.'

总的看他们是满意的。

zǒngde kàn tāmen shì mǎnyì-de
 altogether look 3sg COP satisfied-DE

'All-in-all, they were very satisfied.'

NS7:4-9a

The underlined Object of the first clause refers to two participants newly introduced in the previous sentence. These two participants and the first clause Subject share Subjecthood, and are both unexpressed throughout the following eight clauses. But in the ninth clause, *wǒmen hái qǐngle yīcān* 'we also invited them once', one of them takes an agentive role, and the other is a patient. The agentive participant is overtly expressed once again (in bold). The overt NP is thus used to mark a change in the relationship of the **established and on-going participants** both to each other, and to the action, much as Givon describes (see above). In the tenth clause, the guests are Subject once more, and this argument is now overtly expressed (double-underline).

Far from being a point of thematic disjunction, this clause begins with an explicit discourse marker *zǒngde kàn* 'looking overall', introducing a consequence of all previous clauses, by way of summary. It is here, at the last point of continuity, rather than at a point of thematic disjunction, that the persistent discourse topic (the guests) is once again expressed as a lexical pronoun.

Thus, an overt pronoun can be used to signal switch agency, and may also have an emphatic role in a summary or conclusive sentence, punctuating the closing subtheme rather than opening a new one.

Though Givon's comments about the independence of participant and action continuity from thematic continuity are clearly relevant, his scalar model linking the smallest formal unit to the most accessible referent still leaves some central questions about interpretation unanswered: What **makes** one antecedent more accessible than another? And how do we choose a specific accessible antecedent from among many?

Empirically, zero anaphora in English has been shown to have a mean referential distance of one clause (Givon, 1983) and this sets a rough practical limit on potential antecedents, but it does not serve to identify a specific antecedent. Moreover, Williams (1989) found that zero anaphora has a significantly greater referential range in Singapore English, and the English of Chinese speaking learners. This is presumed to be a

consequence of the influence of Chinese which has regular definite ellipsis. Where referential distances are greater, more potential antecedents will be involved, and as we saw above, there may be more than one continuous topic, making interpretation potentially more difficult. The precise mechanisms by which the specific intended antecedent is selected remain to be identified. One important factor is topicality.

3.2 Topicality

Huang Yan (1994:175), discussing the distribution of zero anaphors with particular regard to correlative structures, observes that 'the choice between the use of a zero anaphor, and that of a pronoun ... seems to be largely dictated by topicness.'

Correlative structures are two-clause constructions with paired conjunctions, one appearing in each clause.

26) 小明越说他越高兴。

*Xiǎomíng₁ yuè shuō, {∅_{1/2} / tā_{*1/2}} yue gaōxìng*
 X more speak {∅ / 3sg} more happy
 'Xiaoming₁, the more [he₁] speaks, the happier {∅₁ / he₂ is}.'

也许小明会来也许(∅/他)不会来。

Yěxǔ Xiǎomíng₁ huì lái, yěxǔ {?∅/tā₁} bù huì lái
 perhaps X may come perhaps {?∅/3sg1 NEG may come
 'Perhaps X. will come, perhaps {?∅ / he} won't come.'

Huang Yan (1994:175) continues: "When both Subjects are in the topic position [i.e. preceding the conjunction - A.H.C.], zero anaphors are obligatory; when one Subject is in the topic position, zero anaphors are optional, when neither Subject is in the topic position, zero anaphors are dispreferred and pronouns are preferred.'

One problem with this description is that it is a moot point whether a ∅ Subject precedes or follows the correlative conjunction, since it is not observable. A second problem is that the correlation between topicality and ellipsis is not consistent for all correlative conjunctions. Table 15 sets out the patterns of grammaticality and reference for three different correlatives. A shaded cell means the structure is ungrammatical with any reference, a * means it is ungrammatical with the reference indicated in the left-hand column. The headings in the second row indicate constituent order: T=Topic CM =

Correlative Marker, S = Subject. From the table it is clear that it is simply not the case that a topical Subject demands obligatory ellipsis of the Subject in the second clause for all correlatives. In fact, correlatives vary as to whether or not they can take two overt topical or non-topical Subjects, and they vary with regard to the prescribed reference of two overt Subjects.

Table 15 Correlatives and overt NPs

Correlative	Topical Overt	Non-Topical	1 topic Ø	1 non-Topic
Disjoint	T CM T CM	CM S, CM S	T CM, CM	CM S, CM
<i>yīnwèi suǒyǐ</i>	*	√	*	√
<i>yuè yuè</i>	√		*	
<i>sūirán háishi</i>	?	√	*	√
Coreferent	T CM, T CM	CM S, CM S	T CM, CM	CM S, CM
<i>yīnwèi suǒyǐ</i>	*	√	√	*
<i>yuè yuè</i>	?		√	
<i>sūirán háishi</i>	√	*	√	*

3.3 Structural Conjunction

Givon (1983:8) points out that thematic continuity is most commonly coded - if at all - via conjunction or clause-subordination particles in the SVO or VSO typology (of, say, English [or Mandarin - AHC])....". In other words, ellipsis is associated with thematic continuity, because this is expressed via syntactic conjunction and subordination, and they involve ellipsis. In Foley and van Valin's view also (1984:266), ellipsis is associated with tight inter-clausal semantic relationships, which are expressed via syntactically inter-dependent structures.

Subordination 'is strictly defined in terms of the part-whole embedding relation.' (Foley and van Valin, 1984:243); embedding pertains when one clause is an argument of another (see Foley and van Valin, 1984:239), or when one clause is an adverbial modifier of the other (Foley and van Valin, 1984:249). According to Foley and van Valin (1984:266), in English subordination, the core arguments of the sub-ordinate clause are always fully specified, i.e. there is no ellipsis.

Co-subordination is considered a tighter relationship (Foley and van Valin, 1984:249); it does not involve embedding, but the two clauses are not entirely

independent. Co-subordination involves ellipsis, and co-subordinated clauses will be unable to have independent operators.

In Mandarin, however, there is no clear connection between the type and level of conjunction and the occurrence of ellipsis. The three correlative conjunctions just shown all produce structures which are dependent at the peripheral level. This can be seen from the fact that only one illocutionary force particle is possible per structure. Illocutionary force is an operator at the peripheral level (Foley and van Valin, 1984:249) so ability to take just one such particle demonstrates dependence at that level.

27) *他越唱歌吗越高兴吗?

**tā yuè chànggē ma yuè gāoxìng ma?*
 3sg more sing.song QUEST more happy QUEST
 "Is he the happier the more does he sing?"

28) *他虽然唱歌吗还是不高兴吗?

**tā sūirán chànggē ma háishì bù gāoxìng ma?*
 3sg although sing.song QUEST still NEG happy QUEST
 "Although does he sing is he nonetheless unhappy?"

29) *他因为唱歌吗所以高兴吗?

**tā yīnwèi chànggē ma suǒyǐ gāoxìng ma?*
 3sg because sing.song QUEST therefore happy QUEST
 "Because does he sing is he therefore happy?"

In 27) neither clause modifies the other, or is an argument of the other, so it is a co-subordinate structure; in 28) the first clause must be a statement, while the second may have an independent illocutionary force (i.e. as a question). The first clause is therefore less independent than the second, so 28) is a sub-ordinate structure; 29) is also a sub-ordinate structure according to the adverbial modifier criterion, but it is ambiguous as to which clause is affirmed and which is questioned.

Now, although 28) and 29) are more similar in terms of subordination type, it is 27) and 28) which are more similar in terms of ellipsis. The co-subordinate sentence 27) and the sub-ordinate sentence 28) both allow an overt Subject in the second clause, though they have different referential effects, see 30) and 31). The subordinate structure

in 29) is supposedly **less** closely linked, and should therefore also allow an overt Subject, but it does not, see 32):

30) 他越唱歌他越高兴吗?

tā₁ yuè chànggē (tā₂) yuè gāoxìng ma?
 3sg more sing.song 3sg more happy QUEST
 'Is it the case that the more he₁ sings, the happier he_{1/(2)} is?'

31) 他虽然唱歌他还是不高兴吗?

tā₁ sūirán chànggē (tā_{1/2}) hái shì bù gāoxìng ma?
 3sg although sing.song (3sg) still NEG happy QUEST
 'Although he₁ sings is he_{1/2} nonetheless unhappy?'

32) 他因为唱歌(*他)所以高兴吗?

*tā yīnwèi chànggē (*tā) suǒyǐ gāoxìng ma?*
 3sg because sing.song (*3sg) therefore happy QUEST
 'Is it because he **sings** that [he] is happy?'
 'Is [he] **happy** because he sings?'

It is clearly not the case that sub-ordination correlates simply with either the semantic tightness of inter-clausal links, or with the choice between ellipsis and overt expression.

Neither topicality nor a hierarchy of inter-clausal syntactic and semantic relatedness are adequate to account for the distribution of ellipsis in correlative structures. Similar problems arise in accounting for interpretation.

4. INTERPRETATION

4.1 'Free' Co-reference

Discourse theories tend to assume that interpretation is achieved by way of semantic and/or pragmatic constraints. Li and Thompson (1979:315 e.g. 3) showed that there is variability in the reference of ellipsed arguments in strings like 33):

33) 倾长老见智深肯去。

Qīng-zhǎng-lǎo jiàn Zhī-Shēn kěn qù
 abbot see Z-S willing go

就留在方丈里歇了……

Ø₁ jiù liú Ø₂ zài fāng-zhāng-lǐ xiē-le...
 Ø then keep Ø at chamber-loc rest-ASP

- a) 'The abbot_i saw Z-S_j was willing PRO_i to go, and PRO_i kept [him]_j in the room PRO_i to rest.'
 b) 'The abbot_i saw Z-S_j was willing PRO_j to go, and PRO_i kept to his room PRO_i resting.'

They also point out that it is sometimes a Subject NP which serves as the referent for a \emptyset Subject, as in 34a), but it is sometimes an Object as in 34d) (Li and Thompson (1979:317).

34a) 行主人应诺 \emptyset 取了水来。
hángzhǔ-rén₁ yìng-nuò. 34b) \emptyset_1 *qǔ-le* *shǔi lái...*
 owner concur \emptyset bring-ASP water come
 'The owner₁ concurred. (He₁) brought the water.'...

34c) 众人道：好了。
zhòng-rén dào: "hǎo-le"
 crowd said: fine-ASP
 'The crowd said: 'fine'.

34 d) 扶着立了起来。
 \emptyset *fú-zhe* \emptyset \emptyset *lì-le* *qǐlai*
 \emptyset support-ASP \emptyset \emptyset stand-ASP up
 '(They) propped (him) up. (He) stood up.'
 (Li and Thompson, 1979:317)

On the basis of these examples and others, Li and Thompson (1979:312) argue: there are no structural properties predicting the interpretation of the referent for zero-pronouns, but ... the interpretation of the referent for the unrealised pronoun is inferred on the basis of pragmatic knowledge.

Similarly, La Polla (1993:773) draws attention to the fact that, unlike in English, a Patient NP in a Mandarin conjoint construction can be deleted without passivisation:

35) 小狗走到山底下那个人就看见了。
xiǎo gǒu zǒu dào shān-dǐ-xià nèi-ge rén jiù kànjiàn \emptyset le
 small dog walk to mountain-bottom that-CL person then saw \emptyset ASP
 a) 'The little dog went downhill and was seen by the man.'
 (La Polla (1993, p 773 e.g. 5, original gloss.)

b) The dog went downhill and the man saw \emptyset . (Literal gloss.)

He argues that this is evidence that there is no Subject in Mandarin. Like Li and Thompson (1979), La Polla concludes (1993:774) that 'co-reference in Chinese is in fact quite free'.

The fallacy of such claims is easily demonstrated by consideration of instances where reference clearly is uniquely constrained. For instance, the co-reference constraints in 36) force an absurd interpretation:

36) 我洗了衣服晾出去。

wǒ xǐ-le yīfu liàng-chū-qù
 1sg wash-ASP clothes air-out-go
 'After I'd washed the clothes [I] went out to air.'
 *'After I'd washed the clothes [they] went out to air.'
 *'After I'd washed the clothes [I] put [them] out to air.'

This is because the \emptyset Subject must be coreferent with the Subject, not the Object, of a verb which is marked with the anterior aspect marker *le*. If co-reference were free, the absurd interpretation would be rejected in favour of a more pragmatically plausible one, but this is not the case. Since Subjecthood is a syntactic attribute, this is clearly an example of syntactic control.

Similarly, in 37) only the topic, and not the Subject or Object of the first clause can control the empty Object position of the second clause, and it can control only the Object position, not the Subject position, producing either incoherence or an absurd interpretation:

37 ?那个人，我盖了房子就拆开了。
nèige rén, wǒ gài-le fángzi jiù chāikai-le
 that-CL person, 1sg build-ASP house then demolish-ASP
 ?That person, I built the house, then demolished [him/her]
 * I built the house then that person demolished it

The ambiguity in Li and Thompson's examples at 33) and 34) arises not because of free co-reference, but because the verbs *liú* 'stay/keep' and *lì* 'stand' can be either transitive or intransitive. This means there is ambiguity as to the GF of the nouns, and it is this which produces the alternative interpretations.

In La Polla's example at 35) the ellipsis of an Object produces quite a different referential effect from the ellipsis of Subjects: it can be altered by a change in context, as shown in 38):

38) 山上有猫站在小狗后。
shānshàng yǒu māo zhàn zài xiǎo gǒu hòu.
 hill-top have cat stand at small dog behind.
 'On the hill-top there was a cat standing behind a small dog.'

小狗走到山底下，
xiǎo gǒu zǒu dào shān-dǐ-xià,
 small dog walk to mountain-bottom
 'The dog went downhill...'

那个人就看见了。
nèi-ge rén jiù kànjiàn Ø_i le
 that-CL person then saw Ø_i ASP
 '... and the man saw [it]_i.'

With no alteration to the sentence in question, the reference of the ellipsed post-verbal NP is now altered; such a change in context could not affect the reference of an NP ellipsed from pre-verbal position.

As to the question of Subject control, Tan (1991), presents a number of arguments in support of the existence of a Subject GF in Mandarin. She uses passive sentences, where Subject and semantic agent are distinct, and Topic sentences, where Subject and Topic are distinct, to investigate the relative accessibility of Subject, Agent and Topic to act as pivots in syntactic processes and finds that i) only Subject, and neither topic nor agent can control reflexives, or be the referent of imperatives, and that ii) Subjects **and** Agents, but **not** Topics can control adjuncts.

Referring to Li and Thompson's (1979) explanation of the reference of an ellipsed **Object**, Tan says (1991:40):

[T]here should not be any comparison between the Subject and the S[entence]-topic concerning coreferential deletion across sentences. What controls the deletion [of Objects] is the chain-topic⁴.

However, Tan does not investigate deletion in co-ordinate sentences, such as those discussed by La Polla (1993,1994). Applying Tan's tests to sentences like La Polla's (at 35) we find that, where topic and Subject are distinct, only one pattern of co-reference is possible: the immediately pre-verbal NP (conventionally the Subject) controls reference of the ellipsed agentive argument, while the topic NP controls reference of the ellipsed patients. Consider:

- 39) 那个人, 我看见 Ø 了 Ø 就认识 Ø 了。
nèi-ge rén_i, wǒ_j kàn-le Ø_i, Ø_j jiù rènshi Ø_i le
 that-CL person_i 1sg_j see-ASP Ø_i Ø_j then recognise Ø_i ASP
- a) 'That person_i, Ø_j having seen [him]_i, I_j recognise [him]_i.'
 b) * That person, I saw him and was recognised by him.

⁴ By 'chain-topic' Tan means a discourse topic that has been established in previous clauses, possibly across several sentences, and may have no overt realisation in the sentence where the ellipsis in question takes place. S-topic refers to a lexical topic in sentence initial position in the standard topic-comment sentence.

If the GF of Subject was irrelevant to coreferent deletion, as La Polla claims, we would expect ambiguity of reference. Even when the only potential controllee for the topic is the dependent Subject, Subject control is preferred: in 40) there are no \emptyset patients available for a topic to control, and the sentence is most likely to be judged incoherent:

- 40) ?那个人，我盖了房子就走。
 ?*nèi ge rén, wǒ gài-le fángzi \emptyset jiù zǒu*
 that CL person 1sg build-ASP house \emptyset then go
 a) *'That person, \emptyset having built the house, I will go.'
 b) ? 'That person, after I've built the house, will go.'

A marginal interpretation is possible, if the first clause is interpreted as a temporal adjunct interposed between the topic and the main verb (gloss b). In this case, the topic would then be indistinguishable from a conventional Subject, and there is actually no co-reference across clauses, since the topic-Subject is now internal to the main clause, not extra-posed.

In conclusion, the variability of reference of \emptyset arguments appears only in certain syntactic positions, moreover, the syntactic position of the \emptyset correlates with a specific controller: ellipsed Subjects tend to have a unique interpretation insensitive to context and are controlled by Subjects; ellipsed Objects, show a consistent sensitivity to context and this is exemplified particularly by a sensitivity to control by sentence Topics. This suggests that there are discrete mechanisms of referential control involved in the determination of reference of ellipsed Subjects and ellipsed Objects in these structures.

The variability of co-reference in Mandarin is clearly structurally conditioned, not free or random.

4.2 Neo-Gricean Pragmatics

The only discourse theories which make a serious attempt to explain the interpretation of ellipsis are the neo-Gricean accounts.

Reference in neo-Gricean accounts is seen primarily as binary. Reinhart (1983) applying a principle of disambiguation suggested by Dowty (1980) proposes that anaphoric terms be understood as belonging to contrastive sets; one member of the set

having a grammatically determined specific co-reference pattern, and the others expressing the 'complement'. Four different referential forms are thus thought to produce only a two-way contrast of reference.

Regardless of which form they consider as basic, most neo-Gricean accounts make use of GB's anaphoric and pronominal binding principles A and B (derived from GB theory, Chomsky, 1981) as the basis of grammatical binding. These involve the concepts of C-command and Governing Categories

4.2.1 C-command and Governing Categories

Strict c-command is a structural relationship between two NPs where the first branching node above the commanding NP dominates the other NP (Reinhart, 1976, Chomsky, 1981). A transitive verb could not be said to c-command its PP in strict terms, because an Object NP would intervene (as the first branching node) between the verb and PP. Such a distinction between verb-types is counter-intuitive (for more detail see Haegeman, 1991), and Chomsky (1981) proposes that c-command pertains in a broad sense, when the first (and consequently every other) 'maximal projection' above the commanding NP dominates the other NP.

A Governing Category (GC) is the smallest constituent containing the anaphor, its 'governor' and an 'accessible Subject'. A potential governor is the lexical head or INFL (the notional node that dominates the tense and agreement features of a verb) of a phrasal constituent. To actually govern a specific NP, a potential governor must be the head of the closest phrasal constituent dominating the NP (m-command). Accessible Subjects are lexical Subjects, (or + tense infl) so long as they are not barred on account of a circularity of co-reference, as for instance when the anaphor is contained within an NP which is itself the Subject of the constituent containing the anaphor.

Reinhart (1983:158) proposes a coindexing rule, which is basically an amalgam of the GB binding principles:

- Coindex a pronoun P with a c-commanding NP α (α not immediately dominated by Comp or S').
Conditions: (a) If P is an R-pronoun (i.e. Anaphor) α must be in its minimal GC
(b) If P is a non-R-pronoun (i.e. pronominal) α must be outside its minimal GC.

The 'neo-Gricean' accounts of ellipsis generally take one formal device, \emptyset , pronouns, or reflexives, as being anaphoric or pronominal in the sense defined above. They then argue that other anaphoric devices will contrast with that which is held to be grammatically determined in ways which can be predicted by reference to Gricean principles such as co-operation, economy of expression, and informativeness (see Grice, 1975).

Levinson (1987a,b) presents a model based on these principles, in which reflexives are held to be grammatically determined as coreferent with their clause-mate co-argument. Pronouns and null arguments are then predicted to contrast with the reference of the reflexive and so be pragmatically preferred to be non-coreferent. So in:

41) Chomsky likes him

the pronoun '*him*' is held to be potentially ambiguous, i.e. it could refer to 'Chomsky' (see Huang Yan, 1994:116 for a discussion).

The criticism has been made (see for example Huang Yan, 1994) that Levinson's model based on a coreferent reflexive, would not predict the occurrence of an overt pronoun which is also locally coreferent, as for example in discourse complements (from Huang Yan 1994: 123):

42) 王先生说他游览过长城。

Wáng xiānshēng shuō tā yǒulán-guò Chángchéng

Wang Mr say 3sg tour-ASP Great-Wall

'Mr Wang₁ says he_{1/2} has visited the Great Wall.'

Levinson (1991) subsequently presents two alternative models, the first taking the reference of pronominals, instead of reflexives, as basic and the second suggesting that the contrast between anaphoric devices need not always be a referential contrast, but could be a contrast in 'logophoricity' or speaker empathy. However these models suffer from comparable problems (see Huang Yan 1994).

Huang Yan's (1994) account, does not refer to binding principles; for Huang Yan (1994) the \emptyset has a 'basic tendency' to be locally coreferent and an overt pronoun to contrast with this. Local co-reference of \emptyset is considered to be an immediate consequence

of Grice's Information-principle, 'say as little as necessary': since speakers say as little as necessary, and \emptyset is semantically least specific, it can be assumed that \emptyset has the most stereotypical, expected, or unmarked interpretation. According to Huang Yan (1994:128) 'except in cases where antecedent and anaphor are clausemate co-arguments, the preference for co-reference is simply stronger than that for non-co-reference'. Where they are clausemates, the Disjoint Reference Presumption (DRP) overrides this preference. The DRP is an idea, formalised by Farmer and Harnish (1987) that there is a preference for non-co-reference between arguments of the same verb.

Huang Yan's definition of 'local', specifically allows a non-c-commanding NP as a local antecedent. Huang Yan (1994) states that the reflexive in the following example is referentially dependent on the possessor NP, which does not c-command it.

- 43) 小明的坏脾气给自己带来了许多麻烦。
Xiǎomíng₁ de huài píqi gěi zìjǐ₁ dài láile xǔduō máfan
 X DE bad temper to self bring much trouble
 'Xiaoming's bad temper has brought much trouble on [him]self.'
 (Huang 1994:131 e.g. 5.39c)

Huang (1994) addresses the selection of a specific antecedent, from among all potential antecedents, by augmenting the basic neo-Gricean approach with an Antecedent Search Procedure (ASP) outlining the sequence in which potential antecedents are assessed as to semantic and pragmatic compatibility. This antecedent search procedure can be represented:

Where n = the number of clauses in the sentence, and failure results from semantic incompatibility, the sequence of selection is: local Subject > local Object > local Subject plus Object > matrix $_{\alpha}$ Subject > matrix $_{\alpha}$ Object > matrix $_{\alpha}$ Subject + Object > matrix $_{\alpha+1}$ Subject > matrix $_{\alpha+1}$ Object ... > matrix $_n$ Subject plus Object, and only if this fails > nearest antecedent (preferably a topic) > arbitrary.

This sequence predicts that an NP later in the sequence will only be selected if all higher NPs are eliminated for semantic or pragmatic reasons.

All the co-reference patterns predicted by Huang's (1994) account are considered only as preferred, not obligatory, and as his examples indicate clearly, there are frequent exceptions. He presents explanations of such exceptions with respect to 8 different syntactic constructions (Huang Yan, 1994). Table 16 shows which of Huang Yan's two

predictions (i.e. local co-reference for zero, and contrastive reference for zero vs an overt pronoun) hold true for each construction. Relative clauses are listed twice since different patterns are found for the relativised constituent and other arguments. Where there is a departure from the predicted values, Huang Yan's explanation is noted briefly.

Table 16. Referential Patterns in 8 Syntactic Structures (Based on Huang Yan's 1994 analysis.)

Construction	∅ local coref	∅ contrast with pronoun	Huang's explanation	Alternative explanation
relative clauses: heads	yes	no pronoun allowed		
control constructions,	yes	some	Verbal semantics/pragmatics preclude pronoun in some places	Structural Ambiguity
correlative constructions,	yes	some	tight interclausal semantic linkage forces pronoun to co-reference in some cases	sentence topic scope and control
topic-comment structures;	yes	no	overt pronouns ignored; for ∅ preferred subj-> subj mapping, semantics sometimes override	sentence topic control
adjunctive constructions,	no	some	none: ignores non-coreferent ∅	discourse topic control
Subjects in s-complement,	no	no	tight interclausal linkage forces pronoun to coref	∅ ambiguous: arbitrary / discourse topic control
simple transitive sentences	no	no	DRP forces ∅ to non-co-reference	
Objects in complements.	no	no	ignores ∅ disjoint reference, 'matrix wins' forces pronoun to co-reference	discourse topic control
relative clauses non-heads	no	no	ignores disjoint reference of ∅	discourse topic control

From this summary table, we can see that Huang Yan (1994) accounts for overlapping referential patterns in the following ways:

- a) ignoring cases where ∅ may have a disjoint reference, claiming that the coreferent interpretation is simply preferred in a 'neutral context' (1994:130); ignoring some cases where an overt pronoun has local co-reference, e.g. in topic-comment constructions (1994: 159-69).

- b) arguing that a \emptyset may be forced to disjoint reference, (thus neutralizing its contrast with a pronoun) because of
- i) the Disjoint Reference Presumption (Huang 1994: 129), that clause-mate co-arguments should not be coreferent, or because of
 - ii) specific lexical semantics (Huang, 1994: 153);
- c) arguing that a pronoun is forced to local co-reference (thus neutralising its contrast with \emptyset) because of
- i) a 'tight inter-clausal relationship' (Huang, 1994:141 after Foley and von Valin, 1984); or
 - ii) a 'matrix wins' constraint (Huang 1994: 141; Levinson 1983: 142-3, 224), whereby whatever reference is implied by the matrix clause wins over the reference implied in a lower clause.

Such a large number of exceptional cases is rather problematic for the theory.

Moreover, the 'neutral context' and the 'forced lack of contrast' are also either problematic, or not generalisable, as shown below.

4.2.2 'Preferred' Interpretations in 'Neutral' contexts.

Huang (1994:130) argues that the more basic interpretation is that which is ascribed in a 'neutral context':

By 'preferred interpretation' is meant the interpretation that is the most favoured one out of a number of other possible interpretations. This interpretation arises without any particular context or specific scenario being necessary. Put in slightly different terms, in terms of the notion of context, the preferred interpretation arises in a default/unmarked context rather than in a specific/marked one. In other words, it is an instance of Grice's generalised (i.e. default) conversational implicature rather than particularised (i.e. context-deduced) conversational implicature.

It should be recognised however that a de-contextualised reading is far from neutral: it forcibly removes all potential antecedents **except** local ones.

In 44) Huang Yan claims (1994:132) that 'by the I-principle, the zero anaphor will be preferentially coreferential with the matrix Subject.'

- 44) 小明说下个月结婚。
Xiǎomíng shuō ∅ xià-ge yuè jiéhūn
 Xiaoming say ∅ next-CL month marry
 'Xiaoming says that (he_{1/2}/I/you/we/they...) will get married next month.'
 (Huang's example 5.31 1994:132)

Actually the I-principle simply states that the interpretation of \emptyset will be the most stereotypical; the stereotypicality of local co-reference is a premise of the model, not something established on independent grounds. Any 'preference' for the embedded Subject's being read as coreferent with the matrix Subject is arguably a consequence of the matrix Subject's being the only potential antecedent in the text. In fact, in natural discourse (surely a more 'normal' context than linguistic exposition) any established discourse topic could serve as antecedent here, just as it could for an overt pronoun.

Consider:

- 45) 老明很着急。小明说他下个月结婚。
Lǎomíng hěn zháojí. Xiǎomíng shuō tā xià-ge yuè jiéhūn
 Laoming COP anxious. Xiaoming say 3sg next-CL month marry
 'Laoming₁ is anxious. Xiaoming₂ says that he_{1/2} will get married next month.'

Here the presence of two potential antecedents reduces the likelihood that some third, more distant NP, is the antecedent, but it does not resolve the basic ambiguity between a local and an external antecedent, regardless of the form of the anaphoric device.

A truly preferred interpretation would be one which is consistently preferred **even in the presence** of a number of potential antecedents, both local and distant, including antecedents beyond the sentence. So Huang Yan's (1994) basic premise that local co-reference is the preferred interpretation for \emptyset , which underlies his whole argument that \emptyset contrasts referentially with a pronoun, is not well established.

4.2.3 Lack of Contrast between \emptyset and Pronoun

Huang Yan's (1994) claim that there is something **basic** about a referential contrast between \emptyset and an overt pronoun is also debatable. Firstly he claims that a zero Object in a simple transitive clause can be interpreted as reflexive, where in fact this contravenes

the DRP, and is not borne out by textual evidence: no instance of a reflexive \emptyset was found in the corpus.

Secondly, apart from restrictions on reference resulting from ϕ features, the reference of a \emptyset and an overt Object pronoun in a simple transitive sentence are virtually identical, in fact, when the entity is inanimate, ellipsis is more common than a pronoun:

- 46) 他去找了 (\emptyset / 她) 一趟没找着 (\emptyset / 她)。
tā qù zhǎo-le \emptyset /tā yītàng, \emptyset méi zhǎo-zháo \emptyset /tā
 3sg go search-ASP \emptyset /3sg once \emptyset NEG.ASP search-ASP \emptyset /3sg
 'He went looking (for her/it) once, [but] didn't find (her/it).'

NS8:12

Similarly in Subject position of a sentential complement of a speech act verb, there is no contrast between \emptyset and an overt pronoun, as Huang Yan (1994:269) acknowledges:

- 47) 张三说 (\emptyset / 他) 不认识李四。
Zhāngsān shuō \emptyset /tā bù rènshi Lǐsì
 Z. say \emptyset /3sg NEG recognise L.
 'Zhangsan says he_{1/2} doesn't know Lisi.'

He explains this variation from his predictions by recourse (1984:269) to Foley and van Valin's Interclausal Semantic Relations Hierarchy. On this hierarchy, indirect discourse complements constitute a tightly linked semantic unit, and Huang Yan suggests (1994:141, after Levinson (1987a)), that this 'give[s] rise to a "same agent/patient as the last clause" effect, hence the suspension of the M-implicature.'

M-implicature refers to an implicature of manner, that a more marked (overt) form is used to contrast with the stereotypical coreferent interpretation of \emptyset . The problem with this is that for both the \emptyset and the overt pronoun, the 'same agent effect' is only optional as 47) above shows.

4.2.4 Objects in Sentential Complements

Similarly, Huang Yan (1994) invokes a 'matrix wins' principle to account for the fact that an overt Object in a sentential complement can refer to the matrix Subject, the same as a \emptyset in:

- 48) 羊大娘担心女儿不肯伺候。
Yáng dàniáng dānxīn nǚér bù kěn cìhou \emptyset / Pron.
 Yang Grandma worry daughter not willing look.after \emptyset / Pron.
 'Grandma Yang₁ is worried that her daughter₂ is not willing to look after (her_{1/3}/ me/you etc...)'

In fact, both the \emptyset and the pronoun could also refer outside the sentence to any established discourse topic. In this case, there would again be no contrast, but the matrix would clearly 'lose'.

4.2.5 Arbitrary Reference when an Antecedent is Available

Finally, Huang Yan's ASP (1994) would rule out examples where reference is to an external or arbitrary NP even when there is a compatible NP locally. In the following extract from the corpus, two verbs in the final sentence have indefinite Subjects despite the referential continuity with the first (and only) lexically represented participant of other zero Subjects in the text.

- 49) 我好想好想看看沙漠，
wǒ₁ hǎo xiǎng hǎo xiǎng kàn-kàn shāmò
 1sg much want much want see-RED desert,
 'I really really want to see the desert.'

Ø想躺在那儿上面感受一下
Ø₁xiǎng qīnyǎn kàn nà cháng cāngmáng
 Ø want own.eyes see that expanse emptiness
 'want to see that expanse of emptiness with [my] own eyes'

想躺在那上面感受一下
Ø₁xiǎng tǎng zài nà shàngmiàn gǎnshòu yíxià
 Ø want lie at that top.face experience a.while
 'want to lie on top of it and feel for a while...'

那份一个
nà fèn yíge
 that share one-CL
 'that...'

独外在苍凉地方的孤独之感，
dúwài zài cāngliáng dìfang de gūdú zhī gǎn
 solitary at remote place POSS lonely ATT feeling
 '...isolation of being all alone in such a remote place..'

那份说不清、
nà fèn Ø_{arb} shuō bù qīng
 that share Ø speak NEG clear
 'that feeling which no[-one] can convey...'

道不明的感觉。

Ø_{arb} dào bù míng de gǎnjué

Ø tell NEG clear ATT feeling

'...no[one] can express.'

NS4:20-29

In fact the only cases where Huang Yan 's (1994) 'preferred' interpretation of Ø as locally coreferent actually holds consistently is in relative clauses, dependent clauses, and Ø arguments that are coreferent with a lexical sentence topic, all of which are prime candidates for some form of syntactic control; and the exceptions which Huang Yan (1994) observes can be readily accounted for by control by discourse topics, as we shall see more clearly below.

5. SUMMARY

5.1 Referential Patterns

The discourse approaches of Li and Thompson (1979), Chen Ping (1984) and La Polla (1993,1994), all recognise a variability in referential effect, but they neither describe the variability in general terms, nor relate it to the existence of structural subtypes of ellipsis.

Huang Yan's (1994) neo-Gricean model does give some recognition to this phenomenon, but sees it as a consequence of semantic rather than syntactic factors. In many cases though, these semantic factors account only poorly for the observed variation. The neo-Gricean accounts claim ambiguity where arguably there is none, and claim a single preferred interpretation where arguably there is ambiguity.

A fundamental problem with the neo-Gricean approach lies in the portrayal of reference as reflecting two basic patterns, the grammatically determined pattern, and everything else, when there are 4 different anaphoric forms: Ø, pronouns, nouns, and reflexives, providing a potential four-way contrast. This inevitably leads to referential overlaps, which weakens considerably the basic premise of contrastiveness on which the models are based.

Huang Yan (1994) makes a partial escape from this predicament by proposing that reference is a matter only of preferences which can be readily over-ridden by semantics and pragmatics. He postulates a norm of local co-reference for \emptyset and addresses the selection of a specific antecedent via an antecedent search procedure. He explains exceptions by reference to semantic incompatibility, variations in the strength of interclausal semantic links, and topicality. However Huang Yan (1994) bases many interpretations on claims of a 'preferred interpretation' in a 'neutral context'. In fact interpretative preferences and contextual neutrality are both difficult to substantiate. Nonetheless, on the basis of these, Huang ignores ambiguities and variations in reference which arise in discourse.

Huang Yan's (1994) model comes to depend to such an extent on judgements as to preferred interpretations, and on lexical semantics and topicality, that the importance to the whole process of basic neo-Gricean principles of economy and quality becomes rather questionable. One of the main weaknesses of Huang Yan's (1994) analysis is that it offers no grounds for a generalisation as to which structural contexts allow this interplay of forces and why, and leaves us with little explanation for the rigidity of interpretation which does arise in many cases.

5.2 Distribution

Discourse approaches have not been successful in describing the distribution of ellipsis in general terms; the main attempt to do so is on the basis of discourse function (in terms of conjoinability), (Li and Thompson, 1979 and Chen, 1984). Huang Yan's (1994) neo-Gricean account makes reference to Foley and van Valin's (1984) observation that the likelihood of ellipsis increases with the 'tightness' of semantic relations between clauses. There does seem to be a relationship between ellipsis and an implication of interclausal conjunction, however not all instances of ellipsis produce this effect.

5.3 Interpretation

The main claims of the discourse approaches are that a) co-reference is free; and b) reference is not controlled by Subjects. Counter-evidence has been presented to show that

the variability of reference is in fact structure-dependent, not random, and that two kinds of reference are evident: context-sensitive reference, which presents no single consistent relationship between referent and anaphoric device, and uniquely constrained reference, where reference is to a specific antecedent only, and semantic contradictions or implausibility cannot alter interpretation. The former tends to be exhibited by \emptyset Objects and the latter by \emptyset Subjects, moreover, Subjects in the structures discussed so far are generally controlled by Subjects, whereas Objects may be controlled by sentence topics or discourse topics.

Chapter Four

Ellipsis in Syntactic Theory

A distinction is needed between pragmatic co-reference and bound anaphora, i.e. the interpretation of pronouns as bound variables, with only the latter directly restricted by sentence-level rules. Consequently, the problem of definite NP co-reference which was the major issue in my dissertation (and was summarised in Reinhart, 1981a), was incorrectly stated to begin with, and many of the facts in this area actually fall outside the grammar.

Tanya Reinhart (1983, Preface)

1. INTRODUCTION

This Chapter begins with a discussion of Shopen's (1972) division of ellipsis into definite and indefinite ellipsis, and positions the current study with respect to that framework. It then reviews the position of the two main syntactic theories Government and Binding Theory (GB) and Lexical Functional Grammar (LFG) with respect to ellipsis. Once again the discussion is organised around the topics of referential effect, distribution and interpretation. The LFG notions of functional and anaphoric control and f-command are discussed.

The main points to be made in this chapter are:

Referential patterns

- 1) There are observable instances of ellipsis which do not fit standard syntactic classifications of empty categories.
- 2) Syntactic models of anaphoric reference have either claimed, quite wrongly, that ellipsed elements cannot refer beyond the sentence, or have admitted that some can, but, since these are outside the domain of syntax, have not directed further attention to them. They tend to assume that where there is no same-sentence antecedent, \emptyset s will have arbitrary reference. This is not the case for Mandarin.

Distribution

- 3) Standard syntactic classifications of ellipsis rely heavily on morphological cues that are absent in Mandarin, this means the distribution of proposed categories of ellipsis is not directly observable in that language.

Interpretation

4) While core GFs (Subject, Object, OBJ2), and adjuncts (sentence topic, possessives, *for*-Objects) are all considered as potential determiners of reference in one or other of these models, the role of a syntactically independent discourse topic is generally overlooked.

2. *DEFINITE AND INDEFINITE REFERENCE:*

Shopen (1972, 1973) defines and discusses two kinds of ellipsis: functional and constituent ellipsis. The former refers to ellipsis of lexical items which have arguments (verbs, prepositions, some nouns) and the latter to the arguments themselves. The focus of my study falls within Shopen's constituent ellipsis. However, my definition of ellipsis is much narrower than Shopen's. By Shopen's definition (1973: 116) "we have constituent ellipsis whenever there is no phonetically realised constituent corresponding to an argument in the propositional meaning of an utterance". His definition of argument (1972: 119) is "a sub-part of propositional meaning capable of syntactic manifestation" and the latter "is to be taken in a broad sense - i.e. 'possible in the language' rather than 'possible in the utterance type' or 'possible in the constructions governed by a particular item'" (1972:124).

Within this broad scope, Shopen (1972) identifies sub-classes of obligatory vs optional ellipsis, and definite vs indefinite ellipsis. Shopen's obligatory category is very different from that used in this study: his is based upon the idea that lexical items form semantic groups which express the same basic proposition. The proposition always has the same semantic arguments, but some lexical members of that sub-group do not permit 'syntactic realisation' of certain arguments. He illustrates this (1972:125) with the unacceptability of the collocation:

50) *the burglary of the crown jewels

Shopen says (1972: 125) that 'burglary' is specified for the obligatory ellipsis of the 'booty' (theme) argument of lexical items belonging to a 'larceny' group. This is because themes can be readily expressed in relation to other semantically similar English words,

yet 'burglary' and its theme cannot co-occur in the same sentence⁵. So Shopen's 'obligatory ellipsis' refers to the inability of a specific lexical item to collocate, within a single sentence, with the overt expression of a given (semantic) argument, compared to the potential for expression of that semantic argument in the language as a whole.

In contrast, the category of obligatory ellipsis used in this work relates to the unacceptability of some overt argument of a specific lexical item in one context, even though the same argument of the same lexical item can be expressed in another context. Shopen's obligatory ellipsis would not constitute ellipsis at all in my framework.

Optional ellipsis in Shopen's framework includes the ellipsis of both arguments which specify agent, theme or patient, such as those in square brackets in 51), and those which specify details like purpose, instrumentality etc., such as those in round brackets.

51) '[I] ate [fish and chips] already, (out of newspaper, with a knife and fork, at Judy's, an hour ago)'.

In my framework, optional non-realisation only constitutes **ellipsis** when the referent is retrievable, or in Shopen's terms, definite; where a specific referent is unretrievable, I would say that reference is simply unspecified.

Fillmore (1969) proposed the need to distinguish between definite and indefinite interpretations of ellipsis. Definite ellipsis refers to omission of an NP which is nonetheless understood as a reference to a specific entity retrievable from discourse, as in 52) below. Indefinite ellipsis refers to a generic or non-specific entity which is not present in discourse, as in 53):

52) Ø Going to the movies? (Ø = 'you')

53) Ø Going to the movies is a popular pastime. (Indefinite, Ø = anyone).

Shopen (1972, 1973) also makes reference to this distinction. This corresponds to my notion of retrievability. It also provides a terminological 'handle' for refining the

⁵ Though what about: 'The burglary I refer to was that of the crown jewels.'?

It seems Shopen may consider this a circumlocution (cf his discussion of 'explanation' (1972: 300)). The relationship between circumlocution and 'possible syntactic manifestation' is not made clear.

problem of what 'fully specified' means. Oblique arguments, e.g. possessors, may be unexpressed, but they are not, it seems readily ellipsed with definite reference.

Core arguments may be ellipsed with either definite or indefinite reference. 'Fully-specified' can thus be usefully defined as 'having an overt representation of all arguments which have definite reference', this allows for certain semantic roles to be unassigned without being counted as ellipsis, in my terms, or needing to account for their interpretation.

Shopen points out (1972, 1973) that Object ellipsis in English is lexically constrained: the acceptability and reference of absent Objects varies with different lexical items in the same syntactic and discourse context:

- 54) a - Watch me!
- I'm (already) watching.
b - Eat your dinner
- I already ate.
c Help me look for your glove!
I already found *(it).

In 54a) the non-lexicalised Object is coreferent with the 'me' of the command, in b) co-reference is impossible and the actual patient of 'I ate' is unspecified, and in c) ellipsis is unacceptable. Unlike Object (and Oblique) ellipsis, Subject ellipsis, does not appear to be subject to lexical variation, but appears to be constrained by syntax and discourse factors, not lexical ones:

- 55) Speaker A₁: What did Tim do yesterday?
Speaker B: Ø Went to Tilley's.
Speaker A₂: * Did you?
56) A: What did Tim do yesterday?
B: I already said that {*Ø /he} went to Tilley's.

The Subject of the lexical item 'went' can be ellipsed in one context, but not in another. In English, Subject ellipsis appears to have a special relationship to turn-taking in dialogue, and to specific syntactic structures. Note that a Ø Subject of an answer (55B) must be coreferent with the Subject of the question it answers, so 55A₂) is bad.

By contrast, Mandarin freely allows both definite and indefinite Subject ellipsis, regardless of syntactic context. And though Mandarin allows even less in the way of

indefinite Object ellipsis (i.e. intransitive valencies of normally transitive verbs) than English, it also freely allows definite Object ellipsis. This characteristic is often referred to as 'pro-drop'.

Shopen's approach illuminates the question of how we know a constituent is missing, and how we know what role that missing constituent plays in the discourse, but it does not bear directly on the question of how we **interpret** the reference of that constituent. GB and LFG do concern themselves with the question of interpretation.

3. GOVERNMENT AND BINDING THEORY (GB)

3.1 Referential Categories

GB takes the position that NPs can be classified on the basis of the structural relationship they have to other coreferent NPs. It identifies four kinds of empty categories (ECs); each is analysed as a combination of two independent referential features, anaphor and pronominal:

Table 17 GB's Empty Categories

Empty Category	OVERT equivalent	referential features
PRO	NONE	+anaphor +pronominal
NP trace	anaphors (reflexives and reciprocals)	+ anaphor -pronominal
wh trace	Referring expressions (Proper and common nouns)	-anaphor -pronominal
pro	pronominals (personal pronouns)	-anaphor +pronominal

3.1.1 Anaphors and Pronominals: Definitions

The definitions of the referential features Anaphor and Pronominal depend upon the concepts of c-command and Governing Category (GC). These have been defined on page 49 above.

According to Principle A of GB, (Chomsky 1981), an anaphor has no independent reference and must be bound within its GC. Pronominals, on the other hand, must be free (more specifically, not bound by an argument) within their governing category (Principle

B, see Chomsky, 1981). To be 'bound' means to be constrained to co-reference with an antecedent which c-commands the anaphor.

3.1.2 PRO and pro

As can be seen from Table 17, the 4 empty categories result from combining plus and minus values for the two referential features. Little pro, is -anaphor +pronominal; -anaphor means it does not need to refer inside its GC, and + pronominal means it must refer outside its GC. 'Little pro' is purportedly common in so-called 'pro-drop' languages like Mandarin (e.g. 57 represented by \emptyset), but can be found also in colloquial English (58):

- 57) 真想跟你聊聊。
 \emptyset *zhēn* *xiǎng* *gēn* *nǐ* *liáoliáo*
 pro really want with you chat
 '[I]'d really like to chat with you.'
 NS1:1

- 58) 'pro Going to the movies?' (pro = 'you')

In fact, there is some doubt as to whether these instances of ellipsis are properly classified as pro, because they do not meet the licensing constraints which GB postulates for pro (see 'Distribution' below).

Big PRO, on the other hand, is +anaphor, +pronominal, so theoretically it must be both free and bound within its GC, a logical impossibility. This contradiction is 'resolved' in GB by proposing that PRO be ungoverned and therefore without a GC. The main support for this proposal is that PRO is said to appear only in positions where a lexical NP cannot. (For arguments against an empty category having positive value for both features see Manzini, 1983; Bouchard, 1983; Aoun, 1985, Borer, 1989, Chung 1989). 'Big PRO' includes both Arbitrary PRO (59) and controlled reference, (60):

- 59) Going to the movies is a popular pastime. (\emptyset = anyone).
 60) I_i hate \emptyset _i going to the movies. (\emptyset is controlled by 'I')

In the absence of a GC, the question of PRO's reference is resolved by invoking a quite separate mechanism of 'control' (see below). This mechanism however allows the

domain in which PRO finds its antecedent to sometimes coincide with that for pro, while at other times it remains distinct.

The result is, on the one hand, two categories, pro and PRO which are not mutually exclusive in terms of referential effect, and on the other hand one category PRO which is not internally uniform in terms of referential effect. In other words, the GB classification of empty categories does not produce a neat system of categories with complementary patterns of reference as it does for overt NPs.

Though this is clearly not the most efficient system, there's no reason to suppose that languages do utilise the most efficient systems and the overlap of referential effect might not matter so long as the two categories, and the two sub-categories within PRO are all distributionally distinct, so that in any case the language user knows which interpretative procedure to employ. But, as we will see below, for Mandarin at least, this is not always the case.

3.1.3 Traces

Traces are -pronominal and must be governed by an antecedent. We will be concerned only with wh-traces and their role in relative clauses and Topic-Comment structures.

3.2 Distribution of GB's Empty Categories

3.2.1 Licensing constraints

3.2.1.1 Big PRO

The requirement that PRO be ungoverned, has the consequence (within the GB framework) that PRO is in complementary distribution with lexical NPs, as the latter must always be governed. For example, PRO cannot appear as Subject of a tensed verb, where a lexical Subject is permitted (or according to GB required):

- 61) { * PRO/ he } should invite the sergeant
(Haegeman 1991:46)

3.2.1.2 Little pro

The distribution of little *pro* is limited only indirectly in GB. This can happen as a consequence of constraints on the structural relationships deemed to hold between it and its antecedent. For the most part these simply exclude a small set of local NPs, but do not narrow the field of long-distance antecedents. More significant is the requirement that *pro* be licensed, for instance by the 'pro-drop parameter' (Rizzi 1986):

62) Pro-Drop parameter (Rizzi, 1986)

- a) *pro* is governed by X^0 ,
- b) Let X be the licensing head of an occurrence of *pro*; then *pro* has the grammatical specification of the features on X co-indexed with it.

One important consequence of the Pro-drop parameter is that X can only govern *pro* if X is in the same sentence as *pro*. This means that in a standard GB account, \emptyset elements which refer beyond the sentence cannot be little *pro*. This constitutes a major problem for a GB-based analysis of ellipsis in Mandarin, because ellipsis in Mandarin routinely refers beyond the sentence.

A later formulation of a Generalised Control Rule, by James Huang (1989) allows that some instances of PRO/*pro* may not have a control domain, and these can refer beyond the sentence, thus allowing a less restricted set of \emptyset elements to be classified as *pro*/PRO. This broader licensing however also allows a certain set of uncontrolled \emptyset elements whose reference is not accounted for within binding or control theory.

3.3 Interpretation: Mechanisms of Binding and Control

3.3.1 Binding

As stated above, anaphors must be 'bound' inside their governing category, while pronominals must not be. Of GB's 4 empty categories PRO has no Governing Category (see above), so binding is irrelevant to its interpretation. Traces are anaphors, and must be bound within their GC, while little *pro* is a pronominal, and must be free within its governing category.

3.3.2 Control

The reference of PRO is determined via 'control', the formal expression of which was formulated by Bresnan (1982:372):

The term **control** is used to refer to a relation of referential dependency between an unexpressed Subject (the **controlled** element) and an expressed or unexpressed constituent (the **controller**). The referential properties of the controlled element...are determined by those of the controller.

Control is optional in certain syntactic positions. According to Haegeman (1991:244), PRO 'may be taken to refer to a specific referent ('you', 'they', etc.) or it may be interpreted as equivalent to the arbitrary pronoun 'one'.'; when PRO is interpreted as referentially dependent on another NP in the sentence it is said to be controlled; when PRO is not controlled it has arbitrary reference (PRO_{arb}).

By and large, sites for optional control are identified by observation of reference. For an explanation of how that reference is determined, independent guidelines as to what characterises sites for PRO and for optional control are required. A recurring theme in the following discussion will be that both GB and LFG analyses rely heavily on morphological cues, all of which are lacking in Mandarin.

3.3.2.1 Obligatory Control.

PRO may be obligatorily controlled by an NP, as in (these two examples from Haegeman, 1991):

63) 'Poirot agreed PRO to go'

or implicitly, as in :

64) 'The operation was abandoned PRO to save money.'

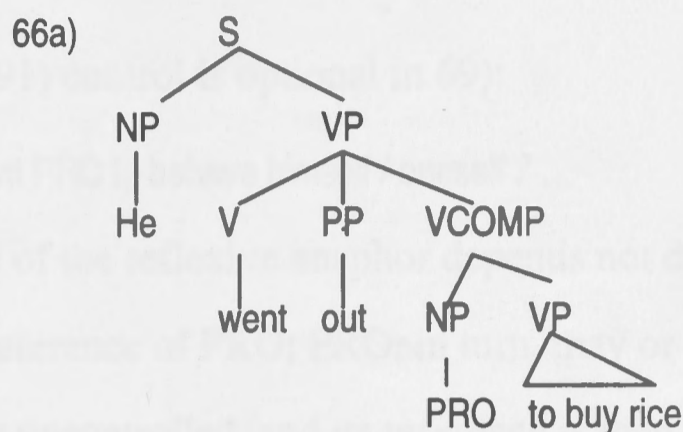
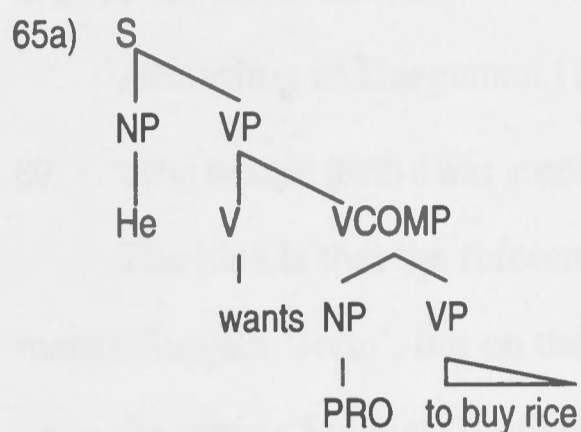
In 64) PRO is not considered to be arbitrary because the money saver must be the agent of 'abandon'. In obligatory control by an NP, the controller must c-command the controllee (PRO). Consider the following sentences:

65) He_i wants PRO_i to buy some rice.

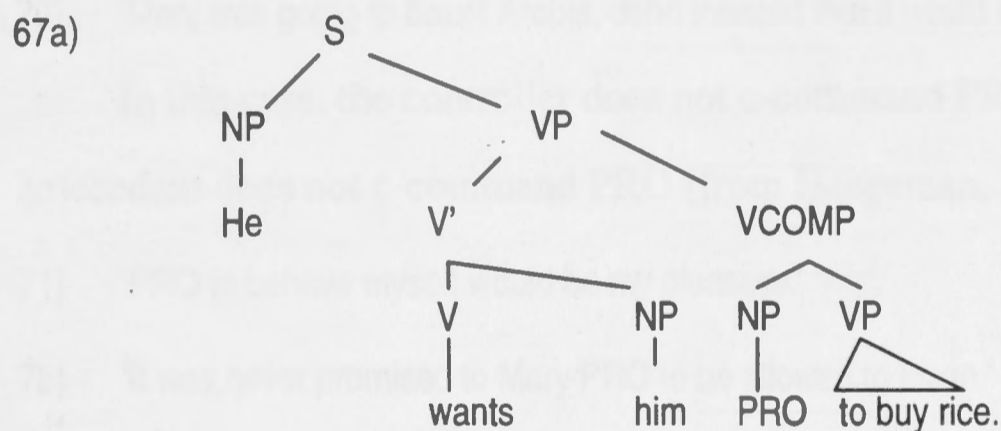
66) He_i went (out) PRO_i to buy some rice.

67) He_i wants him_j PRO to buy rice

65) and 66) have the (simplified) structures shown at 65a) and 66a) below. In both cases, the Subject NP is dominated by S which dominates the \emptyset Subject of the complement clause; the matrix Subject therefore c-commands the Subject position of the complement clause.



In 67) both the Subject and Object of the matrix verb c-command the VCOMP Subject:



The actual controller is determined lexically by the verb, along semantically predictable lines (see for example Curnow, 1992; Kroeger, 1993), some verbs being 'Object-control' verbs, and others being 'Subject-control' verbs.

Basically then, obligatory control affords a concise and accurate description of the reference of the Subject of embedded infinitives in terms of GFs and some lexically encoded disambiguating information.

3.3.2.2 Uncontrolled PRO = PRO_{arb}

Infinitives occurring in Subject position, are uncontrolled, as nothing c-commands the Subject position of the infinitive: (This example from Haegeman, 1991).

68) 'PRO To continue the investigation would be regrettable'.

Sometimes though, the reference of PRO in such structures does appear to be controlled, in the sense that it is constrained to co-reference with a single NP. However, that NP does not always c-command PRO, so control in these structures is held to be optional.

3.3.2.3 Optional Control

According to Haegeman (1991) control is optional in 69):

69) 'John thought (that) it was important PRO to behave himself / oneself.'

The idea is that the reference of the reflexive anaphor depends not directly on the matrix Subject 'John', but on the reference of PRO; PRO, in turn, may or may not depend upon the matrix Subject. If not it is uncontrolled, and its reference is then 'arbitrary'. With the right context, it may be definite, but refer beyond the sentence (this example is from Andrews, p.c.):

70) 'Mary was going to Saudi Arabia, John insisted that it would be important to behave herself.'

In this case, the controller does not c-command PRO. Other examples where the antecedent does not c-command PRO (from Haegeman, 1991:262) include:

71) 'PRO to behave myself would be my pleasure.'

72) 'It was never promised to Mary PRO to be allowed to leave.'

73) 'PRO To be allowed to leave was never promised to Mary.'
(Haegeman, 1991:262)

74) 'John tried to behave himself/*herself/ *oneself.'

In the first case the apparent controller is a possessive pronoun 'my' which cannot c-command PRO, since it is not a head; in the second and third the controller is an oblique NP which is also not head of its constituent (a PP). In these sentences there is only one possible interpretation in each case, even though the structural relationships between controller and controllee differ for each.

Thus, in cases of 'optional control', if a potential controller is present, then co-reference is not optional but obligatory. However, there is no general rule by which to identify controllers generally, and the correct controller in each case.

Haegeman (1991) acknowledges that many aspects of the co-reference constraints on PRO have yet to be adequately explained.

3.3.3 Antecedent Binding of Wh-traces

In GB the empty constituent in a relative clause is a wh-trace, produced by movement of the relativised NP to COMP. The trace is antecedent-bound by the relative pronoun. The co-reference of the wh-trace and the head is assumed to be achieved via co-indexing at the level of logical form (LF) (see Williams, 1980; Chomsky, 1982:92-3; Safir, 1986).

3.4 GB and Mandarin

3.4.1 PRO in Mandarin

In English the three reasons for proposing a PRO element are:

- 1) The existence of expletive subjects 'there' and 'it' suggests that Subjects are syntactically obligatory even when semantically unnecessary;
- 2) Subject clauses do not allow overt Subjects, but
- 3) Subject clauses may contain reflexives, which are anaphors and therefore depend upon a local antecedent. There must therefore be a \emptyset antecedent in Subject position, i.e. PRO.

The evidence for a structural position PRO is not clear-cut in Mandarin.

3.4.1.1 No Expletive Subjects

There are many idioms in Mandarin which have indefinite Subjects but do not allow overt expression of a Subject, even by an indefinite NP, for example

- 75) 总的看。
zǒngde kàn
altogether look
All-in-all...
NS7:9a

- 76) 说实话 ...
shuō shí huà
 speak true speech
 To tell the truth...
 NS2:12a

There is also a widely productive process where indefinite agents are optionally ellipsed; in these cases the patient NP may be placed in Subject position i.e. pre-verbal, but it need not be:

- 77) 盖房子了。
gài fángzi-le
 build house-ASP
 'A house has gone up.'

Though sometimes referred to as post-verbal Subjects, post-verbal NPs like that in 77) cannot control coreferent deletion and therefore do not qualify as true syntactic Subjects (Keenan, 1975).

3.4.1.2 Subject Clauses with Overt Subjects

Though Subjects are generally obligatory in English, a Subject clause can only take an overt Subject if it is (a) a possessor with a gerund or, (b) introduced by the complementiser *for* (Haegeman, 1991).

- 78a) * I introduce myself would benefit John.
 b) My introducing myself would benefit John.
 c) (For me / PRO) to introduce myself would benefit John.

Where there is no complementiser in c), the infinitive is presumed to have a PRO Subject.

In Mandarin, a Subject clause can have an overt Subject without a complementiser:

- 79) 我介绍自己对张三有利。
wǒ jiěshào zìjǐ duì Zhāngsān yǒu lì.
 1sg introduce self for Z have advantage
 a) My introducing myself would benefit Zhangsan.
 b) [For] me to introduce [my]-self would benefit Zhangsan.

The verb in this structure is clearly not a nominal, as it cannot be marked by the possessive particle *de*.

- 80) *我的介绍自己对张三有利。
 * *wǒde jiěshào zìjǐ duì Zhāngsān yǒu lì*
 1sg.POSS introduce self to Z. have profit

As overt Subjects are clearly unnecessary, and there is no evidence of special treatment to allow an overt Subject of a Subject clause, there is no direct evidence for PRO in Mandarin.

One possible explanation for the difference between the English and Mandarin sentences at 78) and 79) could be that Subjects are not governed in Mandarin, and PRO can therefore alternate with a lexical Subject. This is plausible, as Subjects in English are governed by INFL which is generally considered weak in Mandarin as it has neither of the two main components, TENSE nor AGR. James Huang (1982) suggests that this accounts for the long-range reference of Mandarin reflexives (but see Huang Yan 1994 for counter-arguments):

- 81) 约翰告诉玛理介绍自己。
Yuèhàn_i gàosu Mǎlǐ_j Ø jiěshào zìjǐ_{i/j}
 John told Mary Ø introduce self
 'John told Mary to introduce him / herself.'

An alternative possibility is that PRO can be governed in Mandarin, but this would remove the explanation for the existence of positions where a lexical Subject is not permitted (e.g. between an intransitive modal and embedded verb). This is currently explained in terms of lack of government, and an alternative explanation would be required if the licensing constraints on PRO were to be altered for Mandarin.

Yet another possibility is that the ellipsed Subject in Subject clauses represent a different kind of ellipsis from that in English or Mandarin dependent infinitive constructions, i.e. it is not PRO. This is the analysis pursued here. It is discussed in detail in Chapter Five.

3.4.1.3 Reference of Reflexives in Subject Clauses

In support of this argument, it can be observed that, unlike the ellipsed Subject in serial verb or auxiliary constructions, Ø Subjects in Subject clauses may be indefinite, or may refer to an antecedent beyond the sentence, and their reference can be affected by

changes in context. Thus they behave more like pronouns than like obligatorily ellipsed arguments. This is true for Mandarin:

82) 介绍自己对张三有利。

Ø jiěshào zìjǐ duì Zhāngsān yǒu lì.

Ø introduce self to Z have advantage

To introduce [one-/ my- him-/ etc.] self would benefit Z.

and also for English sentences like 78c). Since reflexives are anaphors, they must be referentially **dependent**, and in English their ϕ features must be derived through co-indexing with their antecedent. In cases like 78c), it is argued that the only possible antecedent is PRO, but how PRO's own reference is determined in these cases is yet to be explained. PRO in Subject position has no antecedent and no governor (without INFL) and so cannot be an anaphor. It is also of variable reference: it is not just Arbitrary PRO, but could be specific, given a discourse antecedent.

On the other hand, in declarative complement clauses (e.g. with modals), infinitive adjunct (e.g. purpose) and gerundive adjunct clauses (introduced by a preposition), PRO is 'obligatorily controlled, i.e. has an antecedent which c-commands it and has unique reference; if we allowed that it could be governed, it also has a governor (the Preposition for gerundives, the matrix Subject for infinitives). Thus it would conform to the requirements of an anaphor. So an ellipsed Subject conforms in some contexts to the requirements of an anaphor, and in others to those of a pronominal, but it need not be thought of as conforming to both simultaneously. It would seem plausible to suggest that Subjects of an infinitive in NP role, and Subjects of VCOMPs are therefore two distinct NP types rather than just one.

In other words, despite the surface resemblance to the English sentences, sentences with a VP in Subject position like 82) do not necessarily constitute sites for PRO. At the same time, they cannot be sites for little pro, the only other EC that does not need a local antecedent, because of the licensing constraints on pro. Haegeman (1991:258) states

In (38a) [essentially identical to 78c) - A.H.C.], PRO is not controlled by anything in the sentence: it may be taken as referring to a specific referent which will have been established in the context.....

and (1991:163):

'Various proposals have been formulated to deal with the data described in this section [i.e. on Subject clauses]. However at this stage no completely satisfactory control theory has been developed to cover all the complexities involved.

Even allowing for its existence, the differentiation of sites for PRO in general, and optionally controlled PRO in particular depends upon morphological cues that are entirely lacking in Mandarin. And even if Subjects of Subject clauses are PRO positions, PRO's interpretation in these cases remains to be accounted for.

Moreover, even supposing sites for pro and PRO could be correctly identified in Mandarin, neither can account for the common occurrence of ellipsis which refers beyond the sentence: PRO must refer to a c-commanding NP, or potentially some other specific but unspecifiable NP within the sentence; otherwise it has indefinite reference: PRO cannot have definite reference beyond the sentence. The 'pro-drop' parameter also means that little pro also does not refer outside the sentence. Supposing pro were allowed to refer outside the sentence, then its potential antecedents are unlimited by anything in GB theory, and its specific reference in a given context remains unexplained.

Therefore either i) some instances of ellipsis in Mandarin belong to categories other than pro and PRO, or ii) the mechanisms which determine the distribution and control of these two ECs are incorrectly stated. In any event, an explanation is still lacking for the interpretation of ellipsis which refers beyond the sentence.

3.4.2 Relative Clauses

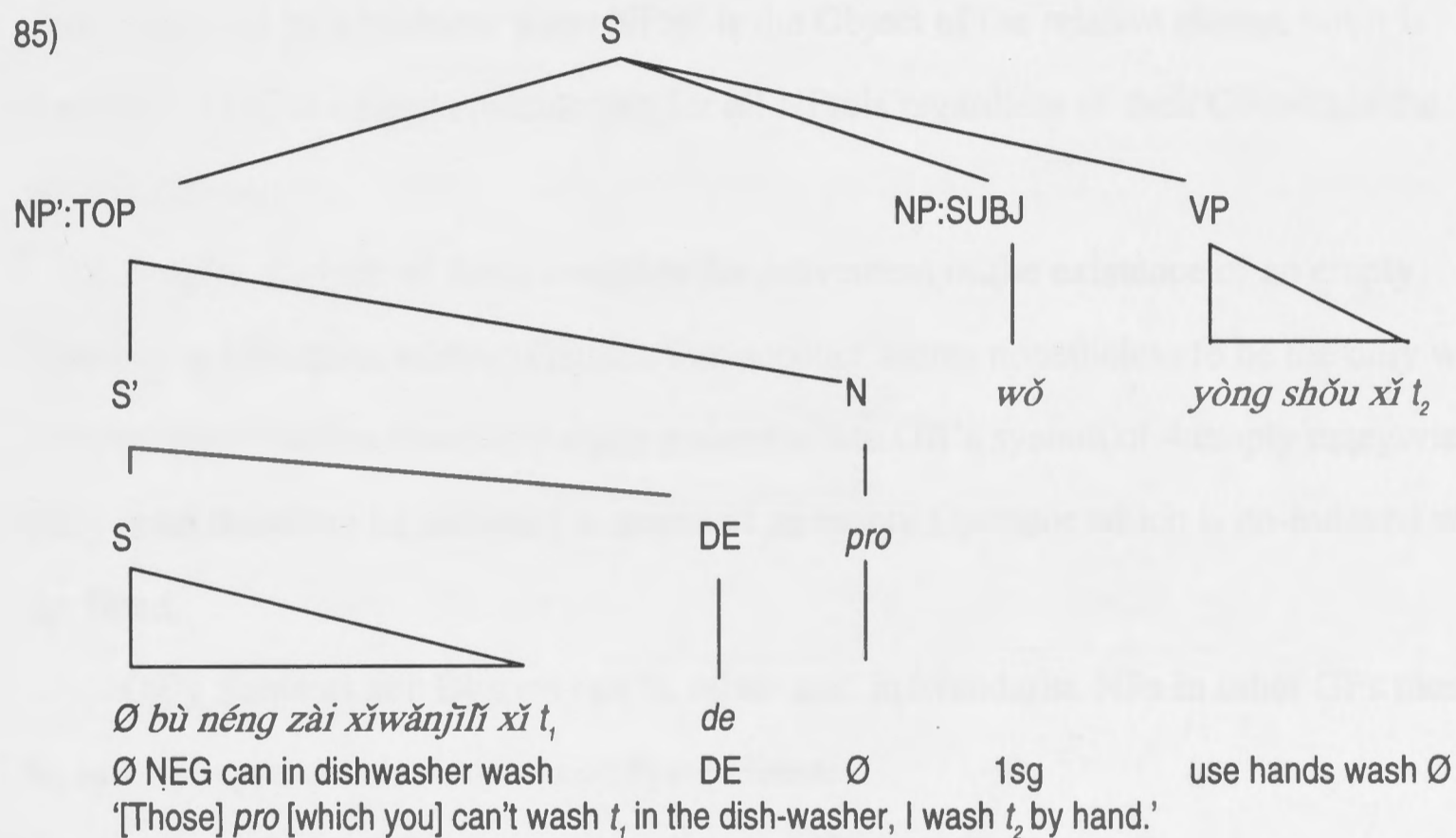
In Mandarin there is no evidence of wh-movement. Firstly, question-words appear in the same position as an NP in declarative sentences:

83) 你看见了他。
nǐ kànjiàn-le tā
2sg see-ASP 3sg
'You saw him.'

84) 你看见了谁。
nǐ kànjiàn-le shéi
2sg see-ASP who
'Who did you see?'

Secondly, Mandarin has no relative pronouns, so nothing appears in COMP position in a relative clause. A relative clause is formed by embedding a clause within the

NP in the usual position for a nominal modifier, i.e. before the head noun. The head may, of course be empty also, as in the following example from the corpus. In this example t_1 is the trace of the relativised constituent ('NPrel' hereafter), i.e. the Object of the relative clause, and t_2 is the trace of the topic clause, which includes the relative clause and the linking particle *de*. In the case of an empty head, represented here as *pro*, the reference of the head is itself determined pragmatically (see Chapter Seven).



As would be expected, since there is no extraction, indirect questions in Mandarin do not constitute *wh*-islands, so 86) is acceptable (cf. : Haegeman, 1991:370, example 69b).

- 86) 这是小王告诉我他什么时候会请的人。
zhè shì Xiǎo Wáng gàosù wǒ tā shénme shíhou huì qǐng de rén
 this is X. told me he what time would invite DE person
 'This is the person who X told me when he would invite [him].'

Since there is no violation of the *wh*-island constraint, there is no reason to propose an empty Operator in COMP position as is the case for English relative clauses with no overt relative pronoun (see Haegeman, 1991:422).

Nonetheless there is an obligatory gap within the relative clause. If this gap were a trace, either *wh*-trace or NP-trace, the only possible antecedent is the head NP, but this bears an independent theta-role, and the antecedents of traces should not have an

independent theta role. So analysis as a trace is problematic. At the same time, the gap cannot be PRO because it is governed. It is also unlikely to be pro. This is because when the relativised constituent is the Subject of the relative clause, the relative clause would not contain an accessible Subject, so the GC would be the matrix clause. As the matrix clause contains the Head NP, which c-commands and therefore binds a \emptyset Subject, the \emptyset Subject would be bound within its GC, which is not permissible for the pronominal pro. This would not be a problem when NPrel is the Object of the relative clause, but it is preferable to have a single mechanism for all NPrels regardless of their GF within the relative clause.

Despite the lack of direct evidence for movement or the existence of an empty Operator in Mandarin relative clauses, that account seems nonetheless to be the only way to fit the 'gap' within Mandarin relative clauses into GB's system of 4 empty categories. They must therefore be analysed as traces of an empty Operator which is co-indexed with the Head.

Only Subjects and Objects can be relativised in Mandarin. NPs in other GFs must be overtly represented within a modifying clause:

- 87) 我借他的书的人。
wǒ jiè tāde shū de rén
 1sg borrow his book DE person
 'The person whose book I borrowed.' (lit: the I-borrowed- his-book person)

Unlike in English, the possibility of regular definite ellipsis in Mandarin means the relative clause may have no overt NP at all. When this happens, ambiguity results (see Huang Yan, 1994:169-72 for a discussion of pragmatics in the disambiguation of relative clauses).

To recap, the trace analysis of the gaps in relative clauses is plausible, but the existence of PRO in Mandarin is uncertain, and classification of \emptyset Subjects of clauses in NP positions is problematic. None of the GB categories provide for definite reference beyond the sentence.

3.4.3 Generalised Control Theory.

3.4.3.1 \emptyset Objects as Traces

It will be recalled that *pro* is generally considered to be licensed by local identification, and therefore will not refer outside the sentence, yet Mandarin has many instances of ellipsis that do refer beyond the sentence. James Huang makes two attempts (1984, 1989) to resolve the difficulties facing a GB account of elliptic reference in Mandarin. In the first attempt, he suggests (James Huang, 1984) that *pro* in Mandarin is locally identified by co-indexing with the closest nominal element. By James Huang's definition (1984, modelled on Chomsky, 1980) the closest NP must c-command the zero element.

Ellipsed Objects are clearly an exception to this: according to James Huang, in sentences like 88), an ellipsed Object can **only** refer outside the sentence, while in 89) the ellipsed Subject can refer either to the matrix Subject, or to a retrievable referent outside the sentence ($e = \emptyset$):

88) 张三说李四不认识 \emptyset 。

a) *Zhāngsān_i shuō [Lǐsì bù rènshi e*_{i/j}]*
Zhangsan_i said Lisi does. not know [him*_{i/j}]

张三希望李四可以看见 \emptyset 。

b) *Zhāngsān_i xīwàng [Lǐsì kěyǐ kànjiàn e*_{i/j}]*
Zhangsan_i hopes Lisi can see [him*_{i/j}]

(James Huang, 1989:19d)

89) 张三说, \emptyset 不认识李四。

Zhāngsān_i shuō [e_{i/j} bù rènshi Lǐsì]
Zhangsan_i said [he_{i/j}] does. not know Lisi

张三希望 \emptyset 可以看见李四。

Zhāngsān_i xīwàng [e_{i/j}] kěyǐ kànjiàn Lǐsì]
Zhangsan_i hopes [he_{i/j}] can see Lisi

(James Huang, 1984:19c)

In 88) the complement clause contains an accessible Subject, and therefore constitutes a GC for the \emptyset Object. If the \emptyset Object were *pro*, binding theory would merely require it to be bound outside its GC, and co-reference with the matrix Subject would not contravene this. So there's no obvious reason why a *pro* Object must refer outside the sentence. James Huang (1984) therefore suggests that the \emptyset Object is not *pro* but a variable (i.e. a non-pronominal *wh*-trace).

By contrast, in 89), where the embedded Subject is absent, the GC for that \emptyset Subject is the matrix clause. James Huang (1984) claims the ellipsed Subject in 89) is *pro*, because, though not free within its GC, it is bound only by an element with an independent theta role. Even so, the \emptyset Subject is only optionally coreferent with the matrix Subject and could refer to some other NP outside the sentence. In this case it would not be controlled by the nearest nominal element and would contravene James Huang's 1984 GCR. In fact many instances of ellipsed Subjects do not refer to the nearest nominal element. Example 90) is a counter-example from Huang Yan (1994:139). The closest nominal element is underlined.

- 90) 医生说病人知道明天给他开刀。
yīshēng shuō bìngrén zhīdào \emptyset míngtiān gěi tā kāidāo
 doctor say patient know \emptyset tomorrow for 3sg operate
 'The doctor₁ says the patient₂ knows that [[I/you/he_{1/3} we.. etc.]] will operate on him tomorrow.'

James Huang claims (1984:539) that such extra-sentential reference is simply 'pragmatics over-riding syntax' in a 'non-neutral' context, where 'neutral' is taken to mean no context whatsoever, and therefore no potential competing antecedents. Since most discourse takes place in, and creates context, the relevance to interpretation, and indeed the 'neutrality' of such a contextual vacuum is highly questionable. In fact, opinion on the acceptability and reference of such a sentence is divided.

In an effort to investigate the issue further, I asked eight informants to comment on their preference for an overt or \emptyset Object in an embedded clause in constructed sentences. The test structure was presented in three contexts to determine whether a \emptyset Object in this

context was taken as obviative, as claimed by James Huang (and incidentally, disputed by Huang Yan (1994).) The contexts were:

i) where the contextually implied referent was represented both by the matrix Subject, and as a potential antecedent outside the sentence.

ii) where the contextually implied referent was not represented by the matrix Subject, but only as an antecedent outside the sentence;

iii) where there was no clear contextual implication, but two potential antecedents: one the matrix Subject, and one a topic established in the preceding sentence.

The context sentences were:

i) 张三碰到李四，总不理他。

Zhāngsān pèngdào Lǐsì, zǒng bù lí tā
Z. comes.across L, always NEG acknowledge 3sg.

'Z_i ignores L_j whenever he_i comes across him_j.'

ii) 张三碰到小王总不理他。

Zhāngsān pèngdào Xiǎo Wáng, zǒng bù lí tā
Z_i comes.across X_j always NEG acknowledge 3sgj.

'Z_i ignores X_j whenever he_i comes across him_j.'

iii) 李四听小王说他不好，

Lǐsì tīng Xiǎo Wáng shuō tā bù hǎo
L. hear X. say 3sg NEG good.

'Lis_i heard X_j say he_{i/j/x} was bad.'

and the test sentences were:

a) 所以李四认为张三不喜欢。

suǒyǐ Lǐsì rènwei Zhāngsān bù xǐhuān
'so L. reckons Z. doesn't like Ø_?'

b) 所以李四认为张三不喜欢他。

suǒyǐ Lǐsì rènwei Zhāngsān bù xǐhuān tā
'so L. reckons Z. doesn't like him_?'

The informants all agreed that an overt pronoun was better in the first context, where co-reference between the Object and a matrix Subject was assumed. Two commented that a Ø would be unacceptable in that context. However, 6 felt that an overt

pronoun was better in the second case also, where the Object was assumed to be coreferent with an external antecedent. One said that either \emptyset or an overt pronoun was acceptable, and one specified that a \emptyset was unacceptable. (There was one non-response to this sentence).

Given a context where the intended reference was deliberately ambiguous, three informants did not respond, possibly because they found it incoherent. Of the five respondents, three found a \emptyset ambiguous, one found it unacceptable, and one preferred it to refer outside the sentence (i.e. James Huang's prediction); for the overt pronoun, three found it ambiguous, one preferred it to refer to the matrix Subject, and one preferred it to refer outside the sentence (see Table 18).

Looking at individuals, one thought both \emptyset and a pronoun were ambiguous, two thought one was ambiguous and the other referred outside (but a different correspondence between form and function in each case), one thought one was ambiguous and the other locally coreferent, and the last thought one was ambiguous and the other bad, see Table 19.

Table 18 Judgements on \emptyset SCOMP Objects:

	Ambiguous	Local	External	Bad
\emptyset	3		1	1
Pron.	3	1	1	

Table 19 Individual Systems of Reference for \emptyset Objects:

	1	2	3	4	5
\emptyset	ambig	ambig	external	ambig	bad
Pron	ambig	external	ambig	local	ambig

So on the individual level there seems to be something approaching a functional division in three cases, but no consistency across informants. These results may suffer from the difficulty of devising realistic contexts and further probing may help define a precise context where an obviative effect does occur with a \emptyset Object.

One clear indication though, is that the use of a \emptyset in this position is actually dispreferred in all the contexts tested.

Supposing for argument's sake that \emptyset Objects in embedded clauses do occur in natural discourse, then, if they are traces, as James Huang (1984) suggests, they must be

locally bound by an antecedent. James Huang suggests the \emptyset Object in 88) is locally bound by a topic (a non-argument), which happens to be empty. This argument rests on the observation that ellipsed Objects in topic-comment sentences are constrained to co-reference with a lexical, sentence-initial topic NP.

However, there are differences in the reference of ellipsed Objects in sentences with lexical topics and those with so-called empty topics. The reference of a \emptyset Object in a sentence with an overt topic is locally bound by that topic, and this co-reference cannot be affected by extra-sentential context; the reference of a \emptyset Object in a sentence with no lexical topic however, is not uniquely constrained and is affected by context outside the sentence. The following illustrates the point:

91a) 这些书, 谁看过?

Speaker A: *Zhèxiē shū_i, shéi kàn-guò e_i?*
 this-few book_i, who see-ASP \emptyset_i ?
 'These books_i, who's read [them_i]?'

我没看过。孩子看过。

Speaker B: *Wǒ méi kàn-guò e_i. Háizi kànguò e_i.*
 1sg NEG see-ASP \emptyset_i . child see \emptyset_i .
 'I didn't read [them_i]. The child read [them_i].'

91b) 这些书, 谁偷了?

Speaker A: *[Zhè-xiē shū, shéi tōule e_i]?*
 this-few books, who stole them?
 'These books, who stole them?'

我没看过。孩子看过。

Speaker B: *Wǒ méi kànguò e_i. Háizi kànguò e_j.*
 1sg NEG see-ASP \emptyset_i child see-ASP \emptyset_i
 'I didn't see [who stole them]. The child saw [who stole them].'

In both utterances by Speaker A, the ellipsed post-verbal NP is constrained to co-reference with the sentence initial lexical topic. Speaker B's utterance is identical in the two passages, but in 91a) the ellipsed element is construed as an NP, coreferent with the ellipsed Object in the previous speaker's turn (which is in turn coreferent with the lexical topic NP in that sentence), while in 91b), the empty category is construed not as an Object NP at all, but as a sentential complement coreferent with the whole of Speaker A's statement. Thus, unlike the ellipsed Object in speaker A's utterance, the reference of

the ellipsis in speaker B's utterance is structurally ambiguous, and its interpretation depends upon the external context. This throws doubt on the comparability of sentences with and without overt topics.

It could be argued that the externally established discourse topic controls the reference of an empty sentence topic which in turn controls the reference of the ellipsed Object. This, however merely transposes the essential problem to another locus: there would still be an instance of an empty category (topic) whose reference is somehow determined from outside the sentence and is therefore not locally bound. Being locally free, an empty topic must be a pronominal, like the Subject NP in 2), but James Huang's (1984) suggestion of local identification by co-reference with the closest NP effectively rules out the possibility of reference for pro outside the sentence, since there can be no c-command across sentence boundaries.

Huang Yan (1994:46) makes a similar criticism of James Huang's analysis of the ellipsed Object as a 'variable', saying:

the proposal that the empty topic can occur in the absence of its chain initial topic has the immediate consequence of forcing us to analyse every sentence in Chinese as having an empty topic, and more absurdly, even to analyse every sentence in Chinese as containing an indefinite number of empty topics...The question that comes up next is how is it locally A'-bound? The answer is likely to be that the empty topic (itself being a variable) is locally A'-bound by another empty topic, ad infinitum. Thus we are led to the implausible position that a sentence in Chinese contains an indefinite number of empty topics. In other words, positing empty topics will result in an infinite regress.

Though Huang Yan's conclusion may be a little strongly stated, the question is an important one: if every sentence has an empty topic, what kind of empty category is it, and how is its reference determined?

In fact, we shall see in Chapter Four, that most \emptyset Objects are optionally ellipsed and have the same reference as an overt pronoun; those in sentences with an overt topic may be bound to that topic, but may also refer to an extra-sentential discourse topic, so long as the sentence-internal topic can control some other semantic role in the sentence.

The sensible resolution seems to be to propose that a \emptyset Object is pragmatically bound: that is it must refer to an element of some specific pragmatic status. Sometimes this pragmatic status is held by a sentence topic, but sometimes not and the obviative

effect observed in sentences like 88) is probably a result of the higher relative topicality of an entity just established in focal position, than the Subject of a speech act verb.

This idea will be developed in Chapter Seven.

3.4.3.2 A Unified pro/PRO

In a later analysis, while maintaining his position on the ellipsed Object as a variable, Huang (1989) suggests an alternative Generalised Control Rule:

Generalised Control Rule 1989 version (C.T. J. Huang, 1989:204)

92) "An empty Pronominal (i.e. pro/PRO) is controlled in its control domain (if it has one)."

This relaxes the requirement for local identification by allowing that some \emptyset elements have no control domain.

Huang (1989) suggests that there is no distinction between PRO and pro in Mandarin, because like PRO, pro occurs only in Subject position (a consequence of his own treatment of ellipsed Objects as bound variables, not pro); must be 'controlled' when it's the Subject of embedded finite clauses (though there is no distinction here between the effects of binding and the effects of 'control') and 'under certain circumstances a Chinese pro may also be free'.

James Huang (1989) compares the ellipsed Subject in 89) above and also in

93) 张三说 \emptyset 很喜欢李四。

Zhāngsān shuō pro hěn xǐhuān Lǐsì

Z say pro very like L.

'Zhangsan₁ says $\emptyset_{1/2}$ likes Lisi₃'

to the 'free' reference of PRO

94) 吸烟有害。

PRO xīyān yǒu hài

PRO smoke have harm

'Smoking is harmful.'

Note however, that this is not free reference in the sense of being able to refer to any contextually retrievable entity, it is **arbitrary** reference, in the sense of obligatorily unspecified, a reading which is not possible for 93). In other words there is a distinction between positions where \emptyset s can only have definite reference and those which can only be

indefinite. The problem with this analysis is that it still fails to account for the ability of empty categories to optionally refer outside the sentence, even when the sentence apparently constitutes a control domain, as in examples 89) and 93) above.

James Huang (1989) suggests (after Rosenbaum, 1967) that this is because say-verbs are subcategorised for an NP rather than directly for S' and so have no control domain for *pro*. But Huang Yan (1992) demonstrates that the evidence for the distinction between say-verbs and other verbs does not hold true for Mandarin. In short then, neither of James Huang's (1984, 1989) attempts to 'save' GB resolve the basic problems of definite reference beyond the sentence, for either \emptyset Subjects or \emptyset Objects.

Battistella (1985) proposes that PRO is governed in Mandarin, but the GC for anaphors is different from that for pronominals. This idea is challenged by Huang Yan (1992) on three counts, firstly a governed PRO would have too wide a distribution; secondly, in 95) PRO could still refer outside the sentence altogether, in which case it is not bound in its anaphoric GC.

- 95) 老王说 [\emptyset 去过美国]。
 [GC_{anaph} Lǎo Wáng_i shuō [GC_{pron} PRO_{i/j} qù-guò Měiguó]]
 Old Wang say PRO go-asp America
 'Old Wang says I/you/he_{i/j}/ they etc. have/has been to America.'

Thirdly, there are elliptic sentences where the \emptyset position is clearly not governed, as it allows no lexical pronoun:

- 96) 老王请小李来。
 Lǎo Wáng qǐng Xiǎo Lǐ \emptyset lái
 Lao Wang invite Xiao Li \emptyset come
 'Lao Wang invites Xiao Li to come'

In Battistella's (1985) schema there is no *pro*, so the identity of this \emptyset position would be problematic.

Huang Yan (1992) argues that neither *pro* nor PRO can exist in Mandarin. PRO cannot exist because there is no systematic way to distinguish finite from non-finite verbs in Mandarin (but see Li Y.A. (1990) for an opposing view). Huang Yan (1992) concludes that neither the variable, nor the *pro*/PRO analysis of Object ECs in Mandarin is tenable. According to him, zero anaphors in Chinese are realised by 'syntactically undifferentiated

gaps rather than by specific empty syntactic categories, and are not grammatically but pragmatically determined.' (Huang Yan, 1992:27)

3.5 Summary of GB

Referential effect

GB identifies four empty categories. However there are inconsistencies of referential effect for PRO, and some overlap of referential effect for PRO and pro. When the reference of \emptyset Subjects is considered, for both Mandarin and English there is variability between Subjects of VPs in Subject position and those of dependent clauses. This could be readily accommodated, and the apparent disparities between English and Mandarin resolved, if the former were not classified as PRO.

More importantly, there are instances of ellipsis in Mandarin which do not fit any of GB's four categories because they refer beyond the sentence.

Distribution

The identification of PRO, and therefore its distribution, does not constitute a major problem in English, or other languages with verbal morphology, because sites for PRO can be effectively defined as the Subject position of a non-finite verb. In Mandarin however, there is no morphological distinction based on finiteness, and lexical NPs can appear in structural positions that look on the surface to be analogous to ungoverned positions, and thus sites for PRO in English constructions. This then creates a problem for identifying sites for PRO in Mandarin discourse and differentiating them from sites for pro. Mandarin also does not comply with the pro-drop parameter, as there is no local identification of ellipsis via verbal agreement.

In terms of a distributional classification, the GB analysis therefore fails to provide the means for accurate identification of EC type on the basis of structural distribution in Mandarin.

Interpretation

Because of the difficulty in identifying empty categories correctly, a GB analysis fails also to explain how a listener/reader can know which interpretative process (anaphoric or pronominal binding or control) to apply in any given case of ellipsis.

Even when identification is possible, the reference of PRO is acknowledged to be quite variable and therefore problematic, and there is no mechanism for the selection of a single antecedent for pro from potentially many that c-command it outside its GC.

In James Huang's (1984, 1989) Generalised Control Theory \emptyset Subjects belong to a single empty category which must be controlled within its control domain, if it has one. Control is by the structurally 'nearest' NP. This works for only some instances of \emptyset Subjects, and cannot account for reference beyond the sentence, because Huang's characterisation of 'nearness' involves syntactic relationships.

\emptyset Objects in this account are traces controlled by topics, which may themselves be null. This is effective where the topic is lexically realised as a sentence topic, but the null-topic hypothesis merely results in a transfer of the essential problem: How is the null topic interpreted? The interpretation of \emptyset Subjects which have no control domain also remains problematic.

4. LEXICAL FUNCTIONAL GRAMMAR (LFG)

4.1 Referential Categories

Categories of ellipsis in LFG are derived from different mechanisms of control, not from observations of referential effect. The reference of covert elements is constrained at the level of f-structure either via a 'rule of interpretation' called functional control, or via anaphoric control (Bresnan, 1982). The former constitutes a kind of direct one-to-one mapping of the bearer of a GF to some other NP in the discourse, which is identified via the argument structures of constituents. It identifies the reference of GFs which are not, and cannot be represented in overt form (Horrocks 1987:254). Since each argument is uniquely represented in f-structure, functional control will produce a unique interpretation of reference.

Anaphoric control is a rule which 'responds' directly to general syntactic structure, as represented by f-structure. It is a response to the occurrence of a verb which assigns a core GF where there is no overt form to bear it. It introduces a null pronominal PRED 'pro' into f-structure. Anaphoric control is implicated in structures where overt realisation would be possible (Horrocks, 1987:255) and operates with respect to unrealised GFs which are not under functional control. Its interpretation is constrained only by the requirement that antecedents must f-command PRED 'pro' (see below). This means that anaphoric control allows for potentially many antecedents for a single ellipsed argument.

LFG then can be seen as a system that predicts the occurrence of ellipsis of both uniquely constrained and variable reference. However, for reasons that will be discussed later (see Table 20 and discussion pg.98), there is not a perfect correlation between control mechanism and referential effect in LFG as currently formulated.

4.2 Distribution

In terms of distribution, LFG essentially allows ellipsis of any argument, and imposes no requirement for a specific type of empty item in a specific structural position.

Horrocks (1987, pg. 239) gives a concise overview of the LFG position:

In LFG predicate arguments are listed independently of grammatical function assignment; this raises the possibility that certain arguments will have no surface grammatical interpretation and that certain surface grammatical functions will have no thematic role. In particular, if there is no *overt* category in surface structures that can be identified with a particular θ -role, there is no need to assume the presence of an empty one. Thus agentless passives, for example, can be represented syntactically as precisely that, because the presence of an (understood) agent follows from the lexical form.

In other words, ellipsis at the morphological level may reflect the absence of a constituent from c-structure, rather than an unfilled position. The choice between functionally controlled ellipsis, and anaphorically controlled ellipsis, is deemed to be determined at the lexical level. The sub-categorisation frame of each lexical item may include information about the functional control, and therefore interpretation, of a given GF in a given syntactic frame. Since Functional control is exercised by specific lexical items and is a facet of lexical knowledge, its distribution in discourse is a function of lexical choice.

4.2.1 PRED 'pro'

If there is no specification of functional control within a sub-categorisation frame, the anaphor PRED 'pro' is automatically introduced into f-structure. Anaphorically controlled ellipsis thus constitutes a kind of ever-present default value.

These mechanisms seem to offer little in the way of constraints upon the occurrence of ellipsis, since theoretically any GF can either be unassigned and its reference determined by an interpretation rule, or be represented by the introduction of PRED 'pro' into f-structure. This is an advantage in accounts of languages like Mandarin, with regular definite ellipsis, on the other hand, it means that other languages need to be constrained in some way. In fact, functional control operates in a limited set of syntactic structures only (see 4.3.2 below). Recall that, empirically speaking, obligatory ellipsis is similarly restricted suggesting a connection between obligatory ellipsis and functional control.

4.3 Interpretation

In all cases where a 'morphologically unexpressed pronominal' is assigned an antecedent, whether obligatorily as in Functional control, or optionally as in Anaphoric control, the assignment of reference must comply with the requirement of f-command (Horrocks, 1987: 240). Bresnan (1982) credits this requirement to Mohanan (1981) and states it as a 'Universal Condition on Anaphoric Control':

'If A is a grammatically assigned antecedent of P, where the value of P is [PRED 'PRO', U+] then A must f-command P.' (Bresnan, 1982:333).

Where 'U+' means morphologically unexpressed.

4.3.1 F-command

'For any occurrences of the functions α , β in an f-structure F, α f-commands β if and only if α does not contain β and every f-structure of F that contains α contains β .' (Bresnan 1982: 334).

The requirement that every f-structure that contains the antecedent α , must also contain the \emptyset element β , constrains the reference of β to an antecedent in β 's own

sentence: since there is no f-structural unit larger than the sentence, elements can only f-command other elements within the same sentence. This means that LFG, like GB, provides no account of ellipsed NPs whose only antecedents are outside the sentence.

In fact, LFG appears to equate the lack of an antecedent within the sentence to indefinite or arbitrary reference: Horrocks mentions (1987:254) that anaphoric control “also allows for cases where there is **no overt controller within the sentence** and the reference of the controllee is **therefore arbitrary**.” (emphasis added). These two points foreshadow LFG’s inability to resolve the twin problems of arbitrary reference even in the presence of a potential controller, and definite reference to an antecedent outside the sentence.

Functionally controlled elements are subject to a further ‘default’ rule of interpretation. For functionally controlled GFs, a specific antecedent in the matrix clause is identified by way of the antecedent’s GF in the order OBJ2> OBJ> SUBJ. (Note that this is virtually the reverse of Huang Yan’s (1994) ASP, but it is applicable to a more restricted set of cases).

4.3.2 Functional Control

The specific syntactic frames in which Functional control operates are verbal complement structures where the matrix verb may be intransitive, or transitive, but cannot take an oblique complement or adjunct such as a *for* complement (Horrocks, 1987: 254-268) . Compare:

- 97) Leon persuaded his colleague to think again
* Leon₁ persuaded his colleague₂ for him_{1/2} to think again.
(After Horrocks 1987:255)
- 98) Leon hesitated to take the matter into his own hands.
* Leon₁ hesitated for him_{1/2} to take the matter into his own hands.

The formal statement of the Functional control rule is as follows:

‘If V-COMP is assigned as a G(rammatical) F(unction) in a lexical form L, then (GF-_θ) = (V-COMP SUBJ).

This rule identifies the non-thematic grammatical function of a lexical form with the Subject of the V-COMP.' (Horrocks 1987: 240). This special kind of VCOMP which cannot take an overt Subject is called an 'XCOMP'.

Verbs which produce a different pattern, e.g. 'promise' where the Subject controls the dependent Subject function even in the presence of an Object NP are simply assumed to have a specific lexical entry which blocks the application of this rule.

4.3.3 Anaphoric Control

There is much less restriction on the operation of Anaphoric control. Horrocks (1987:262) gives the rule of Anaphoric control:

[47] For any verb and any semantically unrestricted GF which it governs, assign optionally the following equations to the verb's lexical entry:

- (i) (\uparrow GF PRED) = 'PRO' (For English GF = {SUBJ})
- (ii) (\uparrow FIN) =_c α (For English α = —)

and states further:

Given [47] any infinitive that may be introduced by *for*, and so have a lexical Subject, (this excludes all cases that fall under the heading of functional control, where *for* is obligatorily absent), and any gerund, all of which may have lexical Subjects, may also appear without a syntactic Subject and have the PREDS of their SUBJ functions interpreted as 'PRO' by 47.

Note that it is the possessive, in the case of the gerund, and the *for*-Object in the case of the infinitive, which are viewed as 'overt Subjects' in Horrocks' examples:

- 99a) Leon decided that for him to sell off the family silver was his only chance
- b) For her to involve herself was risky for Maggie
(Horrocks, 1987:255).

Structurally speaking, these examples involve gerunds and infinitives in NP positions.

As for complement positions, any infinitive can be introduced by *for*; the occurrence of *for* is dependent on the matrix verb. So it is verbs that can **introduce** a *for*-complement which are considered not to impose functional control. We might then expect that the reference of \emptyset Subjects in the VCOMPs of these verbs would be fairly free. However, **antecedents** are restricted by the Obviation Principle. This states that

if P is the pronominal SUBJ of an obviative clause (in English this means an infinitival clause that may be introduced by *for*), and A is a potential antecedent of P and the SUBJ of the clause immediately containing the obviative clause, P is bound to A if P is not morphologically expressed, and P is not bound to A if P is morphologically expressed. (Horrocks 1987:262)

In other words, in English, the sites where anaphoric control is held to determine the reference of a \emptyset Subject of a dependent verb should be precisely those sites where the dependent \emptyset must refer to a matrix Subject. The obviation principle thus restricts the reference of the \emptyset Subject to precisely that which would obtain under functional control anyway. This seems unnecessarily complicated, but it is apparently considered necessary to account for the ability of some VCOMPs to take 'overt Subjects' when others can't. It is, however, not entirely accurate. Consider:

- 100) Mike signalled (for him) to quit.
(Horrocks, 1987:255).

In this example if the *for*-comp in brackets were ellipsed, the 'morphologically empty pronominal' would not generally be interpreted as coreferent with the matrix Subject.

Moreover, the disjoint reference predicted by the Obviation Principle in English appears to be a question of semantic roles, rather than syntactic functions: *for*-Objects that co-refer with the matrix Subjects are only dispreferred, **when they introduce an active clause:**

- | | | | |
|------|------------------------------------|-----|-----------------------|
| 101) | I asked for me to be allowed to go | cf. | *I asked for me to go |
| 102) | ?I hope for me to be allowed to go | cf. | *I hope for me to go |
| 103) | I like for me to be allowed to go | cf. | ?I like for me to go |

The role of semantics in control relations will be important in the discussion below (pg. 120) of Mandarin verb types, where morphological cues are absent.

Basically then LFG has one control mechanism whose **potential occurrence** is restricted to one or two syntactic frames, and whose actual presence is lexically encoded, and a complementary control mechanism with a much wider distribution, whose **interpretation** is restricted in certain syntactic frames.

4.3.3.1 Anaphoric Control and Object Ellipsis

Objects are explicitly excluded from anaphoric control in English (where Object ellipsis is not standard). It is unclear whether they are also excluded in Mandarin by the phrase in the Anaphoric-control rule, 'for any **semantically unrestricted** GF' which a verb assigns. Object GFs are sometimes characterised as more semantically restricted (by

verbal semantics) than Subjects are. But, in any event, there is no other mechanism in LFG to ascribe reference to ellipsed Objects.

Supposing anaphoric control to be the relevant process, its flaws with regard to \emptyset Objects are essentially the same as those for Subjects: they often refer outside the sentence, and though they often have ambiguous reference, which fits nicely with the fact that Obviation does not apply to Objects, they sometimes have a unique reference which cannot be accounted for by the loose constraints of f-command (see for example, the \emptyset Objects in 91) above).

4.3.4 Constituent Control

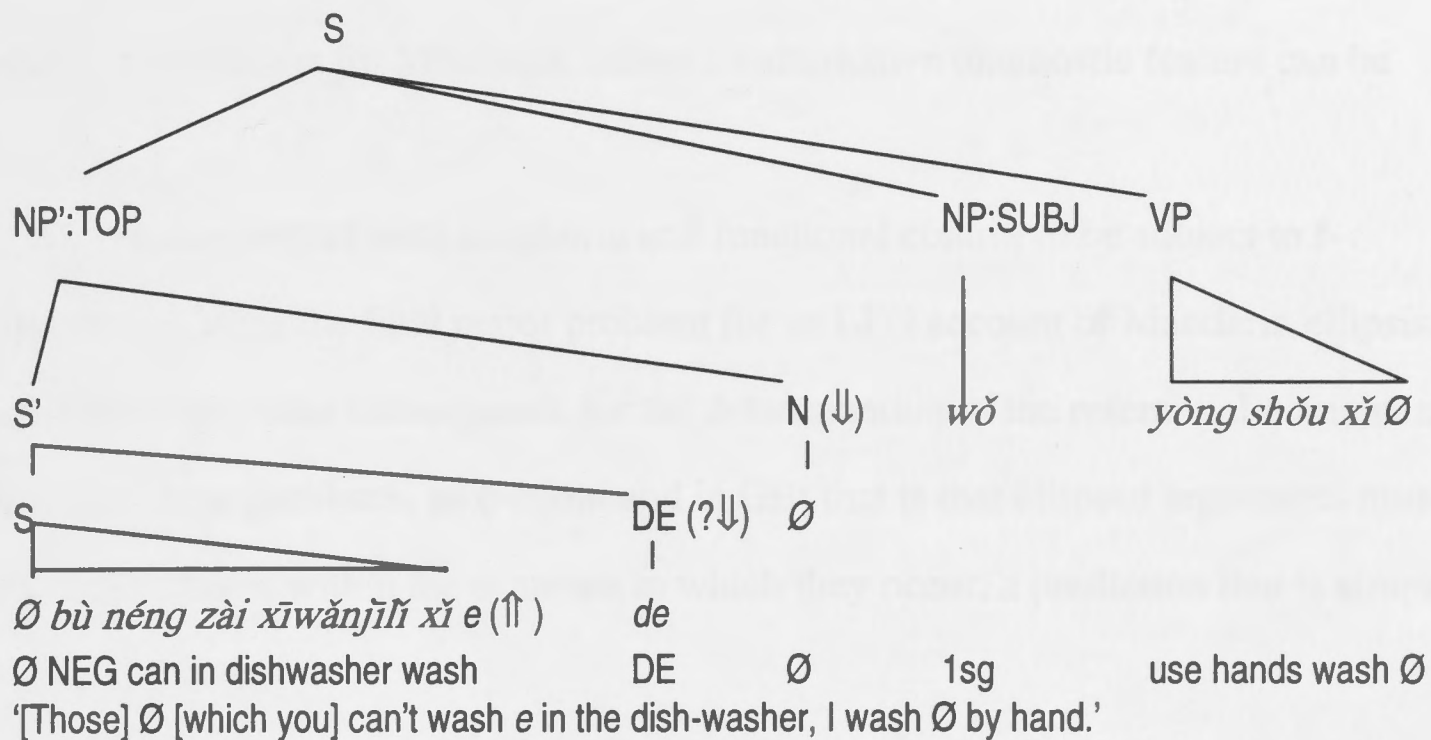
In addition to anaphoric and functional control, LFG has a process called constituent control involved in the interpretation of the empty position (or gap) in relative clauses.

4.3.4.1 Relativised Constituents

Constituent control operates through 'bounded domination variables' which are representations at the level of c-structure. These are meta-variables which 'are instantiated with actual variables when the f-description is formed.' (Kaplan and Bresnan, 1982:234). The 'gap' within the relative clause and the controller (the relative pronoun, or in its absence, the head (see for example Pinker, 1982:663), bear matched meta-variables and are therefore parsed as coreferent. Reference is constrained because of constraints on the structural relationships between controller and controllee, as well as specifications of constituent type 'marked' on the meta-variables. (For a detailed discussion see Kaplan and Bresnan, 1982, and Pinker, 1982). In other words, there is assumed to be some kind of cross-referencing or co-indexing during construction of the sentence.

Since there is no relative pronoun in Mandarin, the gap in a relative clause must be assumed to be linked to the head, (or possibly to the Node dominating the particle DE):

104)



In essence both the GB analysis (see pg. 76 above) and that of LFG reflect the intuition that there is a recognisable gap within a relative clause, that its GF can be deduced from the linear position of the gap, and that it will be coreferent with the head or relative pronoun. Mandarin, being without relativisers and morphological agreement, provides no basis on which to reject or confirm either analysis. At the same time, there is no reason to suppose that the mechanisms by which a relativised constituent is linked to its Head in Mandarin, whatever these may prove to be, are any different from those in any other language.

From the viewpoint of acquisition and interpretation, relative clauses in Mandarin are readily recognisable as modifiers of a head noun since they occur in the same syntactic position as adjectives and take the same nominal linker DE.

4.4 Summary of LFG:

LFG identifies three mechanisms by which \emptyset elements are interpreted: constituent control, functional control and anaphoric control. For English, functional and anaphoric control are differentiated on the basis of the possibility of the occurrence of a *for*-complement to express the agentive role of a verb. In VCOMP position, the reference of a \emptyset Subject under anaphoric control is further restricted by the Obviation Principle, to co-reference with the matrix Subject. Both anaphoric and functional control are subject to f-command.

The differentiation of these two control mechanisms via morphological cues is potentially problematic for Mandarin, unless an alternative diagnostic feature can be found.

The requirement of both anaphoric and functional control to be subject to f-command constitutes the final major problem for an LFG account of Mandarin ellipsis. It has essentially the same consequence for the determination of the referential domain, and the same resulting problems, as c-command in GB; that is that ellipsed arguments must find their antecedent within the sentence in which they occur, a prediction that is simply untrue for Mandarin.

In fact this requirement is redundant for functional control, as all functional controllers must be arguments of a matrix verb, and therefore will, of necessity, be within the sentence whose f-structure contains the f-structure of the \emptyset element. But some alteration to the requirement of f-command for anaphoric control is necessary before it can account for both the rigidity in some cases (e.g. forced external reference of \emptyset Objects), and the freedom and flexibility in other cases (e.g. optional external reference of \emptyset SCOMP Subjects) of elliptic reference in Mandarin.

5. COMPARISON OF APPROACHES

5.1 Referential Patterns

Shopen (1973) distinguishes between definite and indefinite ellipsis, and this corresponds in part to the distinction between pro and PRO in GB: small pro has definite reference, while arbitrary PRO is indefinite. However, the PRO involved in 'equi deletion' as in (60) above, has definite reference. This level of distinction between an unspecified bearer of some semantic role, and a constrained reference to the bearer of a semantic role is one that the GB categories fail to capture.

GB has 4 empty categories, each with supposedly different referential properties and complementary distribution, but for certain instances of ellipsis in Mandarin, it is often difficult to distinguish to which category, if any, it belongs. GB's category PRO,

produces both unique and contextually variable reference, suggesting that contextually, if it is a valid category at all, it is not a homogeneous one.

5.2 Distribution

Shopen sees obligatory and optional ellipsis as a consequence of lexical variation; distribution is thus a consequence of lexical choice in a given discourse context. GB analyses distribution in syntactic terms, but, unfortunately its claims for licensing of PRO and pro simply do not hold for Mandarin. In LFG ellipsis is scarcely constrained at all, allowing virtually any core argument to be ellipsed. The fact that obliques are not generally ellipsed 'falls out' neatly from the LFG mechanisms of control, since they depend upon the non-assignment of **core** GFs to recognise ellipsis and 'trigger' the interpretative mechanisms. This seems closest to the empirical evidence of languages like Mandarin.

5.3 Interpretation and Mechanisms of Control

5.3.1 What Control Mechanism in what Structure?

We have seen that in LFG, certain control mechanisms correlate with certain types of dependent structures. In GB, control mechanisms are associated with thematic role assignment. Because of this, the two theories produce different groupings of syntactic structures (see Table 20 below).

In English, GB, verbs like 'seem' and 'believe' do not assign a semantic role to one or other of their arguments. This can be seen from their ability to take an expletive argument 'it' or 'there'. The status of 'there' as a syntactic Object in 105 ii) is clear because it can appear as Subject in a passive sentence 105 iii) e.g.:

- 105) i) It seems there is a good reason for this.
ii) I believe there to be a good reason for this.
iii) There is believed to be a good reason for this.

In GB, when these verbs take a verbal complement, the Subject of the dependent verb is thought to be optionally 'raised' to the position of a non-thematic argument in the

matrix clause. This leaves the dependent verb with no apparent argument (see examples in Table 20).

Table 20. Control Mechanisms in GB and LFG

Syntactic Frame	GB	LFG	Example
Complement structures: 'Raised' Subjects	Trace Binding	F-control	'Chomsky seems to like LFG'
other	PRO (Obligatory control)		'It seems Chomsky likes LFG'
∅ Subj of subord / adjunct clause	PRO (Optional control)	A-control	'He wants (Bresnan) to use GB'
∅ Subj of Subject Nominal Predicate			'Abandoning GB, he transformed.' 'He signalled for Henry to quit'
∅ Subj co-ord clauses			pro

The structural 'gaps' left by 'raising' are considered (in GB) to be 'traces' and their reference is determined by binding, and must be antecedent-governed, whereas the Subject of other dependent verbs is thought to be represented by PRO, whose reference is determined by control.

In LFG, 'raising' constructions are said to be under functional control. Thus LFG's functional control, corresponds in part to the binding of an NP-trace in GB, and in part to Obligatory control. At the same time, Anaphoric control corresponds in part to GB's optional control and in part to pro. This distribution of control mechanisms across syntactic types is shown in Table 20.

5.4 Raising in Mandarin

Raising in Mandarin appears to be less common than in English. Firstly, Mandarin has no overt expletive element: if a verb assigns no semantic role to a GF, the GF is simply unrepresented at the phonological level at least. Semantic counterparts for the English verb *seem* include *xiǎnde* 'to manifest-RESULT.particle', *hǎoxiàng* a compound of the adverb *hǎo* 'much' and the verb *xiàng* 'resemble' and *kànlái* literally 'look-come'. The first of these, *xiǎnde* is subcategorised for a Subject and a stative-verb complement only, with no alternative valencies. It clearly assigns a semantic role to its Subject as it cannot take an arbitrary or non-referential null Subject, the Subject is ellipsed only if

retrievable. Nor does it allow its sole NP argument to appear in post-verbal position as Subject of a dependent VCOMP as 'seems' does in English :

- 106) 'He seems happy.' 'It seems he's happy.'
 他显得高兴。 *显得他很高兴。
 Tā xiǎnde gāoxìng **xiǎnde tā (hěn) gāoxìng*

Hǎoxiàng and *kànlái* can occur in sentence initial, or in second position:

- 107) a) 他看来高兴。
 Tā kànlái gāoxìng
 3sg look-come happy
 'He looks happy'

- b) 看来, 他高兴。
 Kànlái, tā gāoxìng
 look-come 3sg happy
 'By the look of things, he's happy.'

- 108) a) 他好象高兴。
 Tā hǎoxiàng gāoxìng
 3sg seem happy
 'He seems to be happy'

- b) 好象, 他高兴。
 Hǎoxiàng, tā gāoxìng
 Seem 3sg happy
 'Apparently, he's happy.'

but they are not verbs: they cannot themselves be negated in the usual way for verbs:

- 109) a) 他(*不)看来高兴。
 *tā (*bù) kànlái gāoxìng*
 3sg (neg) look happy

- b) 他(*不)好象很高兴。
 *tā (*bù) hǎoxiàng (hěn) gāoxìng*
 3sg (neg) seem (very) happy

Only the predication that they introduce can be negated:

- 110) a) 他看来不高兴。
 tā kànlái bù gāoxìng
 he look (neg) happy
 'It looks like he's unhappy.'

- b) 他好象不很高兴。
tā hǎoxiàng bù (hěn) gāoxìng
 he seem neg (very) happy
 'It seems he's not (very) happy.'

In fact the structural positions available to *kànlái* and *hǎoxiàng* are precisely those available to time and space adjuncts, cf.:

- 111) a) 他昨天高兴。
Tā zuótiān gāoxìng
 3sg yesterday happy
 'He was happy yesterday.'

- b) 昨天他高兴。
Zuótiān tā gāoxìng
 Yesterday 3sg happy.
 'Yesterday, he was happy.'

In second position, *kànlái* and *hǎoxiàng* can be pronounced with parenthetical intonation, unlike *xiǎnde* and matrix verbs like *kěyǐ* 'to be able to' or *xǐhuān* 'to like to'. The change of position in the sentence is associated simply with a change of scope and a shift in the topicality of the Subject NP: when it is sentence-initial, the Subject NP is being highlighted as an entity about which a predication is being made, when it follows the adjunct, the entity referred to by the Subject NP is being presented simply as a necessary, but not central or newsworthy participant in a 'predicate focus' structure (Wu, 1992).

Since *hǎoxiàng* and *kànlái* are not verbs, there can be no question that the sentence initial NPs in (107) and (108) are raised Subjects, they are simply the Subject of the main verb. As to verbs which assign no semantic role to their surface Object in a complement structure, where raising to Object might occur, there is no evidence that verbs of knowing or belief in Mandarin allow the Subject of their complement to be raised to a matrix Object. Firstly, prosody suggests that an NP following such a verb in Mandarin belongs to the dependent, not the matrix clause:

- 112) 我相信，他不会去。
 a) *wǒ xiāngxìn, tā bù huì qù*
 1sg believe 3sg NEG AUX go
 'I believe he won't go.'

- ?我相信他, [Ø 不会去] 。
- b) ?*wǒ xiāngxìn tā, [(Ø) bù huì qù]*
 1sg believe 3sg (Ø) NEG AUX go

Only if this surface form were interpreted as a sentential complement with a Ø Subject would it be acceptable, with the interpretation: 'I believe him, [I/ he/she/you etc.] won't go'.

Secondly, consistent with the claim that the postverbal NP is not an Object, the Mandarin verbs cannot be passivised. Passivisation can be achieved in Mandarin by *bèi* (adversative) and *yóu* (neutral) (see Tan, 1991) constructions. Neither of these strategies is acceptable for speech-act or epistemic verbs in Mandarin:

- 113) 他们认为, John 不对。
- tāmen rènwéi John bù duì*
 3pl think J. NEG right
 'They think J. is wrong.'

- 114) * John 被 / 由他们认为不对。
- * *John {bèi/yóu tāmen} rènwéi bù duì*
 * J {by/by 3pl} think NEG right

However, the presentative *yǒu* is a likely candidate for a raising verb, since the semantic agent of the dependent clause can appear as Subject of the matrix clause:

- 115 a) 我有很多看不懂。
- wǒ yǒu hěny duō kànbùdǒng*
 1sg have very much see.NEG.understand
 'There was a lot I didn't understand.'

- b) 有很多我看不懂。
- yǒu hěny duō wǒ kànbùdǒng*
 have very much 1sg see.NEG.understand
 'There was a lot I didn't understand.'
- NS5:11

However, this verb can also assign a semantic role to its Subject. When the Subject is animate the verb is interpreted as a possessive 'have' but the verb *yǒu* with a locative Subject is also a standard way to express a person's place of residence, or the time or place of an event's occurrence. Such a usage occurs in the corpus and is shown at 116).

- 116) 山东有个农民老人，
Shāndōng yǒu ge nóngmín lǎorén
 Shandong have CL country old-man,

用自己节省了十几年的钱买了一台电视机。

yòng zìjǐ jiéshěng-le shíjǐ nián de qián mǎi-le yītái diànshìjī

use self save-ASP ten.some years DE money, **buy**.ASP one.CL t.v.set.

'There was an old peasant from Shandong, (lit. = Shandong had an old peasant) [who] used the money he saved for more than ten years **to buy a t.v. set.**'

NS2:3-5

In this sentence the Subject of *yǒu* could not be raised from the dependent clause, because the Subject position of the dependent clause is controlled by *yǒu*'s Object, and the other arguments of the matrix clause are overt. The main verb of the dependent clause is in bold and its arguments underlined:

So the Mandarin presentative is not completely analogous to the English raising verb 'seem' which never assigns a semantic role to its Subject.

In Mandarin then, there is no evidence for raising with respect to epistemic verbs; the evidence for raising with the presentative is ambivalent. This question will be reviewed in the section on 'Verb Complement Types in Mandarin' pg. 120.

5.5 Identification of Categories

The differentiation of sites for GB's PRO in general, and optionally controlled PRO in particular depends upon morphological cues that are entirely lacking in Mandarin. The *for*-complement diagnostic for LFG's Anaphoric and Functional control also appears to depend upon the potential for paradigmatic alternations of morphological forms, but arguments will be presented in the next chapter that the control processes themselves are motivated by semantics and therefore potentially 'discoverable' in any language, regardless of its morphological make-up.

5.6 Reference beyond the Sentence

In Mandarin, as we have seen, \emptyset elements may, and often do, have definite reference beyond the sentence. Both GB and LFG, as well as later approaches building on them (e.g. James Huang 1984; Reinhart, 1983; Levinson, 1991), mistakenly suggest that \emptyset elements that have no syntactically related antecedent **must** be indefinite and

conversely, arbitrary (indefinite) reference is simply a consequence of the lack of any plausible local antecedent. The 'pro-drop' parameter, and c-command in GB, and f-command in LFG, all restrict the potential antecedents of \emptyset elements to such an extent that their control mechanisms and categories cannot possibly be applied to \emptyset elements which do refer beyond the sentence. Not only is an explanation of these elements lacking, they are essentially prevented from being explained in terms of the existing theoretical machinery. While such constraints, or similar, may be required to account for the facts of a language like English, which favours explicit realisation of arguments, they are unlikely to be universal features. It seems self-evident that constraints on the reference of some \emptyset element must be relaxed in these theories, if the empirical data is to be fully accounted for.

LFG does provide two distinct interpretative mechanisms, one of which links a \emptyset element to a single antecedent and the other allowing the possibility of multiple antecedents. This fits most closely to actual observations on the nature of the variation of referential patterns in Mandarin.

6. CONCLUSION

Early (e.g. Li and Thompson, 1979) and some more recent, (e.g. La Polla, 1993) discourse and pragmatic approaches, tend to deny, ignore or misinterpret the role of syntax in the control of at least some cases of ellipsis. This is made easier by assumptions that all ellipsis is of essentially the same character, and that a lack of morphological complexity in Mandarin reflects, if not a lack of syntactic complexity, or semantic and referential specificity, then at least the irrelevance of these factors.

On the other hand, syntactic accounts tend to lean heavily on morphological signals, and thereby risk irrelevance to an account of referential phenomena in languages where complex morphology can be best described as extraneous.

The underlying attitude of the neo-Gricean approach appears to be essentially correct: that some instances of ellipsis are best explained in terms of grammatical determination, while others depend upon more subtle semantic and pragmatic factors. In

my view, their main failing is in associating a specific mode of interpretation with a specific formal device, rather than a specific syntactic structure. In addition, the general consistency constraints of neo-Gricean conversational implicature do not help to limit potential antecedents. The DRP disprefers co-reference of clause-mate co-arguments but in any structure other than a simple transitive clause this is insufficient to narrow the potential antecedents to a unique NP. Huang Yan (1994) suggests an antecedent search procedure, but it has been shown to be inadequate, since it predicts that arbitrary and non-local interpretations will only occur when all other potential antecedents are ruled out for semantic or pragmatic reasons.

Huang Yan's (1994) work does represent a major advance in that it approaches the problem of interpretation from the viewpoint of sets of syntactic structures, not morphology, draws on pragmatic and semantic factors, and allows that the explanations it offers are not intended to replace syntactic analyses entirely, but merely complement, or in some places offer alternatives to them. Huang Yan's work however unfortunately suffers from the same constraining influence as the more syntactically motivated attempts: that straight-jacket called a sentence. This is reflected in Huang Yan's belief that context, though clearly not peripheral in the determination of reference, can nonetheless be 'neutralised' (see Huang Yan, 1994: 130).

7. MOVING FORWARD:

The work now, and the work of this thesis, is to determine clearly, where the boundaries lie between the syntactic and pragmatic control of elliptic reference, and what processes are at work in each.

On the syntactic side, LFG offers the most promising way forward, in that it potentially predicts the kind of variation of referential effect evident in Mandarin: functional control is essentially semantically motivated and produces a unique linking of anaphor and antecedent via GFs; anaphoric control is less constrained and provides a mechanism potentially consistent with contextually variable, pragmatically sensitive interpretation. The main task then is to determine, for Mandarin, which sites involve

functional, and which anaphoric control, and how anaphoric control might be modified or complemented to better reflect observable patterns of co-reference in Mandarin.

Chapter Five

Interpretation of Dependent Clauses

1. INTRODUCTION: SYNTACTIC AND PRAGMATIC CONTROL

In Chapter Two it was shown that Mandarin ellipsis can be divided into referential sub-categories of variable and bound, and that these correlate strongly with paradigmatic sub-types (obligatory, contrastive and optional) on the one hand, and with structural context on the other.

In this chapter I propose that there is a consistent relationship between syntactic structure and control mechanism which can account efficiently for observed patterns of reference. I argue that obligatory bound ellipsis, which is impervious to changes in extra-sentential context, is interpreted via syntactic control, while variable ellipsis is interpreted via pragmatic and semantic cues. These two categories together constitute 60% of all ellipsis in the corpus (see Tables 4 and 5, p 25). Bound optional and bound contrastive ellipsis make up the remainder. They are discussed in Chapter Six.

This proposal of a consistent relationship between syntactic structure and control mechanism represents a departure from the standard LFG account of dependent clauses. Dependent clauses in serial verb constructions represent the largest single contributor to ellipsis in the corpus, accounting for 1/3 of all ellipsis, 1/2 of obligatory ellipsis and almost all bound ellipsis, so they are central in a discussion of ellipsis in discourse, and are the major focus of this chapter. The first half of the chapter concentrates on general issues regarding the processes of control in dependent clauses, and the second half discusses dependent clauses in Mandarin in general, and the corpus in particular.

As discussed in Chapter Four, in LFG some dependent Subjects are considered to be under anaphoric control, while others are under functional control. In this chapter I argue that there are no established tests which can reliably distinguish between anaphoric and functional control in English dependent clauses; that variation of referential effect

(i.e. bound vs variable) is best explained in terms of a distinction between syntactic and pragmatic control mechanisms, and that functional control is syntactic, while anaphoric control is more plausibly associated with the pragmatic determination of reference.

This means that, contrary to the standard LFG account, all ellipsed Subjects of dependent clauses should be under functional control. This in turn, means that verbs which take an XCOMP argument, invoking functional control may still take an SCOMP, i.e. a dependent clause with an **overt** Subject, as an alternative.

Support for this proposal is found in the analysis of Mandarin verb types. Moreover the variations in valency options are attributable to differences in verbal semantics.

This proposal provides an explanation of control which is a) motivated by semantics in a transparent and consistent manner; b) consistent with observations of constituency; c) depends upon fewer control mechanisms and interpretative principles than the standard LFG account, while maintaining a 1-1 correspondence between syntactic structure and control mechanism.

The Chapter concludes with a discussion of dependent clauses in the corpus, demonstrating that this model of control accounts for them all effectively.

2. CRITERIA FOR DISTINGUISHING SYNTACTIC AND PRAGMATIC CONTROL

When interpretation depends upon a syntactic relationship, it stands to reason that a) the anaphor and antecedent must be within the same sentence; b) the relationship between them must be consistent and describable in syntactic terms; c) if the syntactic relationship is significantly altered, the interpretation will alter; and d) co-reference will be assigned regardless of semantics. This last point predicts that absurd interpretations may arise when semantically inappropriate elements are combined in the relevant positions in a structure where syntactic control operates.

Conversely, where alteration of the context **outside the sentence** produces an alteration in the interpretation of reference, we can assume that reference is not syntactically determined. This is because syntactic relationships do not reach beyond the

limits of a single sentence. These observations can serve as criteria for determining the absence, if not the presence of syntactic control.

2.1 Optional Ellipsis and Pragmatic Control

When the two classifications of ellipsis based on referential effect and alternation with an overt NP are correlated (see Chapter Two), it can be observed that unique reference can be found in all distributional categories, while variable ellipsis only occurs where ellipsis is optional (see Table 3, pg. 25). Variable reference was established on the basis of a sensitivity to contextual changes beyond the sentence, so optional ellipsis cannot be syntactically controlled.

One qualification to this is that, in some cases, optional Subject ellipsis produces a more restricted set of possible interpretations than an overt pronoun would in the same structure, that is, the reference of \emptyset is a sub-set of the possible references of an overt pronoun. This suggests that there may still be some measure of syntactic control over the \emptyset element, however, in Chapters Six and Seven I propose that this results from the alternation of two similar structures: an extended structure with two Subject positions, and a reduced structure with only one.

2.2 Obligatory Ellipsis and Syntactic Control

By the same token, reference which is not variable need not necessarily be controlled by syntactic means; it might be pragmatically determined, either from outside the sentence, or within. Determination of the precise mechanisms requires a more in-depth investigation - the subject of this chapter and the next.

The major contributor to obligatory ellipsis is dependent Subjects in serial verb constructions, and in LFG these are considered to be under two types of control mechanism. While functional control clearly places syntactic constraints on dependent Subjects, strictly speaking anaphoric control does not, of itself, constrain interpretation at all. The correct interpretation of dependent Subjects under anaphoric control is achieved by way of the Obviation Principle. This does make reference to syntactic functions, but

the fact remains that LFG proposes two different control mechanisms for one syntactically, and referentially uniform class of syntactic structure.

Unfortunately, as I show in the next section, the tests and guidelines for distinguishing between anaphoric and functional control are either inconclusive, or because they depend on morphological signals, inappropriate to Mandarin. Before discussing the diagnostic tests I will review the LFG position on dependent Subjects.

2.3 Control Mechanisms for Subjects of Dependent Clauses: The Standard LFG Account

It will be recalled that while GB groups together all infinitival Subjects except those in raising constructions, LFG classes some Subjects of dependent infinitival clauses together with those in raising positions as functionally controlled, and the rest together with those in NP positions as anaphorically controlled.

Functional control verbs take an argument called XCOMP which cannot have an overt Subject; its Subject is controlled by a matrix argument. I will adopt Andrews' (in preparation) notation: XCOMP<OBJ>, where the GF in angle brackets is that of the controller. Anaphoric control verbs take an SCOMP argument, which can have an overt Subject.

The type of control associated with a verb is said to be lexically determined: only certain verbs invoke functional control, and for all others, regardless of the syntactic structure in which they occur, if a core NP is unexpressed, anaphoric control automatically introduces PRED 'pro' into f-structure.

Two tests have been proposed as diagnostics for control type: the ability to take a dependent clause with an overt Subject (Bresnan, 1982) and ability to appear in the wh-cleft structure (Andrews, in preparation).

2.3.1 Tests for Anaphoric vs Functional Control.

2.3.1.1 Functional Control and Overt Subjects

Bresnan (1982:331; and see also Horrocks, 1991) suggests that any verbs which **can** take a complement with an overt Subject, do not also invoke functional control.

English verbs which can appear with an overt dependent Subject are those like 'like', 'want', and 'ask'. Note that when the dependent Subject is not overt, it is constrained to co-reference with the matrix Subject:

117) "I_i like Ø_i to go."

Overt dependent Subjects can be realised in the form of NPs in accusative case (i.e. apparent Objects) and *for*-complements. These are shown to be overt subjects (not matrix arguments) by the evidence of passivisation in the first case, and extraction in the second. True direct Objects receive a semantic role from their verb, which permits passivisation. By this criterion, 'him' is an Object in 118) but not in 119) or 120):

118) I asked him to go
He was asked to go (by me).

119) I wanted him to go
* he was wanted to go (by me).

120) I like (for) him to go
* (For) him/he is liked to go (by me).

Since *for*-Objects are overt Subjects, any structure in which they actually appear involves neither functional nor anaphoric control, because there is no Ø argument to be controlled. There is therefore no self-evident connection between the ability to take a *for*-Object and a specific mechanism of control. However, there appear to be two possible reasons for supposing that functional control is incompatible with the ability to take an overt dependent Subject. Firstly, as functional control assigns the Subject GF of the dependent verb, allowing functional control in a verb that can also take an overt dependent Subject could lead to that GF being assigned twice, a contravention of the bi-uniqueness principle (Bresnan, 1982).

Secondly, if an overt Subject can appear, this verb clearly has at least one valency option where functional control is not invoked. It is best to postulate the minimum number of valency options possible for a lexical item, therefore, if such a verb also appears in a frame where no overt dependent Subject is possible, it is better to explain this, if possible, via some other existing control mechanism, and not postulate functional

control. Another control mechanism is available in the form of anaphoric control, which is independently necessary to account for the interpretation of Subjects of VPs in Subject position, where there is no functional controller available, for instance in the wh-cleft structures (see below pg. 114).

Alternative Valency Options

In fact, so long as XCOMPs, which trigger Functional control, alternate, rather than collocate with SCOMPs, where anaphoric control might operate, a single verb could take either argument, and still not contravene the bi-uniqueness principle.

There are some verbs which usually invoke functional control of a dependent Subject, but under certain circumstances cannot do so, and instead take an SCOMP with an overt Subject. Consider:

- 121 a) I forbid them to fight.
b) They were forbidden to fight.
c) *I forbid for them to fight.
d) I forbid there to be any fighting.
e) *There was forbidden to be any fighting.

In 121a) the NP 'them' is the Object of the matrix verb, as it can be passivised (see 121b); the dependent Subject is coreferent with the matrix Object. Example 121c) shows that 'forbid' does not take a *for*-complement, as anaphoric control verbs are supposed to do. In 121a) the verb 'forbid' can therefore be supposed to take an XCOMP where the XCOMP's Subject GF is controlled by the matrix Object, that is it is a functional control verb, XCOMP<OBJ>.

In 121d) the expletive 'there' is not a matrix Object, as evidenced by the fact that the passive sentence 121e) fails. Therefore, in 121d) 'there' is an overt dependent Subject and the verb complement is an SCOMP, i.e. it does not invoke functional control.

Therefore '*forbid*' must have two valency options:

forbid: {SUBJECT, XCOMP<OBJ>}
{SUBJECT, SCOMP}

The main characteristic of expletives like 'there' is that they cannot bear a semantic role. This suggests a correlation between specific semantic roles and the assignment of functional control. An SCOMP of 'forbid' clearly has special restrictions (i.e. the Subject must be expletive) and I suggest that this is a consequence of the semantics of the verb 'forbid' itself: the matrix Subject is not semantically compatible with the role of dependent Subject, because the verb specifies reduced volition of some agent with respect to a dependent event; its own Subject is volitional, in the sense that it voluntarily imposes a restriction on another entity, and therefore it is not an appropriate referent for the dependent Subject. In this structure there is no matrix Object, so the SCOMP subject is unable to receive a semantic role from the matrix verb, hence the use of the existential form. This semantic motivation for control-type is a point we will return to later.

There is nothing unusual about alternative valency options for a single verb. Verbs like 'want' can take a direct Object, to which it assigns a semantic role:

122) I want him

or an SCOMP with overt Subject:

123) I wanted him to go
* he was wanted to go (by me).

So 'want' also clearly has alternative valency options:

want: {SUBJECT, OBJ}
 {SUBJECT, SCOMP}.

Since valency options are a common-place, there is no *a priori* reason why XCOMPs and SCOMPS cannot alternate as complements of the same verb. On this basis, sentences like:

124) I want to go.

might plausibly involve the sub-categorisation frame:

want: {SUBJECT, XCOMP<SUBJ>}.

In other words, the alternation of overt and covert Subjects in dependent clauses can be readily accounted for by two different valency options. There is no need to

postulate two control mechanisms, one for verbs that can take a dependent clause with an overt Subject, and one for verbs that can't.

2.3.1.2 The Wh-cleft Sentence

Wh-cleft sentences, have been proposed as a test for functional control, (or rather the lack of it) (e.g. by Andrews, in preparation). In a wh-cleft structure the complement is 'dislocated':

125) What John wants is to go.

Raising verbs do not assign a Subject GF, and when they have a lexical Subject its GF must be assigned by the dependent verb via functional control. These verbs cannot occur in a wh-cleft structure:

126) * What John seemed was to understand.

It has therefore been proposed (Andrews, in preparation) that, conversely, all verbs which **can** be dislocated from their complement in this structure **do not** involve functional control. However, it is doubtful that functional control is actually blocked in a wh-structure. Note that an adjective can still be controlled by the matrix subject:

127) What John seems is understanding.

Moreover, sub-categorisation provides an adequate explanation for why 'seem' cannot occur in this frame: 'what' is a nominal and 'seem' cannot take a nominal complement. In other words, it can be assumed that verbs which cannot appear in wh-cleft constructions with a verbal complement have only XCOMPs as verbal complements, but it does not follow that verbs which can appear in wh-cleft constructions can **never** take an XCOMP. They may take an XCOMP in other structures, just not in this one.

So, neither of the overt Subject 'test' nor the wh-cleft test is a reliable diagnostic for control-type. Given that direct evidence is inconclusive, account should be taken of which analysis is more efficient and has the greater explanatory power.

3. EFFICIENCY OF EXPLANATIONS

3.1 The LFG Account: Inconsistent Categories and Additional Principles

As mentioned above, the standard LFG account requires that some dependent Subjects are under anaphoric control. This suggests that they are comparable to Subjects of clauses in NP positions, where control is also anaphoric (Bresnan, 1982). What anaphoric control does is simply introduce an element PRED 'pro' into f-structure. The predicate value of PRED 'pro' is held to be the same whether it is overt (a pronoun), or morphologically unexpressed; that is what the symbol PRED 'pro' means. In other words, the element introduced by anaphoric control is the semantic equivalent of an overt pronoun.

Now, the reference of an overt pronoun is relatively unconstrained; it can refer outside the sentence, and its reference must therefore be pragmatically determined. When the Subject of a VP in Subject position is morphologically unexpressed, and PRED 'pro' is introduced by anaphoric control, it may also refer beyond the sentence, and so must also be pragmatically determined, just as an overt pronoun is. This is shown in (128)

(128) "Ø going to Tilley's won't take long. I/you/she'll be back soon."

But when morphologically unexpressed PRED 'pro' occurs as Subject of a dependent clause, an additional principle, the Obviation Principle is required to constrain its reference to the matrix Subject.

In other words, though the reference of the semantic element PRED 'pro' is pragmatically determined, and anaphoric control functions specifically to introduce that semantic element into f-structure, the association between PRED 'pro' and pragmatic interpretation is countermanded in dependent clauses.

At the same time, though all dependent Subjects exhibit similarly constrained reference, this is achieved via two discrete mechanisms and an additional constraining principle. Overall this is an unnecessarily cumbersome account.

In fact there is no real reason to group some dependent clauses together with VPs in NP positions. The only similarity between VPs in NP position and dependent clauses,

apart from the morphological similarity in English, is their ability (in English) to assign the semantic role of agent to a *for*-Object. But I have already shown above that this does not imply the operation of a specific control mechanism. An account based on one control mechanism for one syntactic structure accounts for the data much more economically.

3.2 Alternative Proposal

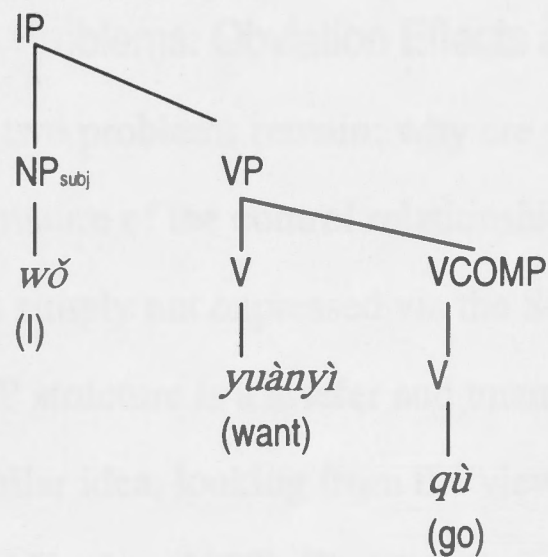
Since \emptyset dependent Subjects all exhibit similarly constrained reference patterns, I propose that they be assumed to be under a single control mechanism: functional control. They are represented in f-structure, but have no independent structural position. Since the reference of the Subjects of nominalised VPs in Subject position is variable, they are clearly not under functional control. Therefore they must be determined under anaphoric control. The ability of some verbs to take an overt Subject in dependent clauses can be explained by the alternation of the complement types, XCOMP and SCOMP as valency options of a single verb. Recall that Object control is the default value for Transitive verbs unless otherwise lexically specified. Some verbs with alternative valency options therefore exercise Object control in their transitive valency, but Subject control in their intransitive valency. These complement types can be derived readily from verbal semantics (see the section "A Semantic Classification of Mandarin Verbs" beginning pg. 130) and where both are available to a single verb, the choice between them can be argued to be pragmatic (see Chapters Six, and Seven), reflecting Gricean principles of avoiding ambiguity, and choosing the briefest means to express a given meaning.

Since valency options are standard in any account of syntax, this proposal requires no additional principles. An optional XCOMP option for verbs like 'want' would explain the obligatory nature of Subject ellipsis when Subjects are coreferent (as in 124) above) without the need to invoke an Obviation principle, or f-command, a clear gain in explanatory efficiency.

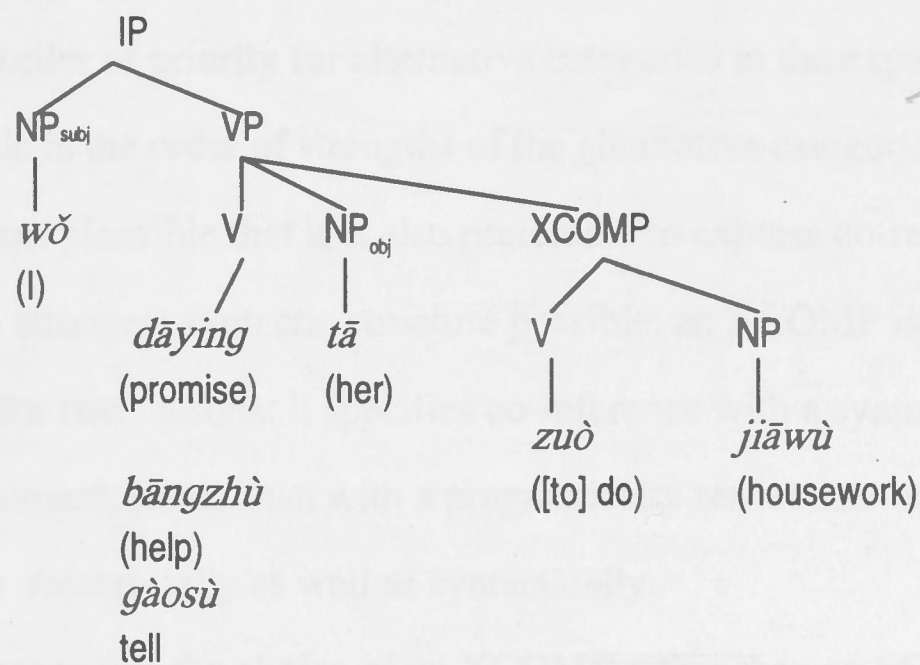
3.2.1 Representational efficiency

The functional control and obligatory ellipsis of dependent Subjects is consistent with the absence of a structural position for dependent Subjects: serial verbs in Mandarin can be readily diagrammed as nested VPs, eliminating a Subject position in either Subject or Object control:

129) Intransitive Matrix Verb



Transitive control:



The dependent verb is the lexical head of the XCOMP argument, and the matrix Subject and Object each stand in the same relationship to the dependent verb as the matrix Subject does to the main verb.

Constituency of the XCOMP can be seen from the fact that the XCOMP, but not the XCOMP-plus-post-verbal NP, can occur in topic position:

130) 做家务，由我帮助他。

- a) zuò jiāwù, yóu wǒ bāngzhù tā.
 Do housework from 1sg help 3sg
 '[As to] doing the housework, it is up to me to help him.'

- *他做家务，由我帮助。
- b) **tā zuò jiāwù, yóu wǒ bāngzhù*
 3sg do housework, from 1sg help.
 '* As to his doing the housework, I have to help.'

In these structures, the matrix argument can simply be considered to have scope over, or be linked to the embedded verb, in the same way that any Subject is associated with its verb.

3.2.2 Two Problems: Obviation Effects and Wh-cleft

Only two problems remain: why are overt dependent Subjects rarely coreferent, and what is the nature of the control relationship in 'wh-cleft' structures. I suggest that co-reference is simply not expressed via the SCOMP option, with an overt Subject, because the XCOMP structure is a briefer and unambiguous way to express this meaning.

A similar idea, looking from the viewpoint of interpretation, is discussed in Ford, Bresnan and Kaplan (1982). They propose a syntactic preference rule (1982:749) which states 'the order of priority for alternative categories in the expansion of a phrase structure rule is the order of strengths of the alternative categories.'

It seems plausible that it is also preferable to express co-reference with a matrix through the strongest syntactic structure possible: an XCOMP is likely to be a 'stronger' alternative for two reasons: it specifies co-reference with a syntactically related (i.e. matrix) argument, rather than with a pragmatically retrievable one; it depends upon the matrix verb referentially as well as syntactically.

In other words the choice of an XCOMP rather than an SCOMP to express co-reference is a pragmatic one. This is consistent with the fact that overt coreferent Subjects **are** permissible when stressed for contrastive purposes. We can say:

131) 'I want ME to leave, not him.'

even if we would not say

132) 'I want me to go to the shops.'

This clearly illustrates that there is **not** a syntactic constraint against an SCOMP with a coreferent Subject, as suggested by the Obviation principle; specification of co-

reference via an SCOMP is simply redundant in the normal course of affairs because of the availability of an XCOMP option with **only** this interpretation. This is consistent with the Gricean (1975) principle of economy.

As to what does control reference in wh-cleft structures, this is problematic in a GB analysis because the missing Subject must be either governed PRO, or a trace whose antecedent bears an autonomous thematic role, neither of which are acceptable. In LFG they can only involve anaphoric control, but this raises the question of how the reference of the introduced PRED 'pro' is constrained to co-reference with the Subject of the verb which is itself embedded within the matrix Subject clause. The matrix Subject constitutes an f-structure which contains the antecedent for PRED 'pro' but does not contain PRED 'pro', a violation of f-command. As it turns out, this construction is in fact quite marginal in Mandarin (see discussion pg. 139), and so is not germane to my topic.

So far then we have no clear evidence that functional control does not operate in the complements of verbs like 'want' and we have a certain amount to be gained in explanatory efficiency, by proposing that it does. Later on we shall see some corroborative evidence for this analysis from Mandarin.

3.3 Control Mechanisms Reviewed

A GB analysis of dependent clauses, of course, insists on a \emptyset element in a given structural position, but this ultimately makes explanation of the obligatory nature of ellipsis in dependent clauses more problematic. When PRO appears in an infinitival clause in Subject position, this position is ungoverned (and uncontrolled), and it is this which is said to account for the fact that no overt NP can appear there. PRO's reference in a dependent clause is determined by c-command but it is describable in these terms only so long as PRO is considered to be structurally present in a specific position. The problem is that this position, embedded under the matrix verb, could readily be analysed as governed. If PRO is structurally present in a governed position it is harder to account for its obligatory absence at the phonological level.

A modified LFG account provides the means for a resolution which seems essentially simpler than the complications of an ungoverned PRO. Anaphoric control can be associated with those positions where we have seen that variable reference ensues, while functional control can be associated with positions where bound reference ensues. There is, therefore a potential link between two structural positions, two patterns of reference, and two control mechanisms.

Though either the GB-based account of ungoverned PRO and optional vs obligatory control, or the LFG analysis of functional control of XCOMPs complemented by anaphoric control of VCOMPs and an Obviation Principle provide an adequate **description** of reference of VCOMP Subjects in English, neither of these accounts is particularly efficient in terms of accounting for both the obligatory nature of ellipsis in this frame, and the non-variable nature of its reference.

The ability to take an overt SCOMP Subject, whether as a 'surface' Object, or a *for-comp*, or to appear in a *wh-cleft* structure has no real standing as a test for the absence of Functional control, since there is no *a priori* reason why SCOMP and XCOMP options should not alternate as options for a single lexical item. There seems, in fact, to be no hard and fast test to distinguish squarely between Anaphoric and Functional control. However, Anaphoric control of dependent Subjects requires the introduction of an additional principle, where Functional control does not.

Though the evidence from English structures is inconclusive, the evidence from Mandarin lends support to this analysis. In the following sections we will see that Obviation does not pertain in Mandarin. In the absence of Obviation, and the presence of regular definite ellipsis, we find evidence for an anaphorically controlled dependent Subject which produces a referential effect that **contrasts** with that of functionally controlled XCOMPs.

4. VERB-COMPLEMENT TYPES IN MANDARIN

The question of the reference and control of dependent Subjects is, of course, complicated somewhat by the fact that Mandarin allows general definite ellipsis. I will

argue that Mandarin serial verb constructions are, by and large, XCOMP constructions involving functional control. Some Mandarin verbs take SCOMPs as a second option, and a few take SCOMPS only. Results of wh-cleft tests, though inconclusive, do not contradict this analysis.

Regular definite ellipsis means that Subjects of SCOMPs might in principle be optionally ellipsed, so long as they are retrievable (i.e. definite ellipsis). This increases the potential variability of surface structures relative to underlying structures and makes the issues of overt Subjects and obviation more complex. Table 21 shows that seven distinct sub-classes of verb can be defined by differences in valency, referential patterns and ability to take an overt Subject in a dependent clause. The following discussion demonstrates how this variability can be accounted for in terms of different subcategorisation frames with respect to XCOMPs and SCOMPs.

Table 21. Collocations of Mandarin verb types

	Transitive				Intransitive		
	dāying	qǐngqiú	shuōfú	gàosù	fǎnduì	níngyuàn	juédìng
Overt dep Subject	X	X	X	√	√	√	X
Subject Control	√	*√	X	X	X	√	√
Object Control	X	√	√	√			
∅ Object OK	√	?topic	topic	topic			

*Note that *qǐngqiú* need not specify its addressee argument. If the addressee is realised as a direct Object, it controls the ∅ Subject; if the addressee is an oblique, or indefinite, the matrix Subject controls the ∅ Subject.

The word 'topic' indicates that Object ellipsis is permissible only when the Object is bound by a sentence initial topic. The line between Transitive and Intransitive is broken because the verb *fǎnduì* can take a direct Object in simple sentences, but not in conjunction with a verbal complement, as will be demonstrated below.

4.1 SCOMPS

SCOMPs can be identified as structures which can have an overt Subject.

4.1.1 Apposed Clauses not Complements:

Since Mandarin gives few clues to indicate the nature of clausal conjunction, it will be useful, before proceeding further, to specifically exclude from the field of discussion certain structures which bear a surface similarity to SCOMPs.

A number of Mandarin speech act verbs and verbs of cognition, including *dāying* 'promise', *juédìng* 'decide', and *shuōfú* 'persuade', can appear apposed to a clause with an overt Subject which might be thought to express the speech-act, and therefore be an SCOMP e.g.:

- 133) 我答应了他我明天就去。
wǒ dāyingle tā wǒ míngtiān jiù qù.
 1sg promise 3sg 1sg tomorrow go
 'I promised him [and] I will go tomorrow.'

In fact the second clause is not an embedded complement of the first; it is a syntactically independent clause. In 133) the clause which follows is not necessarily to be understood as the content of the promise; such an interpretation would be an implicature only; it may, for instance simply express a necessary step in fulfilling the promise. So the collocation at 133) does not exhibit the same syntactic or semantic relationship as that between a speech act verb and a speech complement. Despite appearances then, these verbs do not introduce a dependent SCOMP.

4.1.2 Variation within SCOMPS

Of the seven verbs in Table 21), there are three, *gàosù* 'tell', *fǎndùì* 'oppose', and *níngyuàn* 'prefer', which can take an overt Subject in a complement clause not introduced by the complementiser *shuō*. The status of the post-verbal NP as an SCOMP Subject is reasonably transparent in two cases. Firstly, *níngyuàn*, unlike its English near-counterpart 'would prefer', is intransitive in simple sentences, being able to take a verbal complement only, and not a direct Object:

- 134a) 我宁愿他去北京。
wǒ níngyuàn tā qù Běijīng
 1sg prefer 3sg go Beijing
 'I'd prefer he went to Beijing'

- b) *我宁愿他。
 **wǒ níngyuàn tā*
 1sg prefer 3sg

The post-verbal NP is therefore clearly an embedded Subject. By the same token, *gaòsù* 'tell' is transitive, and has two post-verbal NP positions: one the matrix Object, and the other the dependent Subject:

- 135) 我告诉他我不会来。
wǒ gaòsu tā [wǒ bù huì lái]
 1sg tell 3sg 1sg NEG would come
 'I told him I wouldn't come.'

The third case is less obvious. This verb *fānduì*, can take a direct Object in a simple sentence, but has only one post-verbal NP position in a complement construction. The passive test for Object-hood is not conclusive in Mandarin, since semantic factors mitigate against passivisation in many cases even for true Objects. However direct Objects can be relativised and ellipsed, with the head extracted:

- 136) a) 我吃苹果。
wǒ chī píngguǒ
 1sg eat apple
 'I'm eating an apple.'
- b) 我吃的是苹果。
wǒ chīde shì píngguǒ
 1sg eat-DE COP apple
 'What I'm eating is an apple.'

137a) below, shows a transitive matrix verb with a verbal complement, and in b) the NP following the matrix verb is relativised, and ellipsed; the head is extracted to the position of a Copula complement.

- 137a) 我请求他去。
wǒ qǐngqiú tā qù
 1sg ask 3sg go
 'I asked him to go'
- b) 我请求去的是他。
wǒ qǐngqiú qù de shì tā
 1sg ask go DE is 3sg
 'The one I asked to go is him'

This shows that the extracted NP is an Object of the main clause. But arguments in embedded clauses cannot be relativised and extracted:

138a) 我告诉他我去。

wǒ *gaòsu* tā wǒ qù
1sg tell 3sg 1sg go
'I told him I will go'

b) *我告诉他去的是我。

*wǒ *gaòsu* tā qù de shì wǒ
1sg tell 3sg go DE is 1sg
'*The one I told him will go is me'
?I told him: it is I who will go

In 138a) *gaòsù* 'tell' takes an SCOMP and 138b) shows that the Subject of that SCOMP cannot be relativised, ellipsed, and bound by a Head in COPCOMP position. So this offers a diagnostic test to determine the status of the NP in question.

Now, in 139a) *fǎnduì* 'oppose' is shown with a direct Object in a simple sentence; and in 139b) the Object is relativised and ellipsed; the head is extracted to the Copula complement position. In 140) *fǎnduì* is shown in a serial verb construction. The NP after *fǎnduì* cannot be relativised and extracted, and therefore it is not a direct Object, but belongs inside the dependent clause.

139a) 我反对他。

wǒ *fǎnduì* tā
1sg oppose 3sg
'I oppose him'

b) 我反对的是他。

wǒ *fǎnduì* de shì tā
1sg oppose DE is 3sg
'The one I oppose is him'

140a) 我反对他去。

wǒ *fǎnduì* tā qù
1sg oppose 3sg go
'I oppose his/*him going'

b) *我反对去的是他。

*wǒ *fǎnduì* qù de shì tā
1sg oppose go DE is 3sg
'*The one I oppose going is him.'

So, we can therefore state that *gaòsù*, *níngyuàn* and *fǎnduì* all take some kind of sentential complement with an overt Subject.

4.2 Absence of Obviation

Now, the existence of regular definite ellipsis in Mandarin leads us to expect that these SCOMP Subjects can be ellipsed. In this case, by the standard LFG account, their reference will be determined under Anaphoric control: PRED 'pro' will be introduced into f-structure and interpreted under f-command and, possibly, Obviation. Short of introducing further constraining principles, we would expect all \emptyset SCOMP subjects to exhibit the same referential possibilities.

However, within this group of three, there is some variation as to the possible reference of a \emptyset Subject.

When *fǎnduì* 'oppose' takes a verbal complement with no overt Subject, it does **not** control the reference of the \emptyset dependent Subject:

- 141) 我反对在太平洋地区进行核试验。
wǒ fǎnduì zài tàipíngyángqū jìnxíng hé shíyàn
1sg oppose in Pacific.Ocean.region carry.out nuclear testing.
'I am opposed to nuclear testing in the Pacific region.'

This structure produces precisely the kind of reference that would be expected of a \emptyset SCOMP subject where there is no Obviation principle: indefinite or arbitrary reference.

By contrast, a \emptyset dependent Subject of the Mandarin *níngyuàn* 'would prefer' tends to be interpreted as coreferent with the matrix Subject, even in the face of a compelling context for disjoint reference. In example 142) the topic NP has only one potential semantic role (excluding a vocative usage), that of Subject in the dependent clause, and should therefore control that GF, but this sentence is more likely to be rejected as incoherent than accorded this interpretation.

- 142) *爸爸，我宁愿去。
*Bàba_i, wǒ níngyuàn \emptyset _i qù
Dad_i, 1sg prefer \emptyset _i go
'*Dad_i, I'd prefer \emptyset _i went.'

However, just as in English (see 131) above), an overt coreferent Subject is permissible, with an emphatic or contrastive reading.

143) 我宁愿我去而他不去。

wǒ níngyuàn [_{scomp}wǒ qù ér tā bù qù.]

1sg prefer 1sg go and 3sg NEG go

'I'd rather I went and not him.'

This means that the co-reference of \emptyset in 142) cannot derive from the Obviation principle, as it stands. So the evidence from both *fǎnduì* and *níngyuàn* is that the Obviation Principle does not obtain in Mandarin.

The situation with *gaòsù* is a little more complicated: it can take two somewhat different complements: one aspect-marked, and representing the content of the speech act, and the other being understood as irrealis, and expressing the intent, rather than the content of a speech-act. The appearance of an overt dependent Subject is confined to the first type, and there are different referential constraints on the dependent Subject in each case.

When the complement is irrealis, the dependent Subject must be ellipsed, but unlike *níngyuàn* 'prefer' it **must** refer to the matrix Object:

144) 我告诉他(我/你/他)会来。

wǒ gaòsù tā [wǒ nǐ / tā huì lái]

1sg tell 3sg [1sg/ 2sg/ 3sg may come]

'I told him [I/ you/ (s)he might come.]'

145) 我告诉他来。

wǒ gaòsù tā [lái]

1sg tell 3sg [come]

'I told him [to come].'

Once again, since a \emptyset Subject need not be bound to the matrix Subject, and an overt SCOMP Subject can be coreferent with the matrix Subject, these referential patterns cannot be the result of obviation.

So within the verbs that can take an SCOMP with an overt Subject, there is one verb which produces just the referential pattern that would be expected of an SCOMP with a PRED 'pro' Subject in the absence of obviation, and two different patterns neither of which can be explained by the standard formalisation of the Obviation Principle.

The fact that *fǎndùi* produces a different pattern from that of *níngyuàn* and *gàosù* suggests that those two verbs involve something other than just SCOMPs with optionally ellipsed Subjects, controlled by anaphoric control.

Returning to the case of *gàosù* 'tell', it can be seen that its complement in 144) above is an embedded Sentence, but it is independent of the matrix clause in terms of tense, mood, aspect and illocutionary force. It is comparable to a tensed SCOMP in English. I will call these independent SCOMPs, in contrast to SCOMPS which have restrictions on aspect-marking. The complement in 145) is an embedded sentence which cannot have independent aspect or illocutionary force and is comparable to an infinitival Complement in English. Given that no overt Subject is possible in this second structure in Mandarin, but the reference of the Subject is bound, not variable as with *fǎndùi*, there is no reason to suppose it is anything other than an XCOMP. In other words, *gàosù* can take either an XCOMP<OBJ>, and an SCOMP in complementary distribution.

Similarly, the simplest account of the referential patterns of *níngyuàn* is to postulate that it takes an XCOMP<SUBJ>, as well as an SCOMP. The lack of control by the matrix Subject with *fǎndùi* 'oppose' can be accounted for by the fact that it takes an SCOMP only, not an XCOMP.

The only remaining problem is that of the constraints on ellipsis of SCOMP subjects with *níngyuàn* 'prefer', and *gàosù* 'tell'.

4.2.1 Pragmatic Constraints on SCOMP Subject Ellipsis

Since Mandarin allows definite Subject ellipsis, it might be thought possible for ambiguity to arise as the result of two identical surface structures:

146) 我告诉他来。
wǒ gàosù tā, [Ø lái]
1sg tell 3sg [Ø come]
'I told him, [[they/we/you etc.] are coming]'

and

147) 我告诉他来。
wǒ gàosù tā [lái]
1sg tell 3sg [come]
'I told him to come.'

It seems that, even though structural ambiguity is possible in principle, the ellipsis of SCOMP Subjects is avoided in practice if structural misinterpretation would produce a shift in reference. A \emptyset Subject in a realis SCOMP was rejected by informants when it was not coreferent with the matrix Object, even in a context with a suitable, retrievable antecedent:

148a) 你告诉他 (我 / 他) 应该晚上来吗?

nǐ gaòsù tā_i [scomp wǒ / tā_{i/j} yīnggāi wǎnshàng lái ma?]
 2sg tell 3sg_i 1sg /3sg_{i/j} should evening come QUEST

'Did you tell her_i /she_{i/j} should come this evening?'

b) 我告诉他 (* \emptyset / 我 / 他) 会来。

*wǒ gaòsù tā_i, [scomp * \emptyset / wǒ tā_{i/j} huì lái]*
 1sg tell 3sg_i \emptyset / 1sg/ 3sg_{i/j} may come

'I told her_i [[* \emptyset / / she_{i/j}] should come].'

One informant stated explicitly that a \emptyset Subject in 148b) was unacceptable because the matrix Object should be interpreted as controlling the dependent Subject in this case. It is important to realise that this holds true for independent (i.e. tensed) SCOMPs in Mandarin, where ellipsis would not normally arise and the Obviation Principle would not normally be invoked in English, just as much as for dependent (untensed) SCOMPs.

Again this may be explained by Ford, Bresnan and Kaplan's (1982) 'syntactic preference rule' (see above pg. 115).

5. ANAPHORIC CONTROL RECONSIDERED

It will be recalled that in standard LFG, anaphoric control is subject not only to Obviation, but to a requirement of f-command. The main purpose of f-command seems to be to restrict the potential controllers of \emptyset dependent SCOMP subjects under anaphoric control to just the Subject or Object of the matrix verb, while obviation restricts reference further to just the Matrix Subject, as well as having consequences for the reference of overt Subjects. In languages which don't have regular definite ellipsis, the main role for anaphoric control is with regard to the reference of \emptyset Subjects of VPs in NP position, and coreferent deletion in conjoined clauses. But in languages like Mandarin, PRED 'pro'

must be invoked in many other structural positions, for instance in independent clauses, and in Object positions. When it comes to the reference of PRED 'pro' in these positions, the f-command requirement is too restrictive.

We have already seen the idea of Obviation undermined by the fact that the disjoint reference of overt Subjects is subject to exceptions. Moreover, these exceptions can be accounted for in terms of a pragmatic choice between two similar structures. If we can find a pragmatically motivated account of reference in conjoined clauses, and an alternative account of VPs in NP positions, then we can safely assume that all \emptyset dependent Subjects are actually under functional control, and neither f-command nor obviation would be necessary; anaphoric control could be allowed to be relatively unconstrained, as the variability of relationships between optionally ellipsed elements and their nearest coreferent NP suggests it should. What I proposed, is that all instances of PRED 'pro' introduced by anaphoric control are actually pragmatically controlled by reference to notions like topicality and focus. The evidence for these pragmatic influences on the interpretation of \emptyset elements is presented in Chapters Six and Seven.

To conclude the argument on control mechanisms then, there is some evidence for variation in the reference of \emptyset SCOMP subjects in Mandarin, but it does not conform to the predictions of Obviation associated with Anaphoric control. Rather it reflects a situation where verbs may take either an XCOMP or an SCOMP argument, but the use of each tends to be referentially specialised so as to avoid structural ambiguity resulting in referential mis-interpretation or incoherence.

The lack of formal distinction between complement types in Mandarin creates no problem for interpretation because the XCOMP/SCOMP alternation, would be triggered by whatever subtle aspects of pragmatics and semantics usually trigger the choice of one valency option over another, without recourse to additional principles. The next section suggests a possible semantic motivation for this distinction.

6. A SEMANTIC CLASSIFICATION OF MANDARIN VERBS

The variability of ellipsis and co-reference shown in Table 21 can now be readily accounted for in terms of a variation of available valency options. These are presented in the chart below.

Table 22. Valency Options of Mandarin Verb types

	SCOMP ONLY	SCOMP / XCOMP		XCOMP ONLY		
Transitive	fǎndù 'oppose'	<SUBJ> ?	<OBJ> <i>gaòsù</i> 'tell'	<SUBJ> <i>dāyìng</i> 'promise'	<OBJ> <i>shuōfú</i> 'persuade'	<OBJ>SUBJ> <i>qǐngqiú</i> 'ask-request'
Intransitive	?	<i>níngyuàn</i> 'prefer'		<i>juéding</i> 'decide'		

So far it has been suggested, on the evidence of surface structures, that different verbs have different valency options in terms of SCOMPs and XCOMPs. The central question addressed in this thesis is how listeners know which complement type is found with which verbs, and so which control mechanism to assume. The answer to this question lies in the relationship between complement types and matrix verb semantics.

The distinction between Subject and Object control has been argued to follow readily from the semantic content of the verbs themselves (see Curnow, 1992; Kroeger, 1993). The following section examines the semantic motivation for the invocation of functional control of the Subjects of dependent clauses. I assume along the lines proposed by Hopper and Thompson (1980) that notions like agency, volition and affectedness are relevant to levels of transitivity. I suggest that different levels of semantic transitivity are expressed syntactically by different valency options. In discussing the semantic make-up of verbs I adopt Wierzbicka's (1988:9-12; 1992: 7-27) approach based on a Natural Semantic Meta-language (NSM). In this approach, the complex meanings of lexical items can be explicated by way of a number of meaningful components; these are themselves composed of simple terms which are thought to be universal elements of cognition (Wierzbicka, in press:35, 73-74). I will not attempt full explications of verbs here, but simply refer to components of meaning that are relevant to the issue of control. These are components which identify the semantic roles assigned to entities, especially in the realm

of agentivity ('do'), volition ('want'), and communication of desire or intention ('say' / 'want').

6.1 SCOMP Only - Unspecified Agency over Dispreferred Event.

Verbs which cannot take an XCOMP, like *fǎndù* 'oppose' are verbs which do not specify agency with regard to any event within their scope. Their semantic explication (i.e. their PRED value which refers to subcategorised arguments) would contain a component like:

'I do not want Y to happen.'

Moreover, the dependent event is irrealis, so even an explicit Subject has low agentivity. Since agency is uncertain and unspecified, it cannot be construed as controlled by the matrix verb.

Since the matrix verb does not specify the agent of the dependent event, that event must be expressed via a structure that allows the agent to be specified independently, i.e. a structure with a Subject position available for the speaker to fill.

6.2 XCOMP Only - Control of Agency

Intransitive verbs that involve XCOMPs only are verbs, like modals, that represent the matrix Subject as agentive with respect to a second event. Their explication would contain a component like:

'I want to do X'

'I can do X.'

A structural position in which a dependent agent can be independently specified is therefore unnecessary.

Looked at in this way, the appearance of XCOMP or SCOMP complement types can be seen to relate in a straightforward and transparent way to semantic specifications by the matrix verb with respect to agency in a dependent event. Intransitive verbs which take an XCOMP will naturally involve Subject-control, and interpretation is unproblematic. Transitive verbs could specify control by various combinations or alternative GFs.

6.3 Transitive Verbs

6.3.1 XCOMP <SUBJ> - Retained Agency

Verbs like *dāying* 'promise' also explicitly express the determination of the matrix agent to perform a secondary event. They contain a component like:

'I tell you: I will do something.'

So they also specify the identity of the agent (and therefore typically, the Subject) of a dependent activity. These verbs involve another participant (i.e. the Direct Object), but only as beneficiary or addressee and not as agents. Since agency is retained by the performer of the speech act, encoded by the matrix Subject, there is no need to specify separately the agent of the dependent activity, and so there is no need for a Subject position in the dependent clause.

The general pattern is that the XCOMP argument arises as a consequence of the conventional mapping of agent to Subject GF, but it is not simply a matter of semantic control; the control mechanisms are grammaticalised. This can be seen by comparing a passive sentence from the corpus with its active counterpart, note that the main verb is the same lexical item in both the following sentences, though English has no exact active/passive pair:

149) 他被警察赶下车了。

tā bèi jǐngchá gǎn xià chē le
3sg by police rush descend train ASP
He was driven off the train by the guard
NS8:14a

150) 他赶下车了。

tā gǎn xià chē le
3sg rush descend train ASP
He rushed off the train.

In both cases it is the Subject the 3rd person singular pronoun *tā*, which controls the dependent Subject, and not the agent, which is the policeman in the first sentence and the pronoun in the second.

Thus the use of an XCOMP is plausibly derived from lexical semantics which specify retention of agency by a participant encoded by a matrix argument. The speech

act verb *dāying* has Subject control because the matrix Object is not implicated in an agentive role in the dependent event, only in the role of addressee.

6.3.2 XCOMP<OBJ>- Transferred Agency

Verbs like *shuōfú* 'persuade', specify the influence of the matrix agent over the performance of a secondary event **by a different** agent. Moreover, they specify the effective realisation of that event (or the speaker's conviction that it will be realised), so the dependent Subject is highly agentive. They contain a component like:

'I did something to X. Because of this I know X will do something.'

Since agency is specified, they take an XCOMP, and since it is transferred, they are Object-control verbs: XCOMP<OBJ>.

6.3.3 'Dual Control' Verbs: {XCOMP <OBJ> / XCOMP<SUBJ>}

Verbs like *qǐngqiú* 'request' are more subtle in their implications with regard to agency. Recall that *qǐngqiú* invokes Object control when the addressee is encoded as an Object, but Subject control when the addressee is encoded as an Oblique. This can be represented: {XCOMP <OBJ> / XCOMP<SUBJ>}, where the elements inside curly brackets are alternative valency options. *Qǐngqiú* expresses the communication of a **desire** by one person for **another** to act in some way, i.e. to take an agentive role in a dependent activity, but they also specify that the dependent event has not been, and may never be realised. They contain components like:

'I want X to do Y.'
I say something to X.
Because of this X knows I want X to do Y'

Separately neither the matrix Subject, nor the dependent Subject is strongly agentive with respect to the dependent activity.

Again, the XCOMP can be seen to be associated with specification of a (desired) agent, by the matrix verb. When there is no overt Object to which the agentive role is assigned, the role of the addressee is weakened to one of enabling or allowing the dependent event to be performed by the requester (i.e. Subject-control of the dependent

verb). This is also evidence that control is not derived directly from semantics, but is mediated through syntactic functions.

Assuming that functional controllers must be core arguments of the verb, when the addressee is encoded as an Oblique, then Subject control is the only possible interpretation, reflecting a semantic makeup:

wǒ xiàng X qǐngqiú Y

1sg to X ask Y = I ask X [if I can] do Y

'I want to do Y if X says I can.

I say something to X.

Because of this X knows I want X to say I can do Y'.

6.4 Bi-valent Verbs

6.4.1 {SCOMP|XCOMP} - Unspecified Agency over Desired Event

Verbs which can take either an SCOMP or an XCOMP also have complex and subtle semantics in the area of transferred agency.

6.4.2 {SCOMP | XCOMP<SUBJ>}

Verbs like *níngyuàn* 'would prefer' express a positive attitude or affective response by their Subject towards the realisation of an event. However they are semantically unspecified as to who is the agent of that event:

'I want X to happen.'

This is reflected in the Mandarin by the fact that *níngyuàn* takes no direct Object. Since the verb itself does not specify agency, it can be optionally specified either syntactically, via an XCOMP, where however, the only available controller is the Subject, or by way of an independent argument, i.e. not one assigned a semantic role by the matrix verb, in an SCOMP.

Since the XCOMP option does not permit of any variability in agency, i.e. no other entity is specified in its PRED value, it is only the activity itself, and not the agent which falls within the scope of the matrix Subject's attitude. This accounts for the pragmatic contexts in which one form is likely to be chosen over the other: the XCOMP option

allows contrast of actions only, while the SCOMP option allows for co-ordination and therefore contrast of either the agent, or the action.

In the following examples, a \emptyset coreferent (i.e. XCOMP) Subject is only acceptable in a structure which contrasts predicates; an overt coreferent (i.e. SCOMP) Subject is acceptable whether or not it is the focus of a contrast. This means it can be used in, and would most probably be restricted to, contexts of simultaneous contrast of agent and activity. The context could be set outside the sentence, or within it as in 152). Notice that when an SCOMP structure is used for predicate contrast, this is specifically signalled by stress on the verb rather than the NP.

151) 我宁愿去而不呆在家。

a) *wǒ níngyuàn qù ér bù dāi zài jiā*
 1sg prefer go and not stay at home
 'I'd prefer to go and not to stay at home.'

*我宁愿去而他不去。

b) **wǒ níngyuàn qù ér tā bù qù*
 1sg prefer go and 3sg NEG go
 '*I'd prefer to go and not him.'

我宁愿我去而他不去。

c) *wǒ níngyuàn Wǒ qù ér tā bù qù*
 1sg prefer 1sg go and 3sg not go.
 'I'd prefer for ME to go, not him'
 'I'd prefer that I go and he not go.'

152) 我宁愿我去而他呆在家。

wǒ níngyuàn Wǒ Qù ér tā Dāi zài jiā
 1sg prefer 1sg go and 3sg stay at home
 'I'd prefer that I WENT and HE stayed HOME'
 'I'd prefer for me to GO and HIM to stay HOME.'

153) 我宁愿我去而不呆在家。

wǒ níngyuàn wǒ Wǒ Qù er Bù dāi zài jiā*
 1sg prefer 1sg/*1SG go and NOT stay at home
 'I'd prefer that I/*I WENT and didn't STAY HOME.'
 'I'd prefer for me to GO and not STAY HOME.'

Thus, the appearance and reference of a dependent Subject can be readily explained as the consequence of an alternation of XCOMP and SCOMP arguments. The choice between the two can be related in a straightforward manner to the specification of agency

by entities which are themselves specified in the semantic explications of the verbs. Comparison of these categories with English counterparts (left as an exercise for the reader) suggests that *for*-complements are possible in English only where the matrix verb does not specify the agent of a dependent event; where an independent agent is not and cannot be specified (i.e. where Functional control pertains uncontroversially), it is because the semantics of the matrix verb determine agency.

At the same time, the co-reference patterns in Mandarin have clearly been grammaticalised to some extent, since Subject control occurs in passive structures too, where the Subject of an XCOMP is controlled by the matrix patient, not the matrix agent.

6.5 Reduction of Ambiguity in Dual Control Structures.

We saw above in the discussion of the bi-valent verb *gaòsù* (pg. 127) that ellipsis of an SCOMP Subject after a controlling Object is resisted because it could lead to ambiguity. Verbs which alternate between Subject and Object control could also potentially lead to ambiguity: if the controlling (definite) Object of such a verb were optionally ellipsed, the resulting structure would be identical to the intransitive valency with an indefinite Object, and Subject control.

In 154), when the Object is realised, only the Object, and not the Subject can control the XCOMP Subject.

154) 我请求他去。
*wǒ₂ qǐngqiú tā₁ [xcomp Ø_{1/*2} qù]*
 1sg ask 3sg Ø go
 'I asked her to go.'

NOT I asked her and went

But in 155) where there is no overt Object for *qǐngqiú* in the second sentence, it is ambiguous. There is a discursual implicature that the pronominal addressee/ Object of the first sentence *tā_x*, might be the addressee of *qǐngqiú* 'ask' in the second sentence, but there is also the possibility that the addressee role is not assigned, allowing Subject control.

155) 老王害羞，不愿意请求他去，

Lǎo Wáng₁ hàixiū, Ø₁ bù yuànyì qǐngqiú tā_x qù
 L. shy NEG want ask 3sg go
 'Lao Wang is shy and does not want to ask him to go'

所以我请求去。

Suǒyǐ wǒ₂ qǐngqiú (Ø_x) Ø_{2/(x)} qù
 so 1sg2 ask (Ø_x) Ø_{2/(x)} go
 'So I₂ asked [if I₂] could go.' OR possibly:
 '? So I asked [him_x] to go.'

In this limited context, my informant preferred the Subject control interpretation, but the final interpretation would depend on whether the further context suggested that 'I' had spoken to 'him' or that 'I' wanted to go. If the participant represented by the pronoun in 155) continued as a central topic in the discourse, it would be more likely to be assumed to control the Ø Object of *qǐngqiú*, and therefore the Ø Subject of *qù*.

If a potential addressee appears in Sentence-topic position, it will certainly be interpreted as Object of the verb *qǐngqiú* and Object control will result. This is illustrated in the second part of 156):

156) 你去北京，请求你的父母去吗？

nǐ qù Běijīng, qǐngqiú nǐde fùmǔ qù ma?
 2sg go Beijing ask your parents go QUEST
 'When you go to Beijing, will you ask your parents to go'

爸爸，我请求去，妈妈不请求。

Bàba₁, wǒ₂ qǐngqiú Ø:obj₁ [_{xcomp}Ø:subj qù], māma bù qǐngqiú
 Dad 1sg ask Ø Ø go, Mum NEG ask
 'Dad, I'll ask to go, but not Mum.'

NOT: Dad, I asked if I could go.

Though topics can have an oblique relationship to the comment, they seem to be assigned an Object GF wherever possible.

So, rather than ambiguity of reference, there seems instead to be a variability of interpretation derived from the varying degrees with which a Ø Object can be controlled from other NP positions in the discourse; so long as an Object GF is assumed to be specifically assigned, then Object control pertains; participants in previous sentences clearly do not exercise so strong a control on the interpretation of Ø Objects as

sentence initial topics do. This emerges as an important principle of pragmatic interpretation.

6.6 Conclusions: Obviation, Semantics and Argument Type

The solution of permitting options of XCOMP and SCOMPs clearly accounts for a wealth of surface variability in Mandarin ellipsis and reference of dependent Subjects in a reasonably efficient manner.

In general terms we see that the SCOMP valency is associated with verbs which do not specify the identity of agents in a dependent event, and the XCOMP option, which invokes Functional control, is associated with verbs which do. Anaphoric control is only involved if the Subject of an SCOMP is optionally ellipsed, much like optional ellipsis of the Subject of an independent sentence. But ellipsis of SCOMP Subjects is dispreferred where surface ambiguity might result.

Verbs which take both XCOMP and SCOMPs exhibit a change in the precise semantic roles assigned in each frame, specifying agency, either retained or transferred, when no dependent Subject is permissible, but not when an SCOMP with overt Subject appears.

Though there is a strong correlation between semantics and argument type, it is not the case that the patterns can be explained on purely semantic grounds. Passive verbs demonstrate consistency along syntactic rather than semantic lines, exhibiting Subject rather than agent control. In addition, for certain speech act verbs, the choice of syntactic structure appears to be a means to signal the linking or separation of the roles of addressee and agent: these roles are always assigned at the level of lexical semantics, but it is the choice of GF for the addressee which determines whether the roles are shared by the matrix Object or Subject.

Thus functional control can be seen as a grammaticalisation of the semantic specification of agency by a matrix verb, with respect to a dependent event.

6.7 Results of WH-cleft test in Mandarin

The conclusions about syntactic structure reached on the basis of semantic analysis were tested by the wh-cleft test. Recall that verbs with an XCOMP as their only verbal complement cannot occur in the structure.

157) *what John tried was to leave

In order to test the hypothesis that most VCOMPs in Mandarin are functionally controlled, eight native-speakers of Mandarin were asked to comment on the acceptability of six of the controlling verbs discussed above as well as two others, in this syntactic frame:

158) NP VERB *de shì* X.
 NP VERB NOM COP X
 what NP VERBs is to X.

Where X was a VP semantically appropriate to the verb being tested, DE is a nominalising particle and *shì* is the Copula. For example:

159) ? 他请求张三的是离开。
 **tā qǐngqiú Zhāngsān-de shì líkāi*
 3sg request Z.-NOM is leave
 'What he asked Zhangsan was to leave.'

Table 23. Proposed Control Mechanisms: Mandarin and English

Mandarin verb	Informant judgement on Nominalisation test			English semantic counterpart	Control Mechanism in English (by <i>for-obj</i> criterion)
	√	?	X		
<i>qǐngqiú</i>	1	4	3	ask request	Anaphoric/ Functional
<i>shuōfú</i>	1	3	4	persuade	Functional
<i>yuànyì</i>	2	3	3	want	?
<i>níngyuàn</i>	2	2	4	prefer.	Anaphoric
<i>yào</i>	2	4	1	want/require	?
<i>juéding</i>	4	3	1	decide	Functional
<i>fǎndù</i>	4	3	1	oppose	Functional
<i>xīwàng</i>	7	-	1	hope	Anaphoric

It should be noted that semantic similarity does not necessarily mean identical control mechanisms, as the different status of the English verbs 'ask' and 'request' with

regard to taking a *for*-Object indicates; another indication of the subtlety of semantics involved. Mandarin, like many languages distinguishes between asking a question (*wèn*), and making a request (*qǐngqiú*), and since the 'ask for' collocation in English occurs only with the sense of making a request, I chose the latter for the purpose of the investigation. Informants were asked to judge the collocations as acceptable, unacceptable, or odd/borderline.

There was no absolute agreement amongst informants on any verb. However, only one verb, *xīwàng* 'hope' was found generally acceptable in this frame (7 positive responses, one rejection). This verb is similar to *níngyuàn* in all points discussed so far: it is an intransitive Subject control verb which can take an SCOMP with overt Subject. I have no explanation at this stage as to why it was accepted in this syntactic frame with so much higher a level of consensus than other verbs.

Another two verbs, *juédìng* and *fǎnduì* 'oppose' were borderline overall, with half the informants finding them acceptable, three finding them odd, and one rejecting them in each case. Of these, *fǎnduì* has also been proposed as taking an SCOMP only, but *juédìng* was thought to take only an XCOMP.

It should be remembered though that *juédìng* can introduce a tensed SCOMP i.e. a direct 'speech' complement, reporting as it were the content of a decision. Because its complement is in irrealis (in future time) the form of a tensed SCOMP would be identical to a tenseless one, unlike other verbs which introduce realis SCOMPs. It may be this which licensed its appearance in this frame for some speakers.

For the remainder, each verb was found acceptable by only one or two informants, with the balance evenly split between outright rejection and 'questionable'. Another possibly disruptive factor is that all the informants here were bilinguals; this may have affected their judgements.

The clearest result to emerge from this test is that there is no clear correlation between the English verb's ability to take a *for*-complement, and the Mandarin counterpart's acceptability in this frame. Informants themselves had no clear insights into

the reasons for the variation. Further exploration with more variation in contextualisation may yield more meaningful results.

While not clear-cut, these results are not at odds with the suggestion that Verbal complements in Mandarin are indeed more likely to involve functional control exclusively.

7. COMPLEMENT STRUCTURES IN THE CORPUS

So far I have discussed serial verb constructions in a general way only. To conclude this chapter I will discuss the variability of complement structures found within the corpus.

7.1 Directional Complements as Compounds

The directional verbs *qǐ* 'rise' *chū* 'go/put out' *qù* 'go to' and *lái* 'come' and combinations of them, are generally attributed with aspectual, or metaphorical meanings, and are so widely used, that I have treated them as elements of compounds rather than as dependent verbs. The distinction is, however purely a matter of convenience. Like all Mandarin 'compounds', this one is in fact indistinguishable from a syntactic construction. By way of illustration, the corpus includes a string:

- 160) 我们讨论出一个道理来。
wǒmen tāolùnchū yīge dàolǐ lái
 we discuss-forth a truth come
 we discussed_[xcomp] put-forth truth [_{xcomp} come]]

'We produced a truism through discussion.'

NS7:15a

This is a metaphorical usage of the verb *chū* 'to put forth' which is commonly suffixed to a variety of verbs expressing motion or verbal or visual expression etc.. The directional verb *lái* 'come', often referred to as a locative particle, almost invariably appears with it, coming after the Object NP, unless the latter has been topicalised, in which case *lái* follows directly after *chū* (or *qǐ* etc.):

- 161) 这个道理我们早就讨论出来了。
zhège dàolǐ wǒmen zǎo jiù tāoluunchulaile
 this truism 3pl early then discuss.out.come

In 160) *tǎolùn* can be analysed as a matrix verb; *chū* 'put forth' is both the XCOMP<SUBJ> of *tǎolùn*, and also a matrix verb for *lái*; *lái* in turn, is an XCOMP<OBJ>. Because of the conventional nature of such collocations of directional verbs, I have counted them as part of a single compound.

Another compound which is probably derived from a serial verb construction is the resultative, *kàn-dǒng* 'read-understand'. This differs from a verb complement construction in that the Object of the compound cannot come between the two verbs:

162a) *看这本书懂。

**kàn zhèběn shū dǒng*
read this book understand

b) 看懂这本书。

kàn dǒng zhèběn shū
read understand this book.
'Read and understand the book.'

This potentially represents a further stage of incorporation.

7.2 Semantic Sub-Types of Serial Verb Constructions

In the Corpus, 33 verbs are involved in 63 complement structures. Their distribution is shown in Table 24; the glosses convey the rough meaning of the Mandarin words, but do not necessarily share the same sub-categorisation frames. The structures and their usages in the corpus are discussed in more detail below.

7.2.1 Modals

Of the 64 structures, almost 1/3 are complements of five 'modal' verbs (*kě* can, *yào* need/must, *gǎn* dare, *néng* can, *yīnggāi* should), all are XCOMP<SUBJ>.

7.2.2 Cognition and Affect

The next largest sub-group of complement constructions is 14 speech or 'thought complements' with 9 verbs of cognition, affect and speech: six take SCOMPs (*shuō* 'say', *xiǎng*₂ 'think', *zhīdào* 'know' *júde* 'think/feel', *jiànyì* 'suggest', *míngbái* 'understand', accounting for 11 SCOMP structures; two occur with XCOMP<SUBJ> (*xiǎng*₁ 'want', *yàojiú* 'demand') and one with an XCOMP<OBJ> (*xiāngxìn*

'believe/trust'). The last is interesting because it is clear from the context that the post-verbal NP is a matrix Object controlling the Subject position of a speech-act verb, it is not an overt dependent Subject:

- 163) 他相信小张说他们已经断了。
tā xiāngxìn XiǎoZhāng shuō tāmen yǐjīng duànle
 3sg believe X say 3pl already break-ASP
 'She believed X's saying they had already broken up.'
 after NS6:16

NOT 'She believed **that** X said they had already broken up.'

Nine out of the 11 SCOMPs were independent SCOMPS with their own tense, aspect or illocutionary force, the exceptions were *jiànyì* 'suggest' and *yàoqiú* 'demand' with irrealis complements.

Table 24. Semantic Sub-types in Mandarin Serial Verb Constructions

Modals Ability	Cognition		Movement		Joint agency		Pres.
	Thought/ Speech-act	Affect	self	other	Inst.		
<i>kě</i> can	<i>xiāngxìn</i> believe	<i>xiǎng</i> want	<i>húì</i> return	<i>dài</i> lead	<i>xǐe</i> write	<i>bāng-</i> <i>zhù</i> help	<i>yǒu</i> have
<i>yào</i> must	<i>zhīdào</i> know	<i>juéde</i> consider	<i>(húì)lái</i> come (back)	<i>gǎn</i> drive	<i>yòng</i> use	<i>zhǎo</i> seek	<i>shì</i> be
<i>gǎn</i> dare	<i>xiǎng</i> think		<i>qù</i> go	<i>tǎolùnchū</i> extract in discussion	<i>jǐ</i> jam onto		
<i>yīnggāi</i> should	<i>míngbái</i> understand		<i>zhù</i> reside		<i>dú dào</i> come on while reading		
<i>néng</i> can	<i>yàoqiú</i> demand <i>jiànyì</i> suggest <i>shuō</i> say						
	11	3	10	3	6	2	
20	14		13		8		5
62	Total COMP structures						

7.2.3 The Emphatic Copula

The emphatic Copula is a special complement construction of which there were two instances in the discourse. They have been classified in Table 24 as a kind of presentative, as the Copula highlights the predicate which follows, see 164)

- 164) 好多同学都是分别了三四年后，
hǎo duō tóngxué dōu shì [_{xcomp} [_{adj} *fēnbié-le sān sì nián hòu*]
good many classmates all are separate-ASP three four year after
第一次见面。
[_{Np} *dìyī cì jiànmiàn*]
first time meet
'A good many class-mates were meeting for the first time after having been separated for three or four years.'
NS4:6

The Subject of the copula controls the agentive role of its Complement verb, just as with any other matrix verb. The complement verb is nominalised in that it appears as head of a Classifier Phrase. No overt possessor is possible here, just as no overt Subject is possible in a typical verbal XCOMP.

Thus the Copula appears to function like other Subject control verbs, but its XCOMP must appear as a nominal element; a parallel between VPs in dependent NP and dependent VCOMP positions.

7.2.4 Motion Verbs

Another 10 structures are associated with 3 verbs of motion, (*huí* 'return [to]', *lái* 'come', *qù* 'go'), and one of location, (*zhù* to reside). These are all XCOMP<SUBJ>.

There were three more complements involving transitive verbs which could also be grouped under the semantic class of motion, but involving a transference, or sharing of agency as well (*dài* 'lead', *gǎn* 'drive', *tǎolùnchū* 'extract through discussion' All three are XCOMP<OBJ> usages.

7.2.5 Transferred Agency

Eight more structures also involve a kind of shared or transferred agency, with six being classifiable as 'instrumental' or 'medium' structures (*xǐe xìn gào wǒ* 'write a letter telling me', *jǐ huǒchē húi jiā* 'squeezing on the train to return home'; *yòng qián mǎi diànshìjī* 'used money to buy a tv set'; and, more a medium than an instrument: *dú dào yī tiáo xiāoxi shuō* 'read an article saying'. In the first three examples, the instrument might perhaps be considered as sharing agency in the dependent event, but the clauses could also be considered instrumental adjuncts. What sets them apart from true adjunct clauses though is that the matrix Subject NP cannot be transferred to the Subject position of the second verb. Therefore these are best accounted for as XCOMP<SUBJ> complements.

The dependent Subject in the last example could also be coreferent with the matrix Subject, but only if the first clause is interpreted as an adjunct:

- 165) 我读到这条消息, 说...
- a) *wǒ dú dào zhè tiáo xiāoxi, shuō*
1sg read-ASP this.CL article say...
'reading this article I said...'
- = 读到这条消息, 我说...
- b) *dú dào zhè tiáo xiāoxi, wǒ shuō*
read-ASP this.CL article 1sg say...
'reading this article I said...'

As it stands in the original text, the dependent Subject is coreferent with the matrix Object, and no such movement across clauses is possible for the dependent Subject; so the structure involves an XCOMP<OBJ>.

The other two transferred agency complement structures involve the verb *bāngzhù* 'help' whose primary semantic denotation is specifically an expression of shared agency, and the verb *zhǎo* 'seek' which seems to take an optional 'complement' as an adjunct of purpose. This is unfortunately a confusing juxtaposition of terminology, but the apparent contradiction arises because, while the matrix verb can stand alone with just its Object

NP, and demands no further semantic specifications, so that the clause which follows is a semantic adjunct, that clause is nonetheless syntactically dependent upon it to provide an argument to which it can assign its Subject GF, so it is a syntactic XCOMP.

7.2.6 Presentatives

The last five structures involve the possessive-presentative *yǒu*. In three cases there was an overt Subject, one a locative, one a time-frame (*yìge xīngqī* 'a week') and the other a 1st person singular personal pronoun; the others had no overt or retrievable Subject. In four of these presentative structures, the complement was clearly an XCOMP<OBJ>. However, as discussed above (pg. 98 e.g. 115), repeated as 166) below) the presentative *yǒu* does not always assign a semantic role to its Subject; an overt Subject can be viewed as a Subject raised from the SCOMP, i.e. it is an example of an XCOMP<SUBJ> argument with the presentative.

166) 你给我邮来的信，
 [nǐ gěi wǒ yóulái-de xìn]_{TOP1}
 2sg to 1sg post-come-DE letter
 'The letter you sent me.'

一半都是英文，
 [yībàn dōu shì yīngwén]_{comment1}
 half all COP English
 'fully a half was in English.'

当时我真有很多看不懂。
 [dāngshí wǒ zhēn yǒu [hěn duō]_{TOP2} kànbùdǒng Ø]_{comment2}
 that-time 1sg true have COP much see-NEG-understand
 'At the time there was really a lot I couldn't understand.'
 NS5:11

The verb *yǒu* can therefore take two XCOMP options, one with Subject control and one with Object control. This means, contrary to the standard LFG hierarchy OBJ>SUBJ, it is **not** the case that Subject-control only operates in the absence of an Object. However, when the matrix Object does not control a Ø dependent Subject, it must control a Ø dependent Object, as in 166). Like a Topic, the *yǒu*-Object must have a role in the dependent clause. When the matrix Subject is accounted for, having been assigned from the XCOMP<SUBJ>, the matrix Object can only control the dependent Object position.

But when the matrix Subject is unfilled, the matrix Object may or may not control the dependent Object, depending on the semantics of the dependent verb and its arguments. In 166a) there was Object-Object control but in 167) there is not; the Object refers outside the sentence:

- 167) 只有奶奶看不懂。
zhǐ yǒu nǎinai kànbùdǒng
 only have Grandma see.NEG.understand
 'There was only grandma [who] didn't understand'

7.3 Functional Complements

Finally, the structure illustrated in 167) is similar, in its ambiguity, to relative clauses in Mandarin. However, in Mandarin a relative clause usually precedes the Head, and the particle DE comes between the relative clause and its head. A relative clause with overt head is shown in 168):

- 168) 我有很多看不懂的地方。
wǒ yǒu hěnn duō [rel kànbùdǒng de] dìfang
 1sg have very many see.NEG.understand DE places
 'There are many parts [which] I don't understand.'

This position is not possible for the NP '*hěnn duō*':

- 169) *有我看不懂的很多。
 **yǒu wǒ kànbùdǒng de hěnn duō*
 have 1sg see.NEG.understand DE very much

Therefore the NP '*hěnn duō*' does not seem to be the Nominal head of a relative clause in 166). It is also reminiscent of the infinitival relative in English; Haegeman (1991:427) cites the example:

- 170) I_j want [_{NP} a man_i [_{CP} O_i [_{IP} PRO_j to love t_i]]]
 as an infinitival relative derived from that of a conventional relative clause:

- 171) I_j want [_{NP} a man_i [_{CP} who_i [_{IP} I can love t_i]]].

However, these clauses in English also clearly take a different structural position from infinitival relatives, which modify the NP. Consider:

- 172 a) I wanted an experience to remember.
 b) An experience to remember is what I wanted.
 c) The real experience to remember was bungy-jumping off the Eiffel tower.

- 173) a) I wanted this man to marry
 b) *This man to marry is who/what I wanted.
 c) ? The real man to marry was bungy-jumping off the Eiffel tower.

Nor is it a purpose adjunct, which can occur in Sentence initial position. Consider:

- 174) a) I was looking for a man to keep busy.
 b) To keep busy, I was looking for a man.

Sentence 174a) is ambiguous between a purpose reading: 'I was looking for a man in order to keep **myself** busy' and what I will call the function meaning: 'I was looking for a man whom I could keep busy', (among others which are not relevant here). 174b) can have only the first of these two meanings.

The function reading is ruled out for b) because when an infinitival clause appears in sentence-initial position, its Object is no longer bound to co-reference with the matrix Object, the way it is when the infinitival clause is in sentence-final position. This alters the meaning significantly. These sentences also differ from those discussed in Lasnik and Fiengo (1989) as the \emptyset Objects in those are all bound by Subjects.

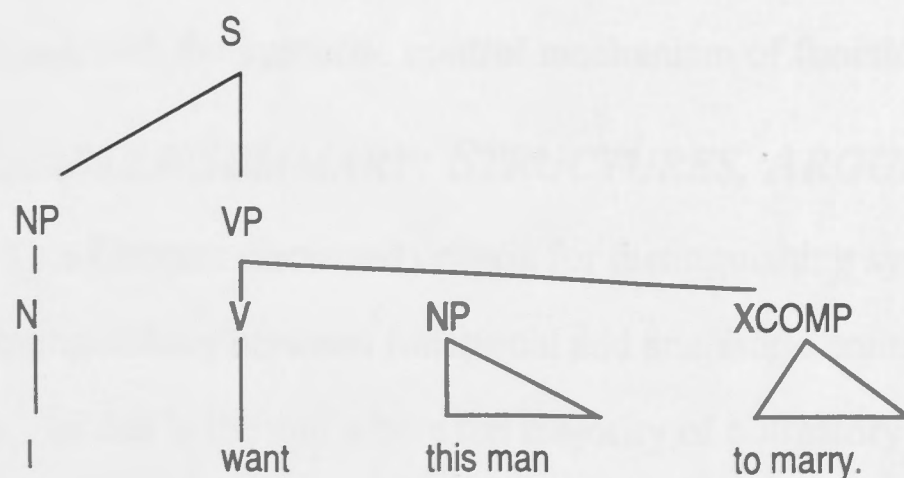
Since the infinitival clause occupies a position not available to relative clauses the binding of its reference is problematic. It can however be readily accounted for by an XCOMP structure where the dependent and Matrix Objects are shared (i.e. Functionally controlled) also. This can be represented: {XCOMP<SUBJ> \uparrow OBJ}.

A position within the VP is supported by distribution:

- 175) a) I want this man to marry. \approx
 b) I want to marry this man. \approx
 c) This man, I want to marry.

I propose the structural position of the function adjunct to be:

176)



The Mandarin example at 166) seems similar to this function complement, and can be reasonably described in terms of an XCOMP<SUBJ>↑OBJ also.

On the other hand, an alternative account based on pragmatic factors is proposed in Chapter Six. The exact status of this construction is still problematic and requires further investigation. There was only one instance in the corpus.

Another similar case had the matrix Object as the **patient** of a passive dependent verb:

- 177) 有一二件衬衫要烫烫。
yǒu yī èr jiàn chénnshān yào tàngtàng
have one two CL shirts need iron
'There are one or two shirts need ironing/ [I] have to iron.'
NS3:21

This can be interpreted simply as an XCOMP<OBJ>, the more usual argument for the presentative verb.

8. CONCLUSION: SYNTACTIC CONTROL OF DEPENDENT SUBJECTS:

The majority of \emptyset dependent Subjects in the corpus are obligatorily ellipsed. Given the reliability of predictions of reference afforded by a syntactic account of argument sharing, it seems reasonable to conclude that this major contributor to bound ellipsis in Mandarin discourse involves functional control. Semantics is nonetheless central to the recognition of sub-categories of verb, based on the kind of arguments they take, and therefore the control mechanisms they invoke, for instance in the distinction between SCOMP and XCOMP, and between Object- and Subject-control.

In conclusion Obligatory bound ellipsis in complement constructions can be clearly associated with the syntactic control mechanism of functional control.

9. CHAPTER SUMMARY: STRUCTURES, ARGUMENTS AND CONTROL

This Chapter discussed criteria for distinguishing syntactic and pragmatic control, and distinguishing between functional and anaphoric control. It focused on dependent clauses, as this is the site where the majority of obligatory ellipsis occurs in the corpus. It was suggested that the two control mechanisms of LFG, Functional control and

Anaphoric control, might be appropriately aligned with the two kinds of ellipsis, bound and variable. This would necessitate postulating functional control for all obligatorily ellipsed dependent Subjects, a departure from the standard LFG account. The benefit of this is that it would allow the constraints of Obviation and f-command to be dispensed with, making the account of control both more efficient, in terms of the number of mechanisms involved, and more 'natural' in terms of plausible semantic motivation, as well as a stronger correlation between control type and referential effect.

As a consequence it was proposed that all dependent clauses where an overt Subject of any reference is not permissible, are XCOMPs, which have no Subject position. Verbs which may take a dependent clause with overt Subject are simply considered to have two alternative valencies.

Support for this proposal was found in the referential patterns of Mandarin verbs, where obviation does not pertain. This allows a contrast between functional and anaphoric control to be observed. A classification of Mandarin verbs is proposed, based on different valency options: SCOMP only, XCOMP only, and SCOMP and XCOMP in alternation.

The reference patterns of 62 complement constructions in the corpus are readily explained within this account. They can be described in terms of 5 basic groups of matrix verbs:

- i) modals (ability and obligation) all intransitive XCOMP<SUBJ>;
- ii) motion and location verbs transitive XCOMP<OBJ> and intransitive XCOMP<SUBJ>;
- iii) verbs of speech or cognition SCOMP; positive affect: volition = XCOMP<SUBJ>, 'trust' = XCOMP<OBJ>;
- iv) presentatives XCOMP<OBJ>/ XCOMP<SUBJ ↑OBJ>;
- v) sharing of agency XCOMP<SUBJ+OBJ>, instrumental XCOMP<SUBJ>, and 'medium' XCOMP<OBJ> structures.

Of the matrix verbs controlling XCOMPs, 30 were intransitive, and 24 were transitive; of the latter 13 were Subject control verbs, 9 were Object control verbs and two involved shared control.

SCOMP structures are clearly in the minority, accounting for only 11 out of the 62 complement structures. Of these, nine were independent SCOMPs, seven had overt Subjects, two had a \emptyset Subject controlled by a topic, one had a \emptyset Subject of indefinite reference, and one was 1st person singular (i.e. referred to the writer), in other words, ellipsed definite SCOMP Subjects were all highly topical and retrievable. Thus ambiguity between SCOMP and XCOMP structures is kept to a minimum.

Chapter Six

Bound Reference

We have taken for granted for too long now that there must be a strong correlation between *main-finite clause* syntax and the foregrounding function in discourse. In the main, our faith in this correlation hinged primarily on Indo-European facts of grammar, and relatively little on cross-linguistic studies of matching grammar with discourse....It is only by wedding the two methodologies - cross-language comparison and function-based empirical studies of text - that we may hope to some day resolve both ends of our semiotic equation - and thus the equation itself. -Talmy Givon, 1987:185

1. INTRODUCTION: STRUCTURE AND PRAGMATIC CHOICES

This chapter examines mechanisms of control in relation to the remaining sites where ellipsis has bound reference. For Subject ellipsis these have been divided into two categories: subordinate constructions, where one clause appears as the argument of another, and co-ordinate constructions, where clauses are relatively independent of each other. The subordinate constructions include relative clauses, which involve obligatory ellipsis and VPs in NP positions, where ellipsis is either obligatory (in Object position) or Optional (in Subject position); co-ordinate constructions include conjunctive and adjunctive structures, which are associated with both contrastive and optional ellipsis; Object ellipsis involves Topic-comment structures ("left-dislocations") where ellipsis is obligatory, and Objects in SCOMPs, where James Huang (1984) claims it is contrastive.

Contrastive ellipsis is shown to be associated with inter-clausal relationships of causality and correlation, regardless of the conjunctive device used. I argue that contrastive ellipsis, and bound optional ellipsis result from pragmatic choices between two similar structures. The semantics of causality prompt speakers to choose a more reduced structure. This line of reasoning is developed in the next chapter.

∅ Objects are held to be pragmatically interpreted, but manifest as uniquely bound to sentence topics because of the pragmatic status of the latter.

2. SUBORDINATE STRUCTURES

2.1 Relativised Constituents

There were only two instances of relative clauses in the data. Since the question of the reference of relativised constituents is both complex and general to all languages (see discussion in Chapter Four pg. 94), and the data available from Mandarin generally, and this corpus in particular is unlikely to throw light on the problem, relative clauses will not be further entered into. It will be assumed that, whether the trace analysis of GB or the metavariable analysis of LFG, or some other analysis yet to be developed is the more correct, the control of the reference of relativised constituents involves a mechanism universal to all languages. Once the learner can recognise the relative clause and its head, they then have recourse to this universal mechanism for interpretation.

One point worth mentioning is the difficulty presented, for production, by the fact that only Subjects and Objects can be relativised in Mandarin. Personal experience shows that this often poses a problem for English-speaking learners: if they ellipsis the NPrel, they produce a structure which appears to have a head, but no gap to which it corresponds; they accurately judge this structure to be ungrammatical, but are unable to formulate an alternative. The possibility of including an overt pronoun does not seem to occur to them, and the general response is to avoid constructions like:

178) 我借他的书的人。

wǒ jiè tāde shū de rén

1sg borrow his book DE person

'[The] person whose book I borrowed' (lit: I-borrowed- his-book person.)

2.2 Equative Copulas

Just two examples of optional bound ellipsis in my texts were heads of relative clauses in copula constructions like 179):

179) 我的论文是关于普通话异读字的。

wǒde lùnwén shì guānyú pǔtōnghuà yìdúzì de Ø

1sgPOSS paper COP about MSC homographs DE Ø

'My paper was one Ø about homographs in Modern Standard Chinese.'

The Ø head of the relative clause is constrained to co-reference with the Subject.

The Mandarin copula is equative, in other words, its Subject and complement must be

coreferent. Because of this, control in this structure could equally be considered semantic control, because of the semantics of the copula, or syntactic control, because the copula construction represents a special syntactic frame.

2.3 Subjects of VPs in NP Positions

In English VPs in NP position are either gerunds or infinitives and all the latter can take a *for*-complement to express their agentive role in this position. In Mandarin, verbs in NP positions do not undergo morphological derivation and they simply retain the ability to take a lexical Subject:

- 180) 他去北京很好。
Tā qù Běijīng hěn hǎo
 3sg go Beijing COP good
 '[For] him to go to Beijing is good.'
- 181) Ø 去北京很好。
 Ø qù Běijīng hěn hǎo
 Ø go Beijing COP good
 'Going to Beijing is good.'
- 182) 最好是他去北京。
zuì hǎo shì (tā) qù Běijīng
 Best is (3sg) go Beijing
 '[What would be] best is ([for] him) to go to Beijing.'

2.3.1 VPs in Subject Position

As discussed in Chapter Five, when VPs appear in Subject positions, Ø Subjects have variable or indefinite reference:

- 183) 一星期里我有三天的课
yī xīngqī lǐ wǒ yǒu sān tiān de kè
 one week-in 1sg have three day POSS class
 'In one week I have classes on three days...'

是下午六点下课。
shì xiàwǔ liù diǎn xiàkè
 COP afternoon six o'clock finish
 ...which finish at six pm.'

下课时天已经很黑了,
xià kè shí tiān yǐjīng hěn hēi le,
 finish class time sky already very dark-ASP
 'When class gets out, it's already very dark.'

∅ 住了离梅西近些是必要的。

∅ *zhù-le lí Méixī jìn xiē shì bìyào-de*

∅ live-COS separate Massey close some COP necessary-DE

'To move closer to Massey was essential.'

In this example the relevant argument is the Subject of the last line. The context makes it clear that it is the writer for whom it was necessary to move closer to Massey, but the only possible lexical, or even structural antecedent is two sentences back, so it cannot possibly be a syntactic controller. With different semantic content, the same structure could readily have the last sentence referring to an entity represented by an element in a different structural position, or not represented at all, and having indefinite reference, e.g.

184) ∅ 改了时间是必要的。

∅ *gǎi-le shíjiān shì bìyào-de*

∅ change-ASP time COP necessary

'To change the time [of the class] was essential.'

This makes sense because VPs in Subject position are not dependent on any verb and therefore their Subjects cannot be determined by functional control. This is all entirely consistent with the analysis of anaphoric control of VPs in Subject position proposed in Chapter Five.

However, VPs in Object positions cannot always have a realised Subject.

2.3.2 VPs in Object and Copula-Complement Positions

When nominalised verbs appear as Objects their agentive role appears to be under the functional control of the verb which assigns them the Object GF, see 185) and 186). The controlling verbs are underlined.

185) 我们带他们参加了一次去房县旅游。

wǒmen dài tāmen cānjiā-le [NP:OBJ yīcì ∅ qù Fángxiàn lǚyóu,

1pl take 3pl [_{xcomp} attend-ASP [NP one.CL ∅ go F._{xcomp} travel]]]

'We took them to join in [_{NP} ∅ visiting Fangxian [_{xcomp:adj} ∅ to go travelling]'

186) ∅ 还无生孩子的打算

∅ *hái wú [np:obj [mod ∅ shēng háizi de] dǎsuàn]*

∅ still lack [np:obj [mod ∅ bear child DE] intend]

'[She] still has no intention of having a child.'

In other words they seem to stand in much the same relationship to their head verb as an XCOMP does to the matrix verb. This suggests that VPs in Object position, unlike those in Subject position, are functionally controlled. Example 186) also exhibits an additional aspect of VPs in NP position, viz. nominalised verbs also appear to exercise functional control over the Subject position of a clause which modifies them. When *dǎsuàn* appears as a main verb, it takes an XCOMP<SUBJ>:

- 187) 她打算生孩子。
tā_i dǎsuàn ∅_i shēng hái_izi
 3sg intend ∅ bear child
 'She intends to have a child.'

and it appears to retain this ability to assign a Subject GF to another verb even from NP position. The lexical item *dǎsuàn* 'intend/intention' in 186) is itself the Object of the negative possessive verb *wú*, and it is modified by the DE clause *shēng hái_izi* 'to have a child'. The head *dǎsuàn* 'intend/intention' is neither Subject nor Object of the verb in the modifying DE clause, instead the Subject position within the modifying clause is controlled by an argument in a relationship of possession to the head *dǎsuàn*. This analysis is quite consistent with the idea that nouns have arguments, and a Noun's Subject is its possessor. Moreover, the DE clause is the standard means for expression of a possessive modifier in Mandarin, so an 'intend verb' takes a genitive complement in Mandarin, just as it does in the English construction 'the intention **of** having a child', cf. 'the intention **of** my sister'.

It appears then that certain nouns can take a kind of XCOMP<POSS> argument. This is comparable also to VPs which appear as Copula Complements, but though it is in an NP position, the complement may lack nominal marking, and so can be analysed as a straightforward verbal XCOMP.

In conclusion, VPs in Object or COP-Complement positions appear therefore to be XCOMPs and their Subjects are functionally controlled, while VPs in Subject positions are independent arguments and their Subjects are pragmatically determined.

3. CO-ORDINATE STRUCTURES: CONJUNCTION AND APPOSITION

There are 48 instances of ellipsis associated with conjunctive structures in the corpus. 31 involved overt conjunctions, and 17 were appositional structures. In all, 11 different lexical conjunctions were used in the corpus, as well as two aspectual suffixes and three clitics indicating temporal relationships.

3.1 Structural Positions of Conjunctions

There are several structural positions available to conjunctions in Mandarin, but individual conjunctions may be restricted to only some of these positions.

3.1.1 Clausal Conjunctions:

Some conjunctions appear between two clauses, but they must **follow** the Subject of the second clause, when it is overt. These include: *jiù* 'then'; *hái* 'still'; *què* 'though, but, however'; *yě* 'also, and, what's more'. For example:

- 188) 我们还请了一餐,
wǒmen hái qǐng-le Ø yī-cān,
 1pl also invite-ASP Ø one-meal
 '...we also invited [them] for a meal.'
 NS7:

- 189) 天越来越冷了Ø也晚得早了。
tiān_i yuè-lá-yuè lěng-le Ø_i yě wǎn-de zǎo-le
 Weather more-and-more cold-ASP, Ø also late-DE early-ASP
 'The weather's getting colder and [it] (sic) gets late earlier'
 NS3:10

Some conjunctions can **also precede** the Subject of the second clause. They include: *huòzhe* 'or, either..or'; *dàn (shì)* 'but, however'; *hòulái* 'afterwards'; *yúshì* 'whereupon/ and so'; *fǎnzhèng* 'anyway'.

- 190) Ø不知Ø从何而来, 反正(它)淡淡的。
Ø_i bù zhī Ø_j cóng hé ér lái, fǎnzhèng Ø_j/ta_j dàn dàn de
Ø_i NEG know Ø_j from where ER come, anyway Ø_j/3sg_j faint.RED.DE
 [SUBJ NEG PRED₁ [SCOMP]] CONJ [SUBJ PRED₂]
 I don't know where [it] came from, [but] in any case (it) was very faint...
 NS4:10

- 191) Ø以后回国Ø也可以教我Ø,
Ø_i yǐhòu huí guó Ø_i yě kěyǐ jiāo wǒ_k Ø
Ø_i After return country, Ø_i also can teach 1sg_k Ø
 'Later, when [you_i] come home, [you_i] can also teach me_k Ø.'

但（你）不能说我笨。
dàn Ø_i/nǐ_j bù néng shuō wǒ bèn
 but Ø_i/2sg NEG can say 1sg thick
 but (you) mustn't call me thick.'
 NS5:23

192) 他应该要求退换或者去打官司。
tā yīnggāi yāoqiú tuìhuàn huòzhe (tā yīnggāi) qù dǎ guānsi.
 Subj modal {[xcomp xcomp] conj (Subj modal) [xcomp [xcomp obj]]}
 3sg ought demand swap or (3sg ought) go hit court
 'He should demand to swap [it], or take [them] to court.'

These conjunctions can all collocate with an overt Subject in the second clause and so Ø Subjects in these structures were classified as optionally ellipsed. The last example is an apparent exception because the conjoined VP is also an XCOMP, that is, the conjunction takes place at the level of the dependent clause, and a single modal has scope over two XCOMPs, an overt Subject would be possible in the second clause, **if** the modal were repeated there also. This is evidence that the ellipsis of these Ø Subjects is, at least potentially, the consequence of constituent conjunction below the level of the VP. That is, there is no second Subject position.

3.1.2 Paired 'Correlative' Conjunctions

There are also paired conjunctions: one appears in the first clause and the other in the second:

193) 不是心脏，就是伤到肺。
bù shì xīnzàng jiù shì shāng dào fèi
 COP - CONJ heart COP - CONJ wound arrive lungs
 'If it wasn't the heart, it was a stab in the lungs.'
 NS8

There is some variation with regard to the possible positions of these conjunctions relative to overt Subjects, within a given sentence, for a given pair in different sentences, and for different pairs. These are discussed in more detail below.

3.2 Constituent Conjunctions and Ellipsis Type

In total, 23 conjunctive structures involved optionally ellipsed Subjects with bound reference, and 24 involved contrastive reference, where, two Subjects can only be realised if they are of disjoint reference. Only one conjunction *jiù* 'then' was associated

with more than one kind of ellipsis: two contrastive and 5 optional. Appositional structures without overt conjunctions, also fell into both categories: 12 contrastive and 5 optional. Only one conjunction was associated with obligatory ellipsis. This was the conjunction *ér* 'and/but/so'.

3.2.1 Obligatory Ellipsis: VP Conjunction

Er is a conjunction from classical Chinese, used mainly in formal or literary writing. It is often described as an 'all-purpose' conjunction, because it corresponds to a number of more semantically restricted conjunctions in English, including concessive or contrastive conjunctions like 'but' and 'though' as well as co-ordinating conjunctions like 'and'. *Er* can appear in co-ordinate structures with two Subjects, or just one.

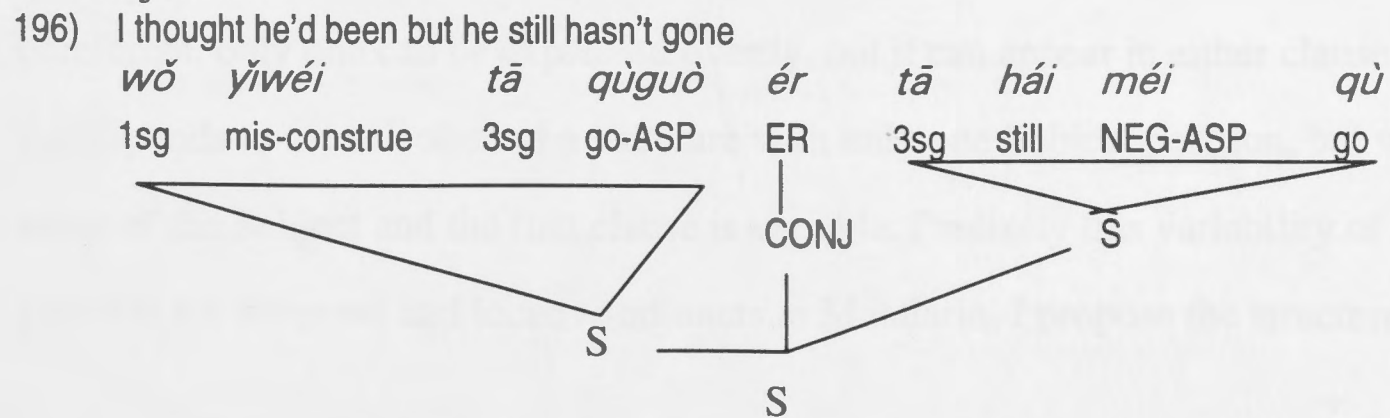
194) 我以为他去过而他还没去。
wǒ yǐwéi tā qùguò ér tā hái méi qù
 1sg mis-construe 3sg go-ASP ER 3sg still NEG-ASP go
 I thought he'd been but he still hasn't gone

195) 当时样子你可想而知了。
dāngshí yàngzi nǐ kě xiǎng ér zhī ∅ le
 that-time appearance you can think ER know ∅ ASP
 lit: [My] appearance at that time, you can consider and know [it]
 NS5:12

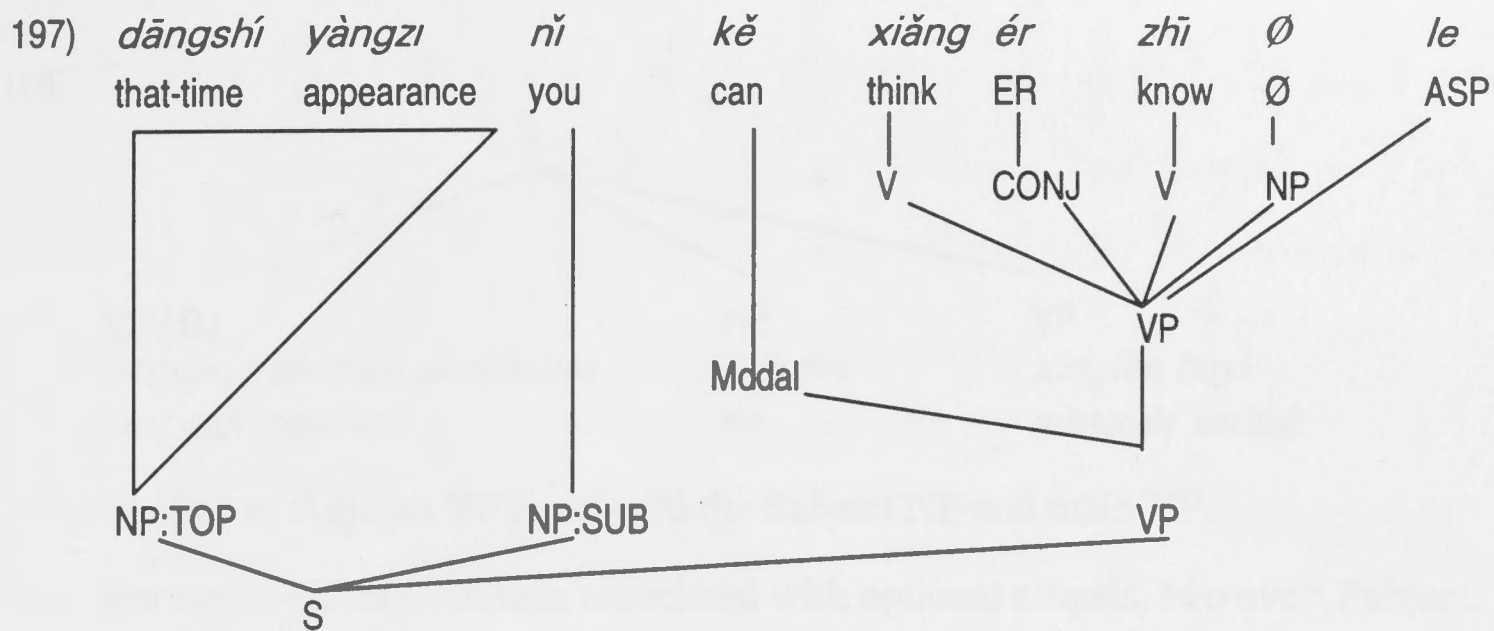
Since *ér* can conjoin sentences, the structure of 195) might be thought to be the same as that for 194), however there are semantic restrictions on the predications which can be conjoined with *ér* and still take two overt Subjects: where the two clauses are in a causal relationship, as in 195), an overt Subject is not permitted in the second clause.

The simplest way to account for this restriction is removal of a Subject position by

VP conjunction:



'[My] appearance at that time, you can consider and know [it].'



I propose that both contrastive and optional ellipsis in conjoined structures also arise from the alternation of two similar structures. There are various indications in support of this proposal. Firstly, conjoined structures are the only ones where optional ellipsis produces bound reference.

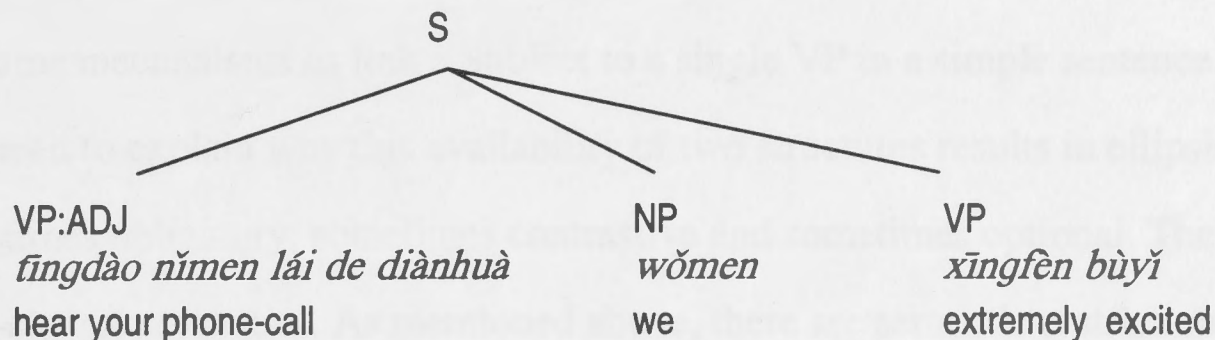
Only 23 out of a total of 86 optionally ellipsed Subjects were in conjoined clauses. Though it's possible that pragmatics might produce a single interpretation, when a consistent relationship holds between anaphor and antecedent, as it does here, it is more likely to be a case of syntactic control.

Secondly, VP conjunction helps account for certain facts about word order in complex sentences.

3.3 Appositions and the Position of the Subject

The division of appositional structures into those involving optional ellipsis and those involving contrastive ellipsis is accompanied by a syntagmatic distinction. For those associated with contrastive ellipsis, when the Subjects of both clauses are coreferent, only one can be expressed overtly, but it can appear in either clause. This corresponds to the selection of a structure with only one Subject position, but where the order of the Subject and the first clause is variable. Precisely this variability of order is possible for temporal and locative adjuncts in Mandarin. I propose the structure:

198)

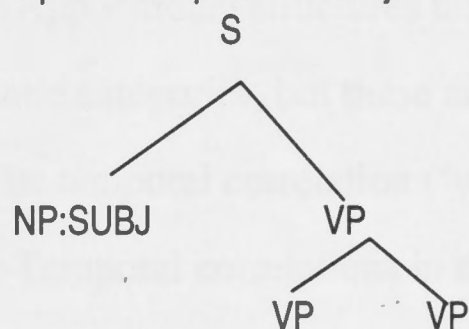


In 198) an Adjunct VP is sister to the Subject NP and main VP.

For appositional structures associated with optional ellipsis, two overt Subjects of any reference are possible. When there is only one, it also controls the reference of both clauses' Subject GFs, but it **must** be realised in the first clause, not the second.

This suggests that these structures consist of VPs conjoined beneath a single node under the Sentence:

199) Optional 'ellipsis': VP conjunction



In fact, providing the referent is retrievable, both Subjects can be ellipsed, as in example 201) above. In this case, where there is no overt Subject at all, co-reference between the two clauses is assumed. In other words, the structure is assumed to be that at 198), not a sentence conjunction. It would not be interpreted as conjoined sentences with **two empty Subject positions**, because referents for both Subject positions would have to be retrievable, and it would be difficult to tell which was which.

In this analysis, apparent co-reference results not from syntactic binding, but from a syntactic structure with only one Subject position. Bi-uniqueness is not violated because the Subject NP bears GFs from two different clauses, just as in functional control. The absence of a structural position in VP conjunction still constitutes ellipsis by the definition used in this work, because there is not a 1-1 correspondence between the GFs assigned by the verbs and lexical NPs.

There is no real problem as to how their reference is interpreted: it is interpreted via the same mechanisms as link a Subject to a single VP in a simple sentence. However, we still need to explain why this availability of two structures results in ellipsis which is sometimes obligatory, sometimes contrastive and sometimes optional. The answer lies in inter-clausal semantics. As mentioned above, there are semantic restrictions on the association between *ér* and obligatory ellipsis. For *ér*, VP conjunction appears to be obligatory only when *ér* conjoins predicates in a causal relationship. Similarly, just a few specific semantic relationships, one of them causal, are associated with **contrastive** ellipsis, regardless of conjunctive device.

3.4 Contrastive Ellipsis - Conditionals, Cause and Correlation

3.4.1 Apposition

Appositional structures that involve contrastive ellipsis fall into three main semantic categories, but these are quite distinct from those associated with optional ellipsis: temporal correlation ('when-then'), conditionals, and causal statements.

Temporal correlations in the corpus were either future tense as in 200), or statements of habitual co-occurrence as in 201) and 202):

200) 以后Ø回国Ø也可以教我。

yǐhòu Ø *húí* *guó* Ø *yě* *kěyǐ* *jiāo* *wǒ*
 after Ø return country Ø also can teach 1sg
 'When you come back, you can teach me.'

201) 上班，和姜丽在一起，

shàng *bān*, *hé* *Jiānglǐ* *zài yìqǐ*
 attend work, with J together,
 '[If/When] [he] comes to work, he's with Jiang Li.'

下班或有时不来，就和妹妹在一起。

xià *bān* *huò yǒu shí* *bù lái*, *jiù* *hé* *Mèimèi* *zài yìqǐ*
 finish work or sometimes not come, then with M together

'[If/ When] [he] finishes work, or sometimes [if/ when] [he] doesn't come, then [he's] with Meimei.'

NS6:13- 14 (2)

202) 没事我们常在一起。

méi shì wǒmen cháng zài yìqǐ

NEG.have matter 1pl often together

'[When] [we] have nothing on, we're often together.'

NS6:9

The generic nature of these statements is evident from the adverbs *yǒushí* 'sometimes', and *cháng* 'often'. The requirement that a coreferent Subject be ellipsed in one clause differs from the situation in English, where this kind of meaning is expressed in a construction where overt co-reference is permissible.

The remaining instances of appositions were causal:

203) 上星期六 Ø 接到你们来的电话，

shàng xīngqīliù Ø jiēdào nǐmen lái-de diànhuà

last Saturday, Ø receive 2pl come-DE phone

'Last Saturday, getting your phone-call,'

听到你们的声音，我们兴奋不已。

tīngdào nǐmen-de shēngyīn, wǒmen xīngfèn bùyǐ

hear 2pl-POSS voice 1pl excited without.end

'hearing your voices, we were extremely excited.'

NS7:1

3.4.2 Overt Conjunctions

One overt conjunction associated with contrastive ellipsis was *jiù* 'then', and this in three sentences only. Just as with appositional structures it is only when *jiù* occurs in contexts where there is a clear causal connection between the two clauses, as in 204), or some kind of habitual correlation between two events, as in 205), that it is associated with contrastive ellipsis.

204) 他很痛苦，就自杀了。

Tā; hěn tòngkǔ, jiù zìshāle

3sg very miserable, then self-kill

'He was very distressed and killed himself.'

NS2: 10-11

205) 我们与他们交流 Ø 就有了困难。

wǒmen yú tāmen jiāoliú Ø jiù yǒule kùnnán

1pl with 3pl communicate Ø then have-ASP difficulty

'[When] we communicate with them we have problems.'

NS7:10c

- 206) Ø 见人就扎。
 Ø jiàn rén jiù zhā
 Ø see people then stab
 '[He'd] no sooner see some-one than stab them.'

When *jiù* was associated with temporal sequence which was not habitual, it was

associated with optional ellipsis, e.g.:

- 207) 他买了一台黑白电视机,
 Tā mǎile yìtái hēibái diànshìjī
 3sg buy-ASP one-CL black.white T.V. set.
 'He bought a black and white T.V. set.'

- Ø 看了两天 Ø 就坏了。
 Ø kànle liǎng tiān Ø jiù huàile
 Ø watch two days Ø then break.down-ASP
 '[He] watched for two days and [it] broke down.'
 NS2

This suggests that the different patterns of co-reference are independent of the meaning of *jiù* and it is the underlying inter-clausal semantics that influences ellipsis in these cases, just as in apposition.

3.4.3 Paired Correlative Conjunctions

Paired correlatives were also associated with contrastive ellipsis. There was only one instance of such a correlative construction in the corpus. This has a meaning similar to that of apposed and conjoined sentences with contrastive ellipsis like 201) and 205), i.e. literally correlative:

- 208) 不是心脏, 就是伤到肺。
 bù shì xīnzàng jiù shì shāng dào fèi
 COR.CONJ heart COR.CONJ wound arrive lungs
 'If it wasn't the heart, it was a stab in the lungs.'
 NS8

(NB: This correlative pair incorporates the Copula, *shì* and assigns a Subject GF only optionally, (cf. the expletive Subject in the English gloss.)) These structures are discussed in detail in Chapter Eight.

3.4.4 Aspectual Subordination

Finally, aspectual subordination also produces contrastive ellipsis. The semantics of the conjunctive suffixes and clitics are also much the same as the semantic implications of the adjunctive VPs discussed under 'Apposition' and under *jiù*, above: temporally

dependent and causal. There were 5 aspectual suffixes in the corpus, *-zhe* indicating simultaneity, and *-le*, anteriority (see Lillian Huang, 1987), as well as the clitics: *X-shí* 'when X', *X hòu* 'after X'; *X-zhīxia* 'under the circumstances described in X / in the light of X'.

For example:

- 209) 我想你看后一定笑我。
wǒ xiǎng nǐ kàn hòu yīdìng xiào wǒ
 1sg think 2sg see after certain laugh 1sg
 'I think, after seeing [it], [you'd] laugh at me, for sure.'
 NS5:13a

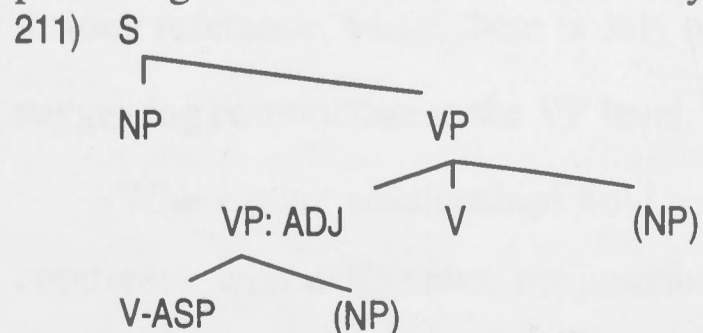
The two post-posed clitics *-le* and *-zhe* are counterparts to the aspectual affixes, *hòu* 'after', and *shí* 'time/when'; the first of each pair indicates anteriority, and the second, simultaneity or overlap.

The clitic *zhī-xià* combines the idea of aspect and causality, expressing the idea that one event occurs as a consequence of and during another event, or the experiential aftermath of it, much like the English construction 'on verb-ing', as in:

- 210) Ø互见之下, Ø确实有一种,
Ø_i hù jiàn zhī-xià, Ø_i quèshí yǒu yīzhǒng
 Ø each-other see POSS-under, Ø really have a.kind
 'On seeing each other, [we] really had a kind of

又惊又喜的感觉。
yòu jīng yòu xǐ de gǎnjué
 both startling both happy DE feeling
 startling and happy feeling.'
 NS4:7a

Of course aspectual markers have the specific function of relating two events in time, and it is not surprising that they should do so by subordinating one verb to another, producing a reduced structure with only one Subject position:



3.5 Optional Ellipsis - Concessives, Sequence and Paraphrase

3.5.1 Apposition

Clauses in the corpus which were **apposed** with optional ellipsis, fell into three semantic classes: a series of actions, paraphrasing, and amplification. For example:
CONJOINED *Bǎ* CLAUSES: A SERIES OF ACTIONS

214) 要把... 衣服晾出去,

Ø *yào bǎ... yīfu liàngchuqu*

Ø must BA... clothes air.out.go

'[I] have to take the clothes and put [them] out to air.'

Ø 把早餐的碗放入洗碗机里。

Ø *bǎ zǎocānde wǎn fàngù xǐwǎnjī-lǐ*

Ø BA breakfast dish put.enter dishwasher-LOC

'Ø take the breakfast dishes, and put them in the dishwasher.'

NS3:17-18

PARAPHRASE

215) 同学还都是老样子Ø没有多大变化。

tóngxué hái dōu shì lǎo yàngzi Ø méi yǒu duō dà biànhuà

class-mates still all are old appearance, Ø NEG have much big change

'The class-mates were all still the same, Ø hadn't changed much.'

NS4:8

AMPLIFICATION

216) 小王调走了、Ø 调回他爸单位去了。

Xiǎo Wáng diào zǒu-le Ø diào huí tā bà dānwèi qù-le

X post.away-ASP Ø post.return 3sg Dad unit go-ASP

'Xiao Wang has been posted away, posted back to his Dad's work unit.'

These are all essentially two clauses representing a single predication, with minor re-statement.

3.5.2 Overt Conjunctions:

Most clauses which were conjoined with overt conjunctions were associated with optional ellipsis. These were illustrated above at 188) to 192). These also involved serial events or actions, or concessive meanings.

3.6 Semantic Motivation for Structure Choice

The correlation between semantics and ellipsis type suggests that for some inter-clausal semantic relationships, such as paraphrase, there is a free choice between sentence conjunction and VP conjunction, while for an interclausal relationship of cause,

correlation or simultaneity VP conjunction is obligatory when only one agent is involved. Concessive relationships can and, in the case of *ér*, must, be expressed by Sentence conjunction. This means that the ultimate explanation of why ellipsis is contrastive, optional or obligatory depends upon an explanation of why speakers choose a certain structure over others. In Chapter Eight I suggest that it is for pragmatic reasons, specifically, to preserve iconicity of syntactic and semantic interdependence, and to reduce ambiguity.

If contrastive and optional ellipsis are simply a consequence of structure choice, why are most conjunctions associated with only one kind of ellipsis or the other? The reasons are two-fold. Firstly some conjunctions can only operate at the highest level of structure, and so, of necessity produce a structure with two Subject positions (see discussion of *dāng* in Aspectual Subordination pg.165) others can operate at several levels, producing more and less reduced structures. Secondly, most conjunctions have semantic content relating to causality, temporality etc.. This restricts their collocation in certain syntactic structures.

3.6.1 *Jiù* Semantically Unspecified

The conjunction *jiù* is the only one in the corpus which can occur with both contrastive and optional ellipsis. This can also be accounted for by lexical semantics: *jiù* is unspecified with respect to causality or other correlations between events: *jiù* invites the listener to attribute causality, and is compatible with a causal interpretation, but it can also be used to conjoin clauses that cannot be considered causally related in any Objective sense, therefore it does not actually **specify** a causal connection. The meaning of *jiù* can be explicated:

217) X *jiù* Y.

Very soon after X, Y happens,
You could think Y happens because of X.

In Chapter Eight it is suggested that it is for semantic and pragmatic reasons, specifically, to preserve iconicity of syntactic and semantic inter-dependence, and reduce ambiguity.

Contrastive ellipsis accounts for only 28 out of 181 instances of ellipsis in the data. All but four were in conjunctive structures; 12 with overt conjunctions, and 12 without.

4. BOUND OBJECT ELLIPSIS

Objects participate in fewer syntactic structures where ellipsis is obligatory, or contrastive. In Chapter Three it was mentioned that James Huang claims that **all** Object ellipsis is essentially of the same character i.e. derived from movement of a post-verbal Object to sentence-initial position. The fronted NP, even in the event that it too has subsequently been ellipsed, binds a trace of the Object NP, via c-command. Despite this claim, \emptyset Objects vary with respect to referential effect and paradigmatic alternatives.

4.1 Ellipsed Objects of SCOMPs - Contrastive?

Chapter Three has a discussion of the possibility, suggested by James Huang (1984) that \emptyset Objects in a speech complement contrast with an overt pronoun:

218) 张三说李四不认识 \emptyset 。

Zhāngsān shuō Lǐsì bù rènshi \emptyset .

Zhangsan says Lisi NEG recognise \emptyset

Zhangsan₁ says Lisi₂ doesn't recognise him_{1/3}

As discussed above (see 55), this structure is somewhat controversial, and it is probably no accident that no such structure appeared in the corpus. It is clear that, for some people at least, the reference of 218) is constrained in some way, but it is not clear that it can be bound syntactically. It is not a candidate for removal by constituent conjunction, since the sentence is not a conjoined one; nor is there any overt NP to bind it, other than the Subject, and Subject binding is ruled out by the DRP. Moreover, the \emptyset Object can refer outside the sentence, so it cannot be bound to a syntactic antecedent.

4.2 Optional Object Ellipsis

18 of the 26 instances of Object ellipsis in the corpus were optional; these have variable interpretation precisely as overt pronouns would. In the next Chapter we will see that their behaviour is precisely parallel to optionally ellipsed Subjects, for which no

method of syntactic control has been proposed (with the possible exception of Anaphoric control).

4.3 Obligatory Object Ellipsis

There were only eight cases of obligatory ellipsis of Objects. Three were gaps in relative clauses, bound by the head; the remaining five cases were in Topic-comment structures.

4.3.1 Topic-Comment Sentences

It is a matter of theoretical perspective whether or not Topic-comment structures are considered to involve ellipsis. The kind of Topic-comment structure relevant here is one with a sentence initial NP which is not coreferent with the Subject, and which is taken as a statement of 'what the sentence is about', i.e. a sentence Topic. In the following discussion a distinction is made between (sentence) Topic, in the sense of a sentence-initial NP, and (discourse) topic, in the sense of an entity or theme whose topicality may be established by a variety of means including presentation as a sentence Topic, fronted Object, Object of a presentative or *bǎ* object.

An example of a Topic-comment structure from the corpus is given below (the antecedent Topic is underlined, the anaphor is bold):

- 219) 当时样子你可想而知了。
dāngshí yàngzi nǐ kě xiǎng ér zhī Ø le
that-time appearance you can think CONJ know Ø ASP
'How I looked at the time, just think ([about it]) and you'll know.'
NS5

In this sentence the Object of the verb *zhī* 'know', and, if we take it as a transitive usage, also of the verb *xiǎng* 'think', is understood to be the sentence Topic. In such sentences as these, the occurrence of an overt Object in post-verbal position is ungrammatical.

4.3.1.1 Moved, Ellipsed, or Never There?

The Topic NP is arguably, outside the comment clause, and, if so, whether we consider the Object has been moved to Topic position, or has been generated, as a

separate GF in sentence-initial position, the comment clause has an empty Object position.

An alternative perspective is that the Topic is not outside the main clause, but receives the Object GF directly from the main verb, i.e. it is an overt Object and the Topic Comment structure is simply a word-order variation. The evidence of coreferent deletion casts some doubt on this alternative.

An Object, in post-verbal position, or an Object fronted with *bǎ* does not control the reference of an empty Object in a conjoined clause:

- 220) 我盖了房子，就烧了。
wǒ_i gài le fángzi_j, Ø_i jiù shaō le Ø_{j/x}
 1sg_i build house_j Ø_i then burn Ø_{j/x}
 'I_i built the house_j then Ø_i burnt it_{j/x}'

- 221) 我把房子盖了，就烧了。
wǒ_i bǎ fángzi_j gài le, Ø_i jiù shaō le Ø_{j/x}
 1sg_i BA house_j build Ø_i then burn Ø_{j/x}
 'I_i built the house_j, then Ø_i burnt it_x'

The Object in the conjoined clause is not *bound* to co-reference with the Object of the preceding clause (unless the VPs are conjoined, which would be indicated by a different prosody), it *could* refer to some other entity, indicated by the subscript 'x'). In a Topic-comment structure though, a Ø Object in both conjoined clauses would be bound to co-reference with the Topic:

- 222) 房子，我盖了就烧了。
Fángzi_j, wǒ_i gài le jiù shaō le Ø_j
 House_j 1sg_i build Ø_i then burn Ø_j
 'The house_j, I_i built then Ø_i burnt Ø_j'

So the Topic does not appear to be simply an Object NP in sentence-initial position, it clearly has additional properties. In fact, though a Topic NP must have some semantic role in the rest of the sentence, it need not be a semantic role associated with a core argument (Li and Thompson, 1981:96). If Topic-comment sentences were simply word order variation, we could not account for topics that have no alternative position in the comment, like Li and Thompson's example (1981:96 e.g. 34):

223) 这场火，幸亏消防队来得快。

zhèi chǎng huǒ, xìngkūi xiāofàngduì láiide kuài
this CL fire, luckily fire-brigade come fast
'This fire, luckily the fire brigade came quickly.'

Clearly topics have some existence independent of the GFs assigned to in the comment. Topic-comment structures, where there is no argument **within the comment** to bear a GF assigned by the verb, are therefore cases of obligatory ellipsis, and the sentence Topic is the preferred choice for controller of a \emptyset Object. Since the semantic role associated with that GF is invariably borne by the Topic, the reference of the \emptyset Object is bound. The question is, what mechanism binds them?

4.3.1.2 The 'Topic-as-Trace' Analysis

In James Huang's (1984) analysis, sentences like 219) have an empty constituent in the Object position, just as with optionally ellipsed Objects. He proposes a single syntactic control mechanism for all \emptyset Objects, i.e. that they are traces bound by a topic which may be empty. This account offers no ready explanation of the variability of referential effect of \emptyset Objects, nor of the variation in paradigmatic alternatives.

Since traces arise as a consequence of movement, a trace cannot be realised; Objects in sentences with Topics also cannot be realised, but Objects in sentences without Topics can.

The status of a null topic is also problematic: it is not a trace, as it has no antecedent; it is not *pro* as it is not governed; and it is not PRO because it can be realised. Thus it does not fit the requirements of any of GB's four empty categories.

4.3.1.3 \emptyset Objects are PRED 'pro'

From the LFG perspective, the missing argument in a comment is unlikely to be under functional control. Functional Control associates one GF with another in a different clause; to do this for all and any arguments would imply alternative functional-control invoking arguments for every argument position of every verb: a kind of automatic functional control. But LFG already has the automatic process of anaphoric control to introduce PRED 'pro' into f-structure in the event of an excess GF. This is more likely to

be the control mechanism in Topic-Comment structures. Additional confirmation comes from the fact that other topicalised NPs can also control \emptyset Objects; the sentence Topic's ability to control the \emptyset Object is a function of its pragmatic status.

4.4 Focus-binding of Objects

4.4.1 Presentatives and Focus: "There is a lot I don't understand."

Another Topic-comment structure in the corpus was that discussed above (see 98):

224) 我有很多看不懂。

wǒ yǒu hě'n duō kànbùdǒng

1sg have very much see.NEG.understand

'There is a lot I don't understand.'

NS5

This \emptyset Object was shown to be bound to co-reference with the Object of the presentative verb in what is potentially a raising or functional control structure (recall that the quantifier *hě'n duō* is an Object NP). Looking at this example in the light of other structures with \emptyset Objects, a parallel can be drawn between control by a sentence Topic and control by the Object of a presentative, which is in a focus position.

4.4.2 Topic and Focus

It is important to realise that topicality and focus position are not mutually exclusive. Lambrecht states (see Lambrecht, 1987: 374):

The *subject* ... is the *unmarked topic* constituent. The *object* NP constituent on the other hand is the *unmarked focus* constituent.... I claim that the focus is to be understood as a formal *scope indicator*, i.e. as a grammatical signal indicating the scope of the assertion expressed by a sentence or proposition. The focus indicates which portions of the sentence are asserted, and which portions are pragmatically presupposed. Thus when the focus is marked on the object or - when no object is present - on some other part of the predicate, this is a grammatical signal that the unmarked relation between presupposition and assertion obtains. A sentence structure in which this unmarked relation holds is a *predicate focus* (PF) structure....When a referent is insufficiently accessible in a discourse,...the unmarked subject-predicate structure or PF structure cannot be used. Instead a *presentational* sentence or *sentence focus structure* is used. The SF structure is a *marked* structure..'

In other words, presentatives and Sentence topics are both structures which shift the focus from its unmarked position in the sentence, i.e. post-verbal Object position, to a position which *includes* a newly established topic.

In simple terms, focusing and topicalisation can be explicated:

225) TOPICALISATION:

I want to tell you something about someone/ something

You have thought about this person before,
not a long time ago
maybe because we have talked about this person/thing before.

- 226) FOCUSsing:
I want to tell you about some-one/something
you don't know who/what I'm thinking of
First I will tell you something
so you will know who/what I want to talk about

These explications show the similarities between placing an entity in focus, and treating them as a topic. Sentence Topics (re-)activate an accessible referent, and establish it within the scope of the predication. An accessible referent is one which has been previously established in the discourse, or is assumed as common knowledge, or is identifiable as a generic class or member of an already established class (see Chafe, 1987: 28-31; Lambrecht, 1987, 376-7; Wu, 1992).

For instance in the corpus, one writer writes that she must 'put the breakfast dishes in the dishwasher, then start the machine'. She then uses a topic-comment structure to refer to a sub-set of the dishes:

- 227) 不能在洗碗机里洗的就用手洗。
[_{top} *bù néng zài xǐwǎnjīlǐ xǐ* Ø *de*] Ø, *jiù yòng shǒu xǐ* Ø
NEG able in dishwasher wash Ø DE Ø, then use hands wash Ø
[Those] which Ø can't be washed in the dish-washer, [I] wash Ø by hand
NS3: 24c

The head itself is already so accessible that it need not be overtly represented, but the NP as a whole is being placed in topic position, and at the same time is in the scope of focus.

A presentative introduces a **new** participant and places it within the scope of the predication, for example the writer relating the story of the old man who bought a faulty t.v. begins with a presentative whose Subject is a place-name:

- 228) 山东有个农民老人,
Shāndōng yǒu ge nóngmín lǎorén
S. has CL 'peasant' old-person
'There was an old 'peasant' in Shandong...'

Note that this clause is, pragmatically speaking, incomplete. It goes without saying that there are old men in Shandong, so the presentative verb with its two arguments does

not satisfy pragmatic requirements that each discourse unit have a topic **and** a predication about the topic; though it is post-verbal, the NP *laǒrén* does not constitute a newsworthy predication about the place *Shandong*. Lambrecht (1987:375) cites this as a characteristic of SF sentences.

He also states (1987:375):

It is important to understand that the notion 'sentence focus structure' defines a grammatical construction in which the *subject* is not a topic...

The presentative verb *yǒu* links a newly introduced participant to an entity of general accessibility (a place or time), but low inherent interest. It can do this because it need not assign a Subject GF; a place-name, or a temporal NP are often used in its Subject position. Since the Subjects of the presentative tend to be generally accessible, and of low interest-value, the Object of a presentative is essentially the first NP in the sentence which is likely to continue as a topic of later discourse; a presentative verb establishes a new participant as a topic, much as a Topic-comment structure does, by placing it in focus. Thus a discourse topic, and a sentence topic, may both be in focus: there is not necessarily a dichotomy of topic or focus.

Since sentence-initial position is where an established topic is re-introduced and highlighted, sentence Topics are obvious candidates for control of later \emptyset elements. Similarly, since post-verbal position is where new participants are introduced, a post-verbal NP which is of higher inherent interest (e.g. human) than any preceding NP in the sentence is also an obvious candidate for control of later \emptyset elements.

This can be stated as a principle:

229) Apart from relativised constituents, all ellipsed Objects of bound reference in the corpus are bound to refer to an entity whose topicality has been established within the same sentence, by placing it in focus.

In the absence of a sentence Topic, optionally ellipsed Objects are controlled by a discourse topic, which may be established by a number of means, but which is almost invariably in the scope of focus.

In presentatives, the semantic-pragmatic implications of introducing a new participant and at the same time, the activity in which it is involved, are expressed by way of functional control of an XCOMP (i.e. syntactic control). As with other dependent

clauses, the relationship between the controlling antecedent and its controllee is very constrained. The relationship between a sentence Topic and its controllee on the other hand, is very variable. A sentence Topic need not have a specific GF in the comment, so long as it has some feasible relationship to the predication.

4.4.3 Conclusions

Sentence-initial Topics are primarily a pragmatic device. They represent the most topical, **and focal** entity at that point in discourse, and they can refer to NPs in an indefinite variety of semantic relationships to the comment clause. Objects of presentatives are also topical and within the scope of focus. In a hierarchy of retrievable antecedents, both Topic NPs and Objects of presentatives therefore rank very high. Moreover, all Objects in canonical word order are in the unmarked focal position, this means that control of \emptyset Objects can be generalised in terms of focal control: a marked focal NP controls an empty unmarked focal position. The reference of all ellipsed Objects can indeed be seen as determined by a single process, as James Huang suggests; but it is a pragmatic rather than a syntactic one.

5. CHAPTER SUMMARY: THE BOUNDARIES OF SYNTACTIC CONTROL

In this chapter we have looked at the mechanisms controlling reference in cases of bound ellipsis other than obligatory Subject ellipsis. This included obligatory Object ellipsis in Topic-Comment structures, the ellipsis of relativised constituents, the ellipsis of the Subjects of VPs in NP positions, bound optional ellipsis in Copula constructions, and contrastive ellipsis in conjoined constructions. Structural positions where ellipsis was found to be under syntactic control included:

\emptyset Subjects of VPs in Object and Copula complement positions, and relativised constituents.

Constructions where ellipsis was found to be under pragmatic control included:

\emptyset Objects in Topic-comment structures; \emptyset Subjects of VPs in Subject position; the Heads of NPs in copula sentences.

5.1 Alternation of Structures

It has been proposed that certain structures exhibiting ellipsis result from syntactic constituent conjunction which removes a structural position. This is true regardless of whether the reduced structure can alternate freely with an extended structure (sentence-conjunction) or can alternate with such structures only when they have certain referential properties (contrastive ellipsis), or only when they have certain inter-clausal semantics (obligatory ellipsis). Clauses are combined in different ways, and this may affect the relative position of Subjects and conjunctions, or clauses, but it does not, of itself, affect the patterns of ellipsis and co-reference. The choice of VP conjunction, and therefore obligatory or contrastive ellipsis, to express interclausal relationships of cause and correlation appears almost mandatory, while other interclausal semantics allow free variation between structural alternatives. In either event, the elliptic structures involve syntactic mechanisms producing bound reference.

5.2 Frequency of Syntactic vs Pragmatic Control

Tables 6 and 7 in Chapter Two presented the frequency with which different structures were represented in the discourse. It is now possible to describe these structures in terms of syntactic vs pragmatic control. The evidence presented above shows that variable ellipsis can be associated with pragmatic control. In addition, \emptyset Objects in Topic-comment structures; \emptyset Subjects of VPs in Subject position; the Heads of NPs in copula sentences are also pragmatically controlled. The frequencies of pragmatically and syntactically controlled structures are presented in Table 25.

Relativised constituents, \emptyset Subjects of VPs in Object and Copula complement positions, and obligatorily ellipsed Subjects of verb complements are under syntactic control. 'Contrastive' ellipsis is ellipsis which, like that in verb complements, results from the sharing of a single Subject argument by two clauses, via conjunction of constituents. Therefore it can also be classified as syntactically determined, though semantics and pragmatics play a central role in the selection of a specific structure.

Table 25: Frequency of Pragmatic and Syntactic Control in NP ellipsis

CONTROL TYPE:	GF ellipsed	STRUCTURAL CONTEXT	Frequency
PRAGMATIC CONTROL	Subject (indefinite)	independent and relative clauses	10
	Subject (definite)	independent clause	30
		tensed SCOMP	5
		VP in SUBJ position	1
	Subject	Total	46
	Object (indefinite)	independent clause	3
	Object (definite)	independent clause	9
		conjoined clauses (sequential)	1
		apposed clauses (paraphrase)	1
	Total Object ellipsis		14
OBJ2 ellipsis	definite	1	
TOTAL		61	
SYNTACTIC CONTROL	Subject (definite)	VP in OBJ position	2
		Heads of Rel Cl. in Copula construction	2
		conjoined verbs (obligatory)	2
		conjoined clauses (contrastive)	12
		apposition (contrastive)	12
		conjoined clauses (optional)	18
		apposed clauses (optional)	5
		Nominal control	2
		XCOMPs	57
		NP rel	2
	Total Subjects		114
Object ellipsis	Topic comment	5	
	NPreI	3	
Total Objects		8	
TOTAL		122	

The proportion of NP ellipsis which is syntactically determined is about 2/3 of the total instances in the corpus (122/183). However, it should be remembered that lexical semantics also play an important role in the selection or reduced structures, as well as in the interpretation of \emptyset elements in correlative and other conjoined structures, so, though they are primary in only 1/3 of the instances of NP ellipsis, semantic and pragmatic factors are clearly important in the interpretation of about 2/3 of the cases. Understanding the details of how pragmatic control functions to select an antecedent is therefore vital to a full explanation of elliptic reference in Mandarin informal written discourse.

Chapter Seven

Pragmatic constraints on Optional Variable Ellipsis

1. INTRODUCTION: PRAGMATICS, DISCOURSE STRUCTURE AND INTERPRETATION

So far, I have accounted for interpretation in cases where ellipsis is bound and/or syntactically determined. This chapter explores how pragmatic, discourse structural and semantic factors contribute to the interpretation of ellipsis of variable reference. Pragmatic factors include knowledge of the world and cultural conventions, which make some interpretations more likely than others; the status of participants in both the linguistic and extra-linguistic context as accessible, focal, thematic, new, and so on. Discourse structural factors refers to formal aspects of a constructed text, including paralleling surface form in adjacent sentences; the use of sentence conjunctions and formulaic conventions to signal relationships between sub-parts within a text, such as openings, and closings, topic change, topic development, re-introduction of a previous topic etc.; the use of specific structural positions, or GFs to signal the pragmatic status of participants. Semantic factors refers to detail of what participants, events and contextual details are specifically encoded in the meaning of lexical items, and therefore contribute to logical and semantic cohesion.

2. ZERO PRONOMINALS AND ACCESSIBILITY

Givon (1984) proposes an accessibility hierarchy whereby the more readily accessible a referent is, the less morphological material is used to encode it. Referential Distance is one of the major measures of accessibility proposed by Givon (1984). Givon states (1984: 131) that zero 'anaphors' (including, both anaphors and pronominals in the GB sense) in English have an average referential distance of one clause.

Since true anaphors (in the GB sense) are bound within their governing category, they could not usually have a referential distance greater than one clause. But there is no syntactic reason why \emptyset pronominals (i.e. the optional variable \emptyset found extensively in Mandarin) should be so constrained. And indeed, it has been observed (Givon, 1984; Williams, 1989) that in the Non-native English of 'pro-drop' language speakers, and in pro-drop languages themselves the average referential distance is considerably longer.

Thus, not only is pronominal ellipsis unconstrained by a syntactic relationship to its antecedent, the antecedent may well be relatively far removed in discourse. How then is the reference of pronominal ellipsis uniquely determined?

3. *OPTIONAL SUBJECT ELLIPSIS OF VARIABLE REFERENCE*

It is sometimes claimed that semantic compatibility accounts for pragmatic interpretation, but this is unlikely to be the sole or primary cue to reference. Though a great many of the verbs in the corpus with ellipsed human Subjects require an animate agent, some do not, and conversely some of those with ellipsed inanimate Subjects would allow animate ones. This means that a reliance on semantics alone would not uniquely identify the intended antecedent in all cases.

3.1 Subject-Subject Co-reference

The hypothesis proposed here is that Subjects are expected to continue as Subjects in the succeeding discourse, and that diversions from this pattern will be given special treatment to facilitate interpretation. In the corpus, there are 59 optionally ellipsed Subjects of definite reference in the corpus. The patterns of co-reference are set out in Table 26. A quarter (15/59) of the optionally ellipsed Subjects of variable reference had no overt lexical antecedent; all but two of these were 1st person; one other was 2nd person (imperative), and the third a reference to the weather (\emptyset *wǎnde zǎo* - '[it] gets late early (sic)'). 1st and 2nd person \emptyset Subjects tend not to have overt textual antecedents because the speaker/writer is naturally continuously present in the non-linguistic context. Two more ellipsed Subjects referred to a Subject in the next (second) clause previous. Of

the remainder with overt antecedents, only three referred to Objects, and one to a Topic NP. One more was ambiguous between Subject and Object reference.

Table 26. Reference of Optionally Ellipsed Subjects

Antecedent	Ø Subject		Total
	Human	Inanimate	
None	15	1	16
Subj - 1	33	3	36
Subj - 2	2	-	2
Obj - 1		3	3
Topic		1	1
Ambiguous			1
Total	50	8	59

Subj - 1 means the next previous Subject, Subj - 2 means a previous Subject further than the closest.

Frequencies were too low for a χ^2 analysis, so Fisher's Exact Test was used and the differences were found to be statistically significant ($p=.000009$; see Table 27).

Table 27 GF of Antecedents for Subjects and Objects

Antecedent	Anaphor		Tot
	Subject	Object	
Subj	38 (E=32.45)	2 (E=7.547)	40
Other	4 (E=10.547)	8 (E=2.452)	12
Total	42	10	52

The general trend (36/59) is for Subjects to refer to the next previous Subject. The Subject antecedent need not be overt, the next example shows co-reference through a string of PRED 'pro' Subjects; the second shows an example where the antecedent is a functionally controlled position, that is it is assigned in f-structure by functional control, but is not an item present at c-structure. (The antecedent is underlined, the relevant Ø is in bold):

230) 小王调走了、Ø 调回他爸单位去了。

Xiǎowáng diàozǒu-le, diàohuí tā bà dānwèi qù le
 X post.away-ASP, post.return 3sg dad unit go. ASP
 'X_i has been posted elsewhere, posted back to his dad's unit.'

有时 Ø 有时间 Ø 还到车间来找我们。

yǒushí **Ø** yǒu shíjiān **Ø** hái dào chējiān lái zhǎo wǒmen
 sometimes Ø have time Ø still to assembly room come seek 3pl
 'Sometimes, [if he] has time, [he] still comes to the assembly room looking for us.'

NS6:5-7

231) 讨论出一个道理来 叫作

[VP *tǎolùnchū* [NP:OBJ *yīgè dàolǐ*] [_{xcomp<OBJ>} PRO *lái*] [_{xcomp<subj>} PRO *jiàozuò*]
 discuss.out a truism PRO come PRO call-be
 [We] brought out through discussion a truism PRO being

中国话好说，中国字难写。

[_{scomp} "zhōngguó huà hǎoshuō, zhōngguó zì nán xiě"]]]
 Chinese talk good-speak, Chinese characters difficult write
 "Chinese is easy to speak but hard to write"

NS7:16

In 231) the understood Subject of *jiàozuò* 'call.be', is clearly *dàolǐ* 'truism'. The first and only overt occurrence of *dàolǐ* 'truism' is as an Object NP (of *taǒlùn.chū* 'discuss.forth'). However, it is not this Object which controls the \emptyset Subject of *jiàozuò* 'call.be'. The GF nearest to the \emptyset Subject of the verb *jiàozuò* 'call.be', and also associated with the 'entity' *dàolǐ* 'truism' is actually the Subject of *lái* 'come'. This is because the compound *chū-lái* 'produce-come' is an XCOMP<OBJ> structure: *lái* 'come' is the dependent complement of *chū*, and cannot take an overt Subject. Its Subject GF is controlled by the Object of *chū*, i.e. *dàolǐ* 'truism', and it is this \emptyset Subject of *lái* 'come' which controls the final \emptyset Subject of *jiàozuò* 'call.be'. Analysed in these terms, 231) is still an example of Subject-Subject co-reference.

3.2 Exceptions to Subject-Subject Co-reference

The first point that stands out is that \emptyset human Subjects always have antecedents in Subject GF; the Subject with a topic for antecedent, and all 3 of the Subjects with an Object for antecedent were inanimate Subjects (see Table 28).

Table 28 GF of Antecedent for Human vs Inanimate \emptyset Subjects

Antecedent	Anaphor		Tot
	Human	inanimate	
Subj	35 (E= 31.8140)	3 (E=6.186)	38
Other	- (E = 4.186)	4 (E=.814)	4
Total	35	7	42

Results of Fisher's Exact Test show that human Subjects are significantly more likely to refer to a previous Subject than inanimate Subjects ($p=.000313$). This is primarily a consequence of the fact that humans tend to be encoded as Subjects, and

inanimates as Objects (Du Bois, 1987). Of course this pattern of co-reference cannot be a tool in the interpretation of reference; the listener-reader does not know whether an ellipsed GF refers to a human or animate entity until they know which entity it refers to. What it does mean is that when a \emptyset Subject does refer to an inanimate entity, the general norm of Subject-Subject co-reference may not apply. Because of this, in each case where an inanimate entity does continue in Subject role, it is foregrounded in some way, establishing or increasing its topicality by way of additional syntactic and discourse features. This makes it more retrievable when a \emptyset Subject reference follows.

3.2.1 Foregrounding of Antecedent

3.2.1.1 Topicalisation

One foregrounding strategy used when the antecedent of a \emptyset Subject was something other than a human Subject, was direct topicalisation of the antecedent. This was achieved in one of two ways: by introduction of the antecedent as object of the presentative verb *yǒu*, as in 232) or by placement of the antecedent in sentence initial position as a Topic, as in 233). The former accounts for two instances in a single sentence; the latter for one instance. The relevant \emptyset arguments are those in bold, the antecedent is underlined.

Antecedent = Object of Presentative

232) 心中有一种失望的感觉。

xīnzhōng yǒu yīzhǒng shīwàng de gǎnjué
heart centre have one.kind disappoint DE feeling

\emptyset 不知 \emptyset 从何而来, 反正 \emptyset 淡淡的。

\emptyset *bù zhī* \emptyset *cóng hé ér lái* fǎnzhèng \emptyset *dàndànde*
 \emptyset NEG know \emptyset from where ER come anyway \emptyset faint.RED

In [my] heart there was a kind of disappointed feeling. \emptyset coming from [I] don't know where, but anyway \emptyset faint as faint.'

NS4:10-11

Antecedent = Sentence Topic

233) 这个住处原来住的一个中国学生,

zhège zhùchù yuánlái zhùde yīge zhōngguó xuésheng,
this dwelling, originally house one.CL Chinese student

他正好要回国 Ø 就建议我来了。

tā zhènghǎo yào húi guó Ø jiù jiànyì wǒ lái-le
3sg co-incidentally must return country Ø then suggest 1sg come-ASP

我最满意的是离梅西很近，

wǒ zuì mǎnyìde shì Ø lí méixī hěn jìn,
1sg most satisfied.DE COP Ø removed M. very close

'This dwelling, originally it housed a Chinese student. By a happy coincidence, he had to go home just at that time, and suggested I come [here]. What pleases me most is that [it] is near M.'

NS3:6-9

3.2.1.2 Individuation

A second strategy for foregrounding an inanimate Object is to individuate it via a numeral and/or classifier. This marks a post-verbal NP as definite and specific, a pre-verbal NP is almost always definite in any case (see Sun and Givon, 1985).

The sentence of ambiguous structure mentioned above (pg. 182) is one where an Object NP is individuated:

234) 目前我正在写本书，

mùqián wǒ₁ zhèngzài xiě běnshū₂
Presently 1sg right-now write CL-book
'At the moment I'm in the process of writing a book'

题目是汉语熟语与中国的人文世界。

tímù shì "Hànyǔ shúyǔ yǔ zhōngguó-de rénwén shìjiè
title COP "Mandarin idioms and China-POSS literary world"
'The title is "Mandarin idioms and China's literary world".'

Ø 三年完成 Ø。

Ø sān nián wánchéng (Ø)

Ø, three years finish (Ø)

'[It] will take three years to finish'

OR Ø₁ three years finish Ø₂

'[I] will finish [it] in three years.'

NS1: 5-7

In 234) the reference of the Ø subject has ambiguous reference, without producing any alteration to the meaning conveyed. This is because the verb could be in active voice, with patient₁ - patient₂ co-reference, mapping onto object₁-object₂ co-reference, or it could be in middle voice, with patient₁ - patient₂ co-reference, mapping onto object₁-subject₂ co-reference. If the latter is assumed, then it can be argued that the inanimate topic 'a book' has been both individuated, and amplified via a partitive reference in the

intervening clause which gives its title, thus establishing 'the book' as a continuous discourse topic.

The fact that patient-patient co-reference obtains, regardless of GF, suggests that it is the prototypical association of human with agent with Subject, and inanimate with patient with Object which produces the outward correspondence of Subject-Subject co-reference. This would explain why the correlation is not absolute, and why the exceptions occur in relation to inanimate \emptyset Subjects.

3.3 Backgrounding of Intervening Topics by Ellipsis

3.3.1 Ellipsis of Human Subjects

In many cases when a \emptyset Subject (human or inanimate) refers to something other than the closest previous Subject, ellipsis is also used as a strategy to remove NPs from any intervening positions, particularly NPs referring to human participants, which might otherwise attract focus.

In this next example, the \emptyset Subject is in the last line, ' \emptyset was forced to alight'.

A syntactic passive was used (one of only two in the corpus) to remove the other human **agent** (the guard) as a potential antecedent.

235) 他去找 \emptyset 了一趟, \emptyset 没找着 \emptyset 。

Tā qù zhǎo \emptyset le yītàng \emptyset méi zhǎozháo \emptyset

3sg go search \emptyset ASP one-trip \emptyset NEG find-succeed \emptyset

'He went looking [for her] once, without success.'

\emptyset 挤火车 \emptyset 回家时, 因火车严重超载,

[_{adj} \emptyset jǐ huǒchē \emptyset huí jiā shí] [_{adj}yīn huǒchē yánzhòng chāozài]

[_{adj} \emptyset cram train \emptyset return home time [_{adj}because train serious over-load]

'Cramming on the train to come home, because the train was seriously over-loaded,

\emptyset 在十堰被警察赶下车了。

\emptyset [_{vp} zài Shíyàn bèi jǐngchá gǎn xià chē le]

\emptyset [VP at Shiyan by guard drive descend train ASP]

'[he] was forced by the guard to get off the train at Shiyan.'

NS8:14

If the sentence had been active, with the guard as Subject, that agentive Subject would have been assumed to be the antecedent. With passivisation, the chain of reference could be continued, even with a \emptyset reference sustaining the topicality of the main participant.

Another example of backgrounding involved an inanimate Ø Subject referring past a human Subject. In this case intervening Subject GFs referring to the inherently topical human participant have all been ellipsed.

- 236) 山东有个农民老人,
Shāndōng yǒu ge nóngmín lǎorén
 S. has CL 'peasant' old-person
 'There was an old 'peasant' in Shandong...
 用自己节省了十几年的钱
 Ø *yòng zìjǐ jiéshěng le shíjǐ nián-de qián*
 Ø use self save-ASP ten-some years-DE money
 ...[who] used the money he'd saved over several decades,...
 Ø 买了一台黑白电视机。
 Ø *mǎile yìtái hēibái diànshìjī.*
 Ø buy-ASP a-CL black-white t.v.
 ...to buy a black-and-white t.v.'
 没想到, Ø 回家 Ø 看了 Ø 两天
méi xiǎngdào Ø huí jiā Ø kànle Ø liǎngtiān
 NEG+ASP think-ASP Ø return home Ø watch-ASP Ø two-days
 ...[who'd] have thought: [after] returning home and watching [it] for two days,...
 Ø 就坏了。
 Ø *jiù huàile.*
 Ø then break-down-ASP
 ...[it] broke down'
 NS 2:2-6

In this example, the ellipsed Subject in question is that of the verb *huài* 'to break-down' or 'become bad'. It refers to a 't.v. set' mentioned four clauses back. There is one overt Object and three Ø argument positions between the anaphor and a previous coreferent NP. The nearest overt NP to the Ø Subject precedes the Ø Subject by four clauses: the locative Object *jiā* 'home'. If this were topical, it would have been individuated and probably introduced as a Subject.

Before I can exclude the intervening Ø arguments as antecedents, some clarification of Mandarin semantic restrictions on collocations is necessary. One of the intervening Ø NPs, is the Object of *kàn* 'watch'. It does refer to the same television as the Ø subject in question does, but its status as a lexical antecedent is slightly problematic. Mandarin distinguishes between the concrete thing *diànshìjī* 'tv set', and the

visual phenomenon *diànshì* 'tv'⁶; the phrase *kàn diànshì* 'watch tv' is acceptable, but the phrase *kàn diànshìjī* 'watch the tv set' is not. On the other hand, the phrase *diànshì huàile* 'tv's broken down' is not acceptable, while the phrase *diànshìjī huàile* 'the tv's broken down' is. In other words, lexical choices for the Subject of *huàile* 'broken down' and the Object of *kàn* 'watch' are in complementary distribution.

Now, in 236) one of the \emptyset s preceding the \emptyset Subject of *huàile* 'broke down', in the last line, is the \emptyset Object of *kàn* 'watch'. It **refers** to the tv set, but for the reasons just stated, it **cannot be a lexical antecedent** for the \emptyset Subject of *huàile*. This means that if the Subject of *huàile* 'break-down' in 246) is deemed to have a structural antecedent, that antecedent would have to be the overt Object NP *diànshìjī* mentioned 4 clauses back. That NP is an overt and individuated Object: the most recently introduced topical NP **in focus** at the time the \emptyset Subject occurs.

The other two intervening \emptyset NP positions in 236), underlined with dots, refer to the old-man who bought the t.v. They are both Subjects referring to a human, who is the main discourse topic. If either of these were overtly expressed, they, and not the previous Object would be understood to control the \emptyset Subject in question, even though the interpretation would be absurd. This is because the aspect marked '*huàile*' 'broke down' is usually only applied to inanimates. If either of the intervening references to the old man were overt, the \emptyset Subject of *huàile* 'broke down' could only refer to the overt Subject pronoun, making the sentences humorous or nonsensical:

237) 他 回家 \emptyset 看了 \emptyset 两天 \emptyset 就坏了。

tā húi jiā \emptyset kànle \emptyset liǎng tiān \emptyset jiù huàile
 3sg return home \emptyset watch-ASP \emptyset two days \emptyset then bad-ASP.
 'He returned home, watched [it] for two days, and 'broke down'.'

238) 回了家 他看了 \emptyset 两天 \emptyset 就坏了。

húile jiā tā kànle \emptyset liǎngtiān \emptyset jiù huàile
 return-ASP home 3sg watch-ASP \emptyset two days \emptyset then bad-ASP.
 '[Having] returned home, he watched [it] for two days, and 'broke down'.'

⁶ Thanks to Picus Ding for reminding me of this.

Note that the use of ellipsis here cannot arise from removal of an NP position by constituent conjunction because the antecedent is in a different sentence. The difference in reference arises depending on whether the structural position is filled or empty. This means that overt representation must endow an NP with some kind of pragmatic status, foregrounding it as an entity in focus.

To summarise, 236) exemplifies a number of foregrounding and backgrounding strategies employed when the norm of co-reference between a \emptyset Subject and the nearest preceding Subject is contravened. Firstly, any overt reference to the inherently topical old man, intervening between the overt mention of the tv set and the final \emptyset Subject would break the intended chain of reference. **Inherently topical NPs which are not the current focus are expressed by \emptyset .**

Secondly **the use of an overt pronoun emerges as a strategy for placing a topical entity in focus.**

Thirdly, and conversely, **ellipsis maintains the topicality of an entity without bringing it into focus.**

And fourthly, an intervening overt Object could break the chain of reference, but the Object is a) not individuated, b) not inherently topical, and c) semantically incompatible. **Inanimate entities in Object position are brought into focus by individuating them with a classifier.**

The foregrounding of the intended antecedent i.e. the **focal** topic, and backgrounding of all the intervening references is clearly crucial to maintaining the chain of reference in a text; where an inherently non-topical entity like an inanimate object, is introduced into discourse as a topic, it is outside the norms of Subject-Subject co-reference exhibited by Human participants which are prototypically agentive and topical. Because of this, it is fore-grounded by topicalisation and/or individuation. At the same time ellipsis is employed to suppress mention of inherently topical entities, like humans, when they intervene between an anaphor and such an antecedent. This bears out the

hypothesis that Subjects are expected to continue as Subjects, and that special treatment is used to highlight the antecedent when this pattern is broken.

4. OPTIONAL OBJECT ELLIPSIS

A very similar situation emerges with respect to optional Object ellipsis. There were three instances of transitive verbs with no Objects which were most readily interpreted as examples of intransitive valency options or indefinite ellipsis.

4.1 Indefinite Ellipsis

The following appears to be an instance of a generic or non-referential Object.

239) Ø下课时天已经很黑了，

Ø xià kè shí tiān yǐjīng hěn hēi le,

Ø finish class time sky already COP black-ASP

'When class gets out, it's already very dark,...'

Ø住了离梅西近些是必要的。

Ø zhù-le lí Méixī jìn xiē shì bìyàode.

Ø live-ASP separate M. close some COP essential

...[it]'s essential to live closer to Massey.'

另外我住在这里，要帮助Ø作家务。

língwài wǒ zhù zài zhèlǐ yào bāngzhù Ø zuò jiāwù

also 1sg live at here must help Ø do housework

'Also, living here, I have to help do the housework.'

NS3 : 14-16

The last line does not specify **who** the writer must help to do the housework.

Though 'help' is commonly used as an intransitive matrix verb in English, it is unusual for *bāngzhù* to be used this way in Mandarin, typically a related intransitive lexical item, *bāngmáng*, is used. The last mentioned (semantically) probable coreferent for the ellipsed Object in 239) appears 13 main clauses earlier, where the family with whom the writer lives is first introduced. This reference is not a very good candidate for a lexical antecedent, because it included a one-year-old boy, unlikely to be involved as an agent of *zuò jiāwù* 'doing the housework' and therefore also as beneficiary (i.e. Ø Object) of the verb *bāngzhù* 'help'. Moreover explicit expression of a controlling Object of *bāngzhù* 'help' functioning also as Subject of *zuò jiāwù* 'do housework' would imply a somewhat

different social contract, where the person/people represented by the pivotal Object-Subject have primary responsibility for the action expressed by the second verb, and the person represented by the Subject of the first verb *bāngzhù* 'help' only a minor role. In fact the actual situation described in the text is the reverse of this. So this example is best understood as an indefinite reference for what is usually a controlling Object, similar to the English usage.

4.2 Object-Object Co-Reference

Besides a few such instances of intransitive usage, there were 11 instances of optional ellipsis of definite Objects. Seven were unambiguously coreferent with the nearest previous Object NP, symmetrical to the proposed norm of Subject - Subject co-reference. The following shows the most typical circumstance under which a definite Object NP is ellipsed:

240) 刀该满好 Ø 见人就扎 Ø。

dāo gāi mǎn hǎo Ø jiàn rén Ø jiù zhā Ø₀₂

knife must full good Ø see person Ø then stab Ø

It must have been a good knife, [he] just saw someone and Ø stabbed Ø.

Ø不是心脏 Ø 就是 Ø 伤 Ø 到肺。

Ø bú shì xīnzàng jiù shì Ø shāng Ø₀₃ dào fèi

Ø NEG COP heart then COP Ø injure Ø to lung

'If [it] wasn't the heart, then [he] got [them] in the lungs.'

NS8:7

The NP *rén* is polysemous in Mandarin, and in this case is comparable to the English indefinite and non-specific 'any-one'. In 240 *rén* appears first as an Object NP; the Ø₀₂ patient of the next verb *zhā* refers back to that Object. Note that while the noun *rén* refers to 'anyone', the Ø₀₂ does refer not to 'anyone' but to 'anyone who was actually "seen"', i.e. the Object of *jiàn*. The next Ø Object in the text, Ø₀₃, refers to 'anyone who was seen **and stabbed**' i.e. the entity represented by Ø₀₂. In other words, each Ø refers to the previous Object, not directly to the lexical antecedent *rén*. This is demonstrated by the following example. In 240a) the word *dāo* 'knife' stands in the same syntactic position as *rén* did in 240), but the co-reference pattern is different. Here, the final Ø refers to a Sentence Topic, not to 'the knife'

240a) 那些人, 流氓见刀 就扎。

nèixiē rén_i, liúmáng_j jiàn dāo_{k o1} Ø jiù zhā (Ø_i)_{o2}

those-few people, hooligan see knife Ø then stab Ø

'As to those people, the hooligan(s) saw a knife, and stabbed [them],'

Ø不是心脏 Ø 就是 Ø 伤 Ø 到肺。

bú shì xīnzàng jiù shì (Ø_i) shāng (Ø_i)_{o3} dào fèi

NEG COP heart then COP Ø injure Ø to lung

'If [it] wasn't the heart, then [he] got [them] in the lungs.'

Firstly, a Topic NP takes precedence over the nearest previous Object as controller of reference, because it is more foregrounded; this results in the Ø Object, Ø_{o2}, being bound to co-reference with the topic, not with the lexical Object *dāo* 'knife'. Secondly, Ø_{o3} is still coreferent with Ø_{o2}, and **it** is now coreferent with the Topic NP as well, not with *dāo* 'knife'. Note also that in both versions another NP intervenes between Ø_{o3} and its antecedent, but this NP *xīnzàng* 'heart' is a syntactic Copula complement, not an Object. The equative meaning of the Copula apparently does not serve to establish its complement as a likely antecedent for later anaphors. This needs further investigation into the discourse function of the copula in Mandarin; examples in my corpus were too few.

Of the remaining Ø Objects, two refer to the nearest previous Subject, and one appears to have no antecedent at all.

4.3 Exceptions to Object-Object Co-reference

4.3.1 Object-Subject Co-reference, No Suitable Object

In one instance of Object-Subject co-reference, there is no Object in the preceding sentence, and the Ø Object co-refers with the nearest previous semantic patient (a passive Subject). The context of this passage is as part of an account of why a peasant went berserk at Shiyan railway station, stabbing 9 people. In this example, the agent of the wife's abduction is unknown, and unimportant in what follows; a passive construction is used, and the patientive wife is represented by a Subject NP because it is she who is the topic of this episode within the larger theme. The Ø Object in the second sentence refers back to the most topical entity **at that point**: the Subject of the previous sentence.

241) 老婆被人骗到四川去。

Lǎopo bèi rén piàn dào Sìchuān.

Old-Lady by someone cheat to Sichuan.

'[His] old lady was stolen away by some-one to Sichuan.'

他去找 \emptyset 了一趟, \emptyset 没找着 \emptyset 。

Tā qù zhǎo \emptyset le yītàng \emptyset měi zhǎozháo \emptyset

3sg go seek \emptyset ASP one.time \emptyset NEG seek.success \emptyset

'He went to look for [her], didn't find [her].'

NS8:11-13

A few words about 'topicality' are pertinent. Topicality is sometimes equated with retrievability, and Givon (1984) predicts that the more retrievable a reference is, the less phonetic material will be used to encode it. By this measure, we might think that the pronominal NP *tā* '3sg' which is Subject of the second sentence in 241) is more retrievable, and therefore more topical, than the full NP *lǎopó*, and therefore a more likely antecedent for the \emptyset Object. There are two reasons why this mapping should not arise. One is simply that same-clause co-arguments are preferred to be of disjoint reference (the DRP), and the second is that topicality is of a complex nature.

While one entity may be the main topic of a text, other entities may be the main topic of sections within that text, and still others may be the topic of specific sentences within the section. Since there are multiple topics operating at different levels within the text, it is clearly not sufficient to talk in simple terms of 'the' topic. What happens in 241) is that the 'husband' has already been established as a major discourse topic in the preceding section (not cited here). That is why *lǎopó* 'old-lady' is understood to be **his** wife without the relationship being made explicit, and why he is referred to by a pronoun, in the second sentence. At the same time, at her first introduction represented by a sentence initial full NP, the *lǎopó* is established as a new important participant. Had the husband been agent of this sentence, the wife would most certainly have been introduced as a *bǎ* object:

242) 他把老婆骗到四川去

Tā bǎ lǎopo piàn dào Sìchuān qù

3sg BA Old-Lady cheat to Sichuan. go

'He tricked [his] old lady into going to Sichuan.'

Instead, the newly introduced participant is given prominence, and so is established as an antecedent for later anaphors.

This exception to Object - Object co-reference occurs when a typically agentive entity, a human, has been presented in patient role. This suggests that it could be the underlying semantic role that is the determining factor, i.e. patient -patient co-reference. However, Ø arguments do not always have precisely the same semantic role as their antecedents; the term 'patient' is used loosely to refer to a range of semantic relationships between participant and event, for instance in 241) the more precise role of the anaphor is as an unaffected target, rather than an affected patient, of the verb *zhǎo* 'search' while role of the antecedent is as an affected 'experiencer'.

In the next example, another one where a Ø Object refers to a preceding Subject, rather than an Object, the precise role of the anaphor is as recipient of an invitation (low in affectedness) while the precise role of the antecedent is as an active participant in an event. The tract also involves two topical human participants, who are being treated in the surrounding discourse sometimes as joint agents, and sometimes as agent and patient. This means that potential ambiguity is high, and a reliance on semantic roles alone is insufficient. Indeed, some ambiguity does result. Consider:

243) 我们带他们参加了一次去房县旅游。

wǒmen dài tāmen cānjiāle yīcì qù Fángxiàn lǚyóu
 1pl lead 3pl attend one-time go F. travel
 'We took them on a trip to Fangxian,...

一次去武当山，市里 Ø 也去了。

yīcì qù Wūdāngshān, shìlǐ Ø yě qù le,
 one-time go W. town Ø also go-ASP
 ...and one to Wudangshan,...

Ø 还参加了一个晚会。

Ø *hái cānjiā-le yīgè wǎnhuì*
 Ø and attend-ASP one.CL party,
 ...[we all/ they] also went into town, and Ø went to a party,...

Ø 也到陈涓的单位去了参加私人的生日宴会。

Ø *yě dào Chénjuān-de dānwèi qù-le cānjiā sīrén-de shēngrì yànhuì*
 Ø also to C.-DE unit go-ASP attend private-DE birthday party

...and Ø went to C.'s work unit to attend a private birthday function.'

我们还请了 Ø 一餐

wǒmen hái qǐng-le Ø yīcān,

1pl also invite-ASP one-meal

'We also invited [them] for a meal.'

NS7:3-8

But essentially the topicality and precise involvement of both participants is maintained by topic chaining. The first verb in this text, *dài* 'lead / take' is a shared control verb, because of this, both its Subject and its Object are joint controllers of any empty Subject positions which depend upon that verb. This includes the first two instances of the verb *qù* in the text. The following three verbs, *qù*, *cānjiā* and *qù*, are not syntactically dependent upon the verb *dài*; their Ø Subjects are pragmatically controlled. There is a measure of ambiguity as to whether both participants were involved in these activities (Subject-Subject co-reference), or only the guests (the most topical participant in the discourse as a whole). However, this ambiguity is inconsequential; in the last sentence cited, where the Ø Object appears, ambiguity is removed completely: the hosts are excluded as potential antecedents for the Ø Object, because they are represented explicitly as the Subject of the same clause.

The guests are then the only available, currently topical human participants: they were last represented in a core argument position (i.e. participating in Subjecthood of the preceding verb) as part of a Ø topic-chain; all other human referents are back-grounded as obliques (possessive-modifiers); all other preceding Objects are semantically incompatible (locatives), and not topical, back as far as the original introduction of the guests as Object of *dài* 'lead / take'.

In this way topicality, backgrounding, semantics and the Disjoint Reference Preference (Farmer and Harnish, 1987) all contribute to the correct interpretation. Reference to poorly delimited notions like 'patient' and 'agent' is clearly not sufficient to account for specific interpretations in such a complex discourse environment. The fact that Obliques are not available as antecedents supports the view that GF and not **just** semantic role is important. At the same time, in general terms, though the conventional

association between patient and Object and agent and Subject does appear to play a role in determining reference, the ultimate interpretation is still mediated through the pragmatic status of the anaphor and the antecedent. As with Subjects, various strategies are available for establishing an NP as a likely antecedent. Another exception to immediate Object-Object co-reference, illustrating just such a strategy, involves control by a *bǎ*-Object.

4.3.2 *Bǎ* as a Topicalisation Strategy

The optionally ellipsed NP in the last line of 244) is coreferent with the *bǎ* Object in the first line. Intervening, there are two overt references to another inanimate NP, 'the dishwasher'.

- 244) Ø 再把他们晚饭的碗放入洗碗机，
 Ø *zài bǎ tāmen wǎnfàn de wǎn_{BAO1} fàng rù xǐwǎnjī_{2ndO1}*
 Ø again BA 3pl dinner DE bowls put-into dish-washer
 'Next [I] put their dinner dishes in the dishwasher...
 Ø 就可开机了。
 Ø_s *jiù kě kāi jī le*
 Ø then can start machine ASP
 ...then [I] can start the dishwasher.'
 不能在洗碗机里洗的
 [_{np} [_{rel cl} *bù néng zài xǐwǎnjī-lǐ* *xǐ de*] Ø_{o3}]...
 [_{nn} [_{rel cl} NEG can in dishwasher-LOC wash DE] Ø]
 'Those Ø which can't be washed in the dishwasher...'
 NS3:25-26.

One of these references to the 'dishwasher' presents it as the third (locative) argument of the di-transitive verb *fàng rù* 'put-enter', the other presents it as an unindividuated patientive Object. The fact that these Objects do not control the Ø Object shows conclusively that the principles of co-reference cannot be stated simply in terms of Object-Object, or patient-patient co-reference. The important factor is that neither of these treatments places the dishwasher in focus. By contrast, introduction by *bǎ* is a topicalising strategy: it moves an Object to pre-verbal position. So this is a signal that the *bǎ* Object is likely to be referred to again, and is therefore a better candidate as a referent for the Ø Head than the other overt NPs.

Ø 没想到，

[_{scomp} Ø *méi xiǎngdào*

Ø NEG+ASP think-ASP

'[who'd] have thought.'

Ø 回家 Ø 看了 Ø 两天 Ø 就坏了。

[_s:adj [_{np:s} Ø] [_{vp:ad} *húijia*] [_{vp} *kàn-le liǎngtiān* Ø]] [_{np:s} Ø] [_{vp} *jiù huàile*]]

Ø return home watch-ASP two-days Ø Ø then broke.down

'..[after] returning home and watching [it] for two days, [it] broke down.'

说实话，我看到这条消息时，

Shuō shíhuà, wǒ kàndào zhè-tiáo xiāoxi shí

Speak truth 1sg see this-CL news time

'To tell the truth, when I saw this piece of news'

对老人的同情不是很多的，

dùilǎorén de tóngqíng bù shì hěn duō-de,

to old-man DE sympathy NEG COP very much-DE

'[my] sympathy for the old man was not great...'

Ø 只是觉得他为了那几个钱就去自杀，

Ø *zhǐshì juéde tā wèile nà jǐge qián jiù qù zìshā*

just reckon 3sg for that several money then go self-kill

'[I] just thought [for] him to go and kill himself over that small amount of money,...'

太不值得。

tài bù zhídé,

too NEG worth

'...wasn't worth it.'

而且他应该要求退换Ø。

érqiě tā yīnggāi yàoliú tuìhuàn Ø,

moreover 3sg ought demand return Ø

'What's more, he should have demanded to return [it].'

The omitted Object is clearly not coreferent with the next previous Object (*xiāoxi* 'news'), nor the one previous to that (*shíhuà* 'the truth'). Although these two Object NPs are embedded in parenthetical statements (c.f. Li and Thompson, 1979; Chen Ping, 1984), nothing would prevent them grammatically speaking, from being the antecedent to this Ø pronoun. For instance, if *xiāoxi* 'news' were replaced with

diànshìjī 'TV', they could be coreferent, e.g. 'when I saw the TV.... I thought the old man should return it'.

It would be reasonable to suggest that the verb *tùihuàn* is being used here in an intransitive sense, much as a noun might be in the English phrase: 'he should have demanded an exchange'. The correct interpretation seems to depend on semantics: in the context of the story, a news item is not something for which the old man would demand a refund or replacement. This shows quite clearly that pragmatic 'control' may involve co-reference, but not necessarily lexical antecedence; it is through reference to things in the world, or in our cognitive representations of them, and not through reference to syntactic positions, that pragmatic interpretation functions. This is, of course consistent with the technical distinction between anaphors, which refer to and depend on linguistic units, and referring expressions, which refer to concepts and semantic representations of things we perceive.

5. CHAPTER SUMMARY: SAME GF CO-REFERENCE AS A COROLLARY OF PRAGMATIC CONTROL

In this chapter I examined contexts where ellipsis arises because a structural position that is present is not filled by morphological or phonetic material. In these cases, the reference of the ellipsed argument must also be determined pragmatically. Even when ellipsis is not syntactically bound, Grammatical Function still has a role to play in the signalling of co-reference. There appears to be an expectation that agents will continue to be agents, and patients, patients, and that the former will be represented as Subjects and the latter as Objects (Du Bois, 1987).

However, with both Subject-Subject co-reference, and Object-Object co-reference it is clear that interpretation depends not so much upon the form, or position or GF of the \emptyset element, as it does upon the establishment of an obvious antecedent. This is achieved by reference to pragmatic norms about the representation of new, established, topical, and focal information.

Indeed the norms of same GF co-reference appear to be a consequence of the use of Subjecthood to signal topicality, and Objecthood to signal focus. When speakers appear to depart from the 'norms' of Subject-Subject co-reference, they will be found to comply with the pragmatic demands in other ways: by highlighting the most topical participant in some way, e.g. by the foregrounding of a participant to whom the speaker intends to refer again, and the backgrounding of participants, including even the most topical participant of the text as a whole, who are less pertinent **at that particular point in the discourse.**

A number of clear principles have emerged:

- 1) An overt pronoun places a topical entity in focus, and conversely, ellipsis maintains the topicality of an entity without bringing it into focus.
- 2) Inherently topical NPs which are not the current focus are expressed by \emptyset .
- 3) Inanimate entities, inherently non-topical, are brought more clearly into focus in Object position by individuating them with a classifier.
- 4) Empty positions associated with Subjects (typically established referents) will be interpreted as referring to the last- 'mentioned' topical but backgrounded participant (typically a Subject);
- 5) Empty positions associated with Objects (typically focal or new referents) will be interpreted as referring to the last- 'mentioned' topical participant in focus;
- 6) Antecedents in focus are chosen according to a hierarchy:
Topicalised Object > Topic > previous Object > overt Subject.

Chapter Eight

The Roles and Functions of Ellipsis

1. INTRODUCTION: EXPRESSING RELATIONSHIP

So far we have been concerned with the reference and interpretation of ellipsis. Some clarity has been achieved as to where and how it is interpreted by syntactic processes, and where and how it is pragmatically controlled; and in which structures it is optional, and in which obligatory. This Chapter addresses the question of distribution further, in terms of pragmatic and discourse functions.

The role of ellipsis is considered in a) signalling conjunction and thereby expressing relationships (causal, aspectual etc.) between events in the discourse; signalling continuity of semantic role (agentivity etc.) and thereby contributing to topical cohesion; and backgrounding references to topical participants which might compete as antecedents for a more topical \emptyset element.

1.1 The Function of Ellipsis

It is fair to say that ellipsis serves a **function**, and is not just as an accidental by-product of syntactic dependence, because in any language narrators have a choice as to whether to represent events by way of inter-dependent structures or full sentences. It is especially true for Mandarin, because in Mandarin, overt markers of syntactic dependence are few; close inter-clausal semantic relationships are to some extent ascribed **because** there is an argument ellipsed.

1.1.1 Ellipsis as a Marker of Interdependence

In English, where ellipsis is obligatory, as in dependent clauses, dependency is signalled by the verb form; where English uses special lexemes to signal conjunction, as in co-ordination or sub-ordination, if ellipsis is acceptable at all, it does not generally

contrast with an overt NP. Mandarin does not signal dependency on verbs (except via aspectual markers) and tends to make greater use of simple apposition. 247) illustrates the kind of co-ordination of actions which in English is represented by a conjunction, and in conventional Chinese orthography, is punctuated by the parallel structure comma 、 .

- 247) Ø晚上回来把衣服收回来、
 Ø wǎnshàng húilái bǎ yīfu shòuhúilái、
 Ø evening return BA clothes take-in
- Ø Ø 叠起来、
 Ø Ø diéqǐlái
 Ø Ø fold-up
 'In the evening [I] come home, bring in Ø [and] fold up the clothes.'
 NS3: 20-21

In these circumstances, the expression or omission of an NP can be construed, from the interpreter's viewpoint, not just as a consequence of syntactic inter-dependence, but as a **marker** of it.

Having established that speakers sometimes have a choice as to whether to express a proposition via a more reduced structure or a more extended one, and as to whether to realise a structural position with an overt form, or leave it unfilled, the question can be raised: What influences this choice? What pragmatic and discourse functions are served by the choice of one particular structure from among alternatives?

1.1.2 Levels of Relationship

According to Foley and van Valin (1984:271):

Given the inventory of syntactic clause-linkage categories in a language, it will always be the case that the strongest semantic relations will be expressed in the most tightly linked syntactic configurations found in the language, the weaker relations in the less tightly linked constructions.

This syntactic linking often consists of conjunction and argument sharing, and therefore, ellipsis.

In Foley and van Valin's IRH (Foley and van Valin, 1984), semantic factors such as causality, temporality and correlativity, etc. are referred to. However, inter-clausal relationships can operate at many levels, semantic, pragmatic and narrative. For instance, continuity of agency across two events, or representation of separate incidents as part of a thematically unified story are also aspects of connectedness.

It is suggested then that ellipsis in Mandarin serves the meta-function of expressing relatedness, be it semantic, pragmatic or narrative.

2. THE PRAGMATICS OF REDUCED STRUCTURES

2.1 Agency

2.1.1 Dependent Clauses

In Chapter Five, it was proposed that in dependent clauses, the selection of the complement type was primarily a consequence of verbal semantics: where the matrix verb specifies continued agency in a secondary event XCOMP<SUBJ> is selected, where it specifies transferred agency XCOMP<OBJ> is selected, and where agency is unspecified SCOMP is selected.

Where the matrix verb can have scope over a secondary event, but does not specify agency, then an XCOMP is not lexically required, but it may be selected by the **speaker**, so long as agency is coreferent; but if the verb excludes agency by either of its own arguments, an XCOMP argument will be incompatible. The same semantic effects influence the choice of structure in conjunctive structures.

2.1.2 Conjoined Clauses: Contrastive Reference and Switch Agency

Huang Yan (e.g. 1994:8) suggested that \emptyset in Mandarin is always given the most stereotypical interpretation. His assumption was that this is always a locally coreferent interpretation, but in fact the stereotypical interpretation depends upon the semantics of the lexical items and structures in which the \emptyset occurs.

In Chapter Six we saw that contrastive ellipsis is associated with the expression of causality and correlativity, precisely the factors represented on the higher end of Foley and van Valin's (1984) hierarchy. However, this correlation clearly cannot be related to the hierarchy in a simplistic way, because different positions on the semantic hierarchy may be expressed in the same syntactic structures. The answer is in the relationship between causality and agency.

In causal constructions, the causer is perceived as having control over a causee in the realisation of an event, therefore, the semantics of cause involve not just the

specification of agency, but its retention, as distal agent, by a single participant, similar to the case with matrix verbs that select an XCOMP. This means that shared agency is essentially the norm for causal constructions, and overt expression of an agent functions essentially as a switch-reference marker.

The control of two events by a single agent can be unambiguously signalled by the reduced structure (i.e. VP conjunction), with only one Subject position. There is therefore no need to use the more extended structure (i.e. conjoined Sentences) to express this idea. Though in a language with regular definite ellipsis, a disjoint agent in a structure formed by sentence conjunction might, in principle, be ellipsed, resulting in ambiguity between structures, in practice this structure is preserved for the expression of disjoint reference. Disjoint agency represents a greater level of independence between two events, and thus the iconicity between syntactic and semantic relatedness is preserved. The exclusion of coreferent agents from the conjoined sentence is therefore as consistent with Foley and van Valin's (1984) model as is the choice of VP conjunction when there is co-reference between agents.

In concessive and contrastive constructions, there is no specification of agency, and so no stereotypical interpretation. Instead there is a subtle interplay of topicality, agency and scope. This is evident in the patterns of ellipsis displayed by correlative conjunctions.

2.1.3 Correlative Conjunctions: Focal Scope and Agency

Where clauses are conjoined by correlative conjunctions, the pattern of ellipsis relates to the more detailed semantic content of the conjunctions.

2.1.3.1 Causals

Where events share an agent, a causal correlative conjunction *yīnwèi.... suǒyǐ* 'because....therefore' can only be used at VP level or below; but where they have different agents, they must be conjoined at the sentence level.

248) 他因为出去了所以没见到他。

tā₁ yīnwèi [vp chūqù-le]suǒyǐ [vp méi jiàndào tā]
 3sg1 because out.go-ASP therefore NEG see 3sg
 'Because he went out, [he] didn't see him.'

249) 因为他出去了所以我没见到他。

yīnwèi [_s*tā chūqù-le*] *suǒyǐ* [_s*wǒ méi jiàndào tā*]
because 3sg out.go-ASP therefore 1sg NEG see 3sg
'Because he went out, I didn't see him.'

Overt correlative conjunctions are more versatile, since they do not depend purely on the signalling of joint agency to produce a semantic relationship - they allow for independent signalling of causality and joint or disjoint agency.

2.1.3.2 Concessives

On the other hand, in the concessive structure *sūirán... háishì* 'although... still', there is no inherent implication of shared or retained agency. Instead there is the implication of contrast. This contrastive force inter-acts with topicality in a rather subtle way. In these structures, according to Huang Yan (1994) a sentence-initial Subject is considered more topical than a Subject which follows a conjunction.

In fact, when the Subject precedes the conjunction it is outside the scope of the conjunction, and because of this, outside the scope of focus. Where there is an overt Topic in the first clause, it should be central in both clauses; thus co-reference is already determined, so there is no need to express the Subject in the second clause. An overt Subject preceding the correlative conjunction in the second clause serves to accentuate that the Subject is **outside** the scope of the contrast, and puts the predicate in focus. This throws emphasis on the predicate, conveying the message that the outcome is counter to expectation:

250) 他虽然唱歌他还是不高兴。

tā₁ sūirán chànggē tā_{1/2} háishì bù gāoxìng
3sg although sing.song 3sg still NEG happy
'Although he₁ sings is he_{1/2} nonetheless unhappy?'

In keeping with this interpretation, a Topic in the first clause is unlikely to be followed by an overt non-topical Subject in the second clause:

251) ?他虽然唱歌还是他不高兴。

?*tā₁ sūirán chànggē háishì tā₁ bù gāoxìng*
3sg although sing.song still 3sg NEG happy
'Although he₁ sings nonetheless it is he₁ (not any-one else) who is unhappy.'

This is because such a treatment would put the second Subject in focus, implying that **it** is being contrasted. If this were so, the Subject of the first clause should not be outside the scope of the conjunction.

When there is a non-topical Subject in the first clause, falling **within** the scope of the contrast, a topical Subject in the second clause is coreferent with it, (unless it is stressed).

252) 虽然他唱歌, (?Ø / 他) 还是不高兴。

sūirán tā₁ chàngē (?Ø) / tā₁ háishì bù gāoxìng

although 3sg sing.song ?Ø 3sg still NEG happy

'Although he₁ sings he_{1,2} is STILL unhappy?'

Once again the position of the second Subject clearly indicates that it is the predicate which is the focus of the contrast, despite the position of the Subject in the first clause, no contrast of agency is assumed.

253) ?虽然他唱歌, 还是 (Ø / 他) 不高兴。

?*sūirán tā₁ chàngē háishì Ø / tā₂ bù gāoxìng*

although 3sg sing.song still Ø /3sg NEG happy

'Although it is he who sings yet it is he who is unhappy.'

An overt non-topical Subject in the second clause is unacceptable. This is because the Subject in the first clause is in focus, and therefore we expect the contrastive meaning of the conjunction to be expressed with respect to the two Subjects, i.e. we expect disjoint reference of two pronominal Subjects within the conjunctions' scope. The lack of agreement features on Mandarin pronouns, accompanied by a lack of topical control makes interpretation of such a structure with two different Ø Subjects difficult if not impossible, so it is rejected as incoherent.

A Ø Subject in the second clause is odd because its reference is not retrievable, there being no established topic at that point in the discourse. This inter-action of topicality and contrastive scope produces a situation where disjunction is the norm for Subjects within the scope of contrast, and co-reference is the norm for Subjects outside the scope of contrast.

2.1.3.3 Correlatives

In the correlative structure *yuè...yuè* 'the more...the more', there is no specification as to agency, shared or otherwise; the primary function of the conjunction is to correlate predicates. For this reason Subjects in clauses conjoined by *yuè...yuè* are restricted to Topic position. Since the conjunction does not itself express contrastive meanings, contrast can be expressed with regard to the Subjects even though they are outside the scope of the conjunction. However, when they do not contrast, there is no need to express the second Subject, since co-reference will be assumed on account of Topic control. That is to say, Topics are always assumed to have scope over both clauses unless a second overt topic is introduced in the second clause. Overt expression of a second Topic will be taken as an indication of disjoint reference.

There was only one instance of a correlative construction in the corpus:

254) 不是心脏，就是伤到肺。

bùshì xīnzàng jiù shì shāng dào fèi
 NEG.COP heart then COP wound arrive lungs
 'If it wasn't the heart, it was a stab in the lungs.'

NS8

This correlative pair incorporates the Copula, and assigns a Subject GF only optionally, (cf. the expletive Subject in the English gloss.) It is also a contrastive structure and as with *súirán* 'although', co-reference varies with topicality:

255) 不是我病了，就是(*Ø/我/他)忙。

*bùshì wǒ bìngle, jiùshì *Ø/ *wǒ tā máng*
 CM SUBJ CM SUBJ
 NEG.COP 1sg sick-ASP, then.COP *Ø/ *1sg/ 3sg busy
 'If it's not that I'm sick, it's that *I'm/ he's busy.'

256) 我不是病了，就是(Ø/*我/*他)忙。

*wǒ bùshì bìngle, jiùshì Ø/*wǒ /*tā máng*
 TOP CM CM TOP
 1sg NEG.COP sick-ASP, then.COP Ø/*1sg /3sg busy
 'I'm either sick, or {Ø/ *I'm / *he's} busy.'

257) 我不是病了，(Ø/我/*他)就是忙。

*wǒ bùshì bìngle, Ø/ wǒ *tā jiùshì máng*
 TOP CM TOP CM
 1sg NEG.COP sick-ASP, *Ø/ *1sg/ 3sg then.COP busy
 'If I'm not sick, I'm / *he's busy.'

Overt co-reference is permissible only if both Subjects are in Topic position, outside the scope of focus; overt disjoint reference is permissible only if both Subjects are within the scope of focus.

2.2 Purpose and Causal Relationships

Some verbs do not of themselves allude to a secondary event. When they are closely combined with another verb as in a complement structure, a strong semantic inter-relationship between events is signalled, and will be ascribed to the two clauses by the listener. In 258) we see an example of conjoint sentences expressing purpose,

- 258) 老人用自己节省了的钱
lǎorén yòng zìjǐ jiéshěng-le de qián
 old-person use self save-ASP DE money
 The old man used the money he'd saved

Ø 买了一台黑白电视机。
Ø mǎi-le yìtái hēibái diànshìjī
 Ø buy-ASP a-CL black-white t.v.
 'to buy a black-and-white t.v.'

In 259) sentences are conjoined with the implication of both shared agency and a causal relationship:

- 259) 我住在这里要帮助做家务。
wǒ zhù zài zhèlǐ yào bāngzhù Ø zuò jiāwù
 1sg live at here must help Ø do housework.
 '[Because] I live here, [I] must help Ø do the housework.'
 (NS3:15-16)

With a matrix verb of motion, whether transitive or intransitive, the complement is generally interpreted as being the purpose of the movement encoded by the matrix verb.

- 260) Ø 也到陈涓的单位去了,
Ø_s yě dào Chénjuān-de dānwèi qù-le
 Ø also to C.-DE work-place go-ASP
 'and also went to C's work,

参加私人的生日宴会。
Ø cānjiā sīrén de shēngrì yànhuì
 attend private DE birthday party
 '...to attend a private birthday party.'

261) 我们带他们参加了一次 Ø 去房县 Ø 旅游。

wǒmen dài tāmen cānjiā-le₈ yīcì Ø qù Fángxiàn Ø lǚyóu
1PL take 3PL attend-ASP once Ø go F. Ø travel
'We took them once to join in going to F to go sight-seeing.'

In the second example, the matrix *dài* 'lead' specifies that both its agent and its patient share agency in a subsequent event.

2.3 Similarity of Predications

When conjoined VPs can alternate with conjoined sentences with no contrast in reference, sentence conjunction is still understood to convey the independence of events, but it is not necessarily in relation to agency. The inter-play of factors now extends to the expression of connectedness at the narrative level as well as the semantic.

Clauses in the corpus which were apposed with optional ellipsis, fell into three semantic classes: a series of actions, paraphrasing (262), and amplification (263), as illustrated in Chapter Six. For example:

262) 同学还都是老样子，

tóngxué hái dōu shì lǎo yàngzi
class-mates still all are old appearance,
'The class-mates were all still the same,...'

Ø 没有多大变化。

Ø méi yǒu duō dà biànhuà

Ø NEG have much big change

'...Ø hadn't changed much.'

NS4:8

263) 小王调走了、Ø 调回他爸单位去了。

Xiǎowáng diào zǒu-le Ø diào huí tā bā dānwèi qù-le
X post.away-ASP Ø post.return 3sg Dad unit go-ASP
'Xiao Wang has been posted away, posted back to his Dad's work unit.'

In these cases, there is no important semantic contrast between VP conjunction and Sentence conjunction. Where two predications have essentially similar content, their expression as juxtaposed entire sentences (i.e. with repeated Subjects) would however produce a rather oddly disconnected effect. This is because they would seem to represent the events as unrelated, perhaps reports of incidents of a similar nature, but occurring on

different occasions. Alternatively the second sentence may seem an after-thought on the speaker's part.

In other words, where the content of the clauses suggests that they are **repeated representations** of a single event, any contrast between Sentence conjunction and VP conjunction is primarily one of structural cohesion. The choice of sentence conjunction will more likely be interpreted as reflecting aspects of production, after-thoughts, clarification, self-correction etc., and not given major semantic significance. Where clauses are understood, on the basis of pragmatic knowledge, to represent **different aspects** of a single event, like cause-and-effect, the choice of sentence conjunction will imply a significant disjunction between the two aspects: disjoint agency.

2.4 Summary

The choice of Sentence conjunction when different agents are involved, but VP conjunction when the same agent is involved in each event is consistent with Gricean principles of avoiding ambiguity, as well as iconicity of a tighter structure for sentences expressing more closely connected events.

The appearance and interpretation of zero anaphors in correlative constructions depends on the semantics of the correlative conjunctions, and the position of the Subject NPs inside and outside the scope of the conjunctions, and therefore of the focus. Topical Subjects are assumed to have scope across both clauses, expressing a same-agent relationship, unless this is countered by the introduction of a second overt topical Subject in the second clause. \emptyset anaphora is associated with the co-reference of agents, and therefore with Topical Subjects in the first clause, but lexical semantics may exclude or disfavour this interpretation, in which case a \emptyset argument may be rejected entirely, or interpreted as disjoint. Conversely, co-reference may be expressed via overt arguments as well as zero anaphora, lexical semantics allowing.

3. *SITUATIONS VACANT*

We have now seen that elliptic structures may be chosen to signal a variety of types of close interrelationship between clauses or propositions. But what of positions where

ellipsis cannot signal syntactic dependence because there is no conjunctive structure which can be ascribed, i.e. in the sites where optional ellipsis has variable reference? It is instructive to see where ellipsis might occur, but in practice does not.

3.1 The Distribution of Overt Pronouns

Chen (1984) proposed that overt pronouns occur where there is disjunction, characterised by turning from background to foreground information, from new to given, digressions, insertion of adverbials etc.. From the discussion so far it should be clear that many of these factors are not associated with disjunction, but with the alternation of different 'players' within a single thematic development.

3.1.1 Backgrounding to Avoid Ambiguity

Ellipsis can often be readily explained by the need to background inherently topical NPs, so they don't subvert the intended chain of reference. Therefore ellipsis is somewhat paradoxically, a strategy for reducing ambiguity, by decreasing the number of highlighted topical antecedents.

3.1.2 Avoidance of Agent-only sentences

The case of SCOMP Subjects was discussed above. Chao Yuen Ren (1968:72) drew attention to the opportunity for ambiguity afforded by ellipsis in sentences where a transitive verb appears with only one argument, which could, semantically speaking, be either the agent or the patient; since word order is somewhat variable, and Objects can be pre-verbal, like Subjects, how can we distinguish GF in a sentence with a single NP?

In this corpus, it was found that where one argument of a two-place predicate is omitted, it is much more likely to be the Agentive argument. This is consistent with Du Bois' (1987) observation that Agents are usually established participants, and Givon's (1983) hypothesis that the most topical (retrievable) participants are the least marked, i.e. the morphologically or phonologically least complex.

Moreover, where Objects were ellipsed, the agent was generally ellipsed also. Sentences where Objects were ellipsed but Agents realised, were very few, and where they did appear, they tended to involve indefinite Objects, i.e. intransitive valencies, or

Topic bound Objects or Objects referring to abstract notions, semantically incompatible as Agents, for example:

264) 你那里过五一吗?

nǐ nàlǐ guò wǔyī ma?
2sg there pass Mayday Q-PART
Do you observe Mayday over there?

你怎样过 \emptyset 的?

nǐ zěnyàng guò \emptyset de
2sg how pass \emptyset DE
'How did you spend [it]?'
NS5:14-15

265) 他们的事, 开始妹妹知道。

tāmen de shì kāishǐ, Mèimei zhīdào
3 pl matter start M. know
'As to their affair, M knew [about it] at the start.'

The scarcity of these structures suggests that speakers are well aware of the structural potential for ambiguity, and avoid it wherever necessary.

3.1.3 Lexical Pronouns and Topic Closure

It was suggested in Chapter One, that optionally realised pronouns also correlate with a change in agency within a single thematic event, and with summarising pre-closers. These observations are in keeping with the idea that ellipsis is employed extensively as a back-grounding device; a pronoun functions to bring an established topical entity back into focus, to highlight it in contrast to the ellipsis of others. This may be because, as a new agent, it has an increase in topicality, or because it is the major topic of a thematic section nearing closure, or for potentially many reasons.

3.1.4 Position of the Overt Subject in Adjunctive Constructions

In contrastive ellipsis, which results from adjunction, the overt Subject may appear in either clause. The examples in the corpus were too few to throw much light on why the Subject may be realised in one clause rather than the other. One point worth investigation is the fact that, in each of these sentences, the overt Subject appears in the clause which is more newsworthy, the other clause in each case being background information already

known to the reader. In 267) the information is known from the context: the first sentence of the text introduces the topic of moving house:

266) 我于三月二十八日搬到 25 Y.Terrace.

wǒ yú sānyuè 28 rì bān dào 25 Y.Terrace.

1sg on March 28 day move to 25 Y.Terrace.

'On March 28th I moved to 25 Y Terrace.'

The next several lines discuss the household, then the writer's fruitless search via advertisements. Then the writer says:

267) 这个住处原来住的一个中国学生，

zhège zhùchù yuánlái zhùde yige zhōngguó xuésheng,

this residence originally house one.CL Chinese student

他正好要回国 就建议我来了。

tā zhènghǎo yào huí guó Ø jiù jiànyì wǒ lái le

3sg just.right must return country Ø then suggest I come ASP

'This place originally housed a Chinese student. He coincidentally had to return home and suggested I come.'

The whole sentence is introduced with a topic, 'this residence', and the choice of *lái* 'come', places the writer's perspective firmly within that residence, so the final clause is clearly less newsworthy. By contrast, the Chinese student's fortuitous return home is newsworthy, and it is in this clause that the overt pronoun appears. In 268) the known information is in the real-world knowledge of the addressee, that she had telephoned the writer:

268) 上星期六, 接到你们来的电话我们兴奋不已。

shàng xīngqīliù Ø jiēdào nǐmen lái-de diànhuà wǒmen xīngfèn bùyǐ

last Saturday, Ø receive 2pl come-DE phone-call 1pl excited no.end.

'Receiving your call last Saturday, we were excited no end.'

Another factor is that when one clause has a stative predicate, it is likely to be that clause which has an overt Subject. Compare 268) where it is in the second clause when that has a stative, to 269) and 270) below, where the Subject appears in the first clause, with a stative predicate.

269) 我不小心 掉到了水沟里...

wǒ bù xiǎoxīn Ø diào dào le shuǐgōu lǐ...

1sg NEG careful Ø fall into ASP ditch into

'I was careless and fell into a ditch.'

NS5:9.

270) 他很痛苦，就自杀了。

Tā, hěn tòngkǔ, jiù zìshāle

3sg very miserable, then self-kill

'He was very distressed and killed himself.'

NS2: 10-11

In 268) the second clause has a Copula and stative predicate *xīngfèn* 'excited', and it would sound very odd if the Subject appeared in the first clause. In 267) above also, the main predicate is not stative, but it is dependent on an auxiliary expressing necessity, so it is not a prototypically active clause. I don't know why stative clauses attract overt Subjects; examples in the corpus were too few and it warrants further investigation.

4. CHAPTER SUMMARY: PRAGMATIC FUNCTIONS OF REGULAR DEFINITE ELLIPSIS

Ellipsis in Mandarin can be seen as having the function of signalling inter-clausal semantic relationships and textual cohesion. It also serves to background arguments which are topical, but not the current focus of the discourse. The complex patterns of distribution of Subjects with respect to correlative conjunctions have been explained on the basis of the interaction of focal scope and semantics. The role of overt pronouns as a topic-closure marker has been proposed, and some evidence presented to show that there are pragmatic limitations on the distribution of elliptic structures, especially those with more than one ellipsed NP. Regular definite ellipsis clearly has potential for numerous discursal applications, and writers of Mandarin manage it in such a way as to optimise communication, and minimise ambiguity.

SUMMARY AND MAJOR CONCLUSIONS

1. PROBLEMS OF PREVIOUS ACCOUNTS

In Chapters Two and Three it was shown that the standard syntactic accounts of the reference of zero elements have difficulty in accounting for the appearance of \emptyset elements which refer beyond the sentence: they fit none of the 4 empty categories proposed in GB; they may refer outside the sentence, and therefore be neither f-commanded as LFG requires of PRED 'pro' nor c-commanded, as GB requires anaphors and pronominals, nor have an antecedent with no independent theta role as GB requires for traces, and still not have indefinite reference; they refer not so much to a specific lexical item, but to a specific participant in the discourse.

The best that syntactic theory can do with these zero elements is to recognise that they are beyond the scope of syntactic control, that other factors determine their reference. Though most syntacticians do indeed recognise this, there is disagreement as to which zero elements are within the scope of syntactic control, and which are not, and to which type of control mechanism operates in different cases within the domain of syntactic control.

The reference of \emptyset elements in 'pro-drop' languages is at the centre of the first difficulty, and James Huang's (1984) treatment of \emptyset Objects represents an area where a conviction that control is syntactic inspires the suggestion of a sentence internal \emptyset operator to control the \emptyset Object, even though the reference of this \emptyset operator itself is then a mystery.

The disagreement as to categorical membership arises in the treatment of \emptyset Subjects of infinitival clauses: in GB they are all seen as formally identical, with the exception of raising constructions, and this propagates difficulty in accounting for their outward variation in terms of reference and control; in LFG, they are treated as belonging

to two separate categories, with some verb complements being treated in a manner identical to raising constructions, and others being treated in a manner identical to \emptyset Subjects of VPs in NP positions. This requires the formulation of additional constraints to explain similarities in the reference of all \emptyset VCOMP Subjects.

Discourse approaches to zero anaphora in Mandarin have also not been entirely successful. They have tended to underestimate the role of syntax in the control of reference, and because of this attempt, to find pragmatic explanations for phenomena which are more readily described in terms of syntax. This leads to unnecessarily complicated accounts.

Overall, the model of syntactic control followed in this thesis is closest to LFG. The LFG model imposes fewer constraints on word order; it allows for some \emptyset elements to be actually absent from f-structure, as well as c-structure, which means both that there is a potential two-way split between types of \emptyset elements, and that there are fewer discrepancies between f- and c- structure; and it offers a more flexible mechanism of control for PRED 'pro', as well as the possibility of uncontrolled PRED 'pro'. These factors conform relatively well to the observable data from Mandarin. However some amendments have been proposed to the Standard LFG analysis, specifically in the area of the control of verb complements, and in the removal of constraints on interpretation of PRED 'pro'.

2. AN ALTERNATIVE MODEL

The analysis proposed here is that ellipsis is essentially binary: either an argument position has been removed by conjunction, adjunction, or in XCOMP arguments, and arguments are shared at f-structure level; or an argument position is present in f-structure, represented by PRED 'pro' but phonetically unrealised. In the first case, reference is syntactically determined, and in the latter it is pragmatically determined.

2.1 Syntactic Control

Removal of an argument position can occur at many levels, via VP embedding (verb complements, serial verbs), VP conjunction (correlative constructions, apposition

etc.) and VP adjunction (aspectual sub-ordination, causal adjuncts etc.), via Verb conjunction and Object conjunction. In the case of dependent clauses, arguments are shared because Subject GFs are assigned to a matrix argument under functional control; in cases of conjunction and adjunction, conjoined constituents simply assign GFs to arguments above them in the phrase structure tree, or are assigned GFs by constituents below them just as they would in independent structures; the only difference being that a single argument may bear GFs in two clauses simultaneously.

2.2 Pragmatic Control

In Mandarin, empty argument positions can, by and large, occur wherever the discourse context renders them interpretable. There are practical constraints on their appearance (or rather non-appearance) however, relating to too many topical referents (ambiguity, or incoherence), or too few (unretrievability).

Since in this model, all obligatorily absent dependent Subjects are under functional control, PRED 'pro' is never involved in the interpretation of dependent Subjects, and there is therefore no need for f-command and obviation to constrain its reference. When PRED 'pro' is introduced by anaphoric control, in independent clauses, it is pragmatically controlled by reference to notions like topicality and focus.

2.2.1 Same GF: a Corollary of Topicality and Focus

Topicality is seen as complex, with multiple topics being fore-grounded and backgrounded at different points within a thematic development. The backgrounding of topical elements that stand between a \emptyset and its nearest previous coreferent NP is seen as central to referent tracking.

Subjects tend to be established topics, and \emptyset Subjects tend to be coreferent with the last Subject in the discourse; new participants tend to be introduced as Objects, and Objects are in focal position. Where speakers depart from these norms, highlighting and backgrounding strategies are used to clarify the intended chain of reference.

2.2.2 Hierarchy of Focus

∅ Objects are always interpreted as referring to the topic most recently in focal position along a hierarchy:

bǎ object > sentence Topic > nearest previous Object > nearest previous Subject.

Note that there are different hierarchies of antecedents for ∅ Subjects and ∅ Objects. This is supported by the fact that a Topic NP will always be preferred over a Subject as controller of a ∅ Object, but a Subject NP will always be preferred over a Topic as controller of a ∅ Subject.

2.2.3 Highlighting Strategies

If a topic is introduced as an Object, but the speaker plans to cast it as a Subject in succeeding discourse, it will be individuated and topicalised, and following potentially topical Subjects will be ellipsed. If a participant is topical but in a patientive role, it will be introduced first as a Topic, a *bǎ* object, or a Subject in a passive construction, thus establishing its topicality, and making it a retrievable referent for subsequent ∅ Objects.

3. APPLICATIONS OF THE MODEL TO COMMON CONSTRUCTIONS

3.1 Syntactic Control

3.1.1 Relativised Constituents

No new analysis of relative clauses has been proposed. One point that does emerge (and is discussed at length by Huang Yan (1994) is that while the relationship between the 'gap' and the relativised constituent or Head is clearly syntactic in nature, where regular indefinite ellipsis allows the production of relative clauses with no overt arguments, the decision as to which ∅ argument is a reflex of the relativised constituent appears to be made on the basis pragmatic considerations. The mechanism suggested here is different from that in Huang (1994): he proposes that the identity of the NPrel is determined according to a hierarchy whereby ∅ Subjects are interpreted before ∅ Objects, and are first assumed to be coreferent with a Subject, and failing that an Object and failing that a topic etc.. The proposal made here is that relative clauses with two gaps are essentially ambiguous, and will only be used, or accepted without question, when

semantic restrictions imposed by the verb of the relative clause on its arguments make the intended reference obvious.

3.1.2 Reduced Structures

3.1.2.1 XCOMPs and Functional Control

The most common source of bound ellipsis in the corpus was the ellipsis of Subjects in dependent clauses. Around 82% of dependent Subjects in the discourse were obligatorily ellipsed, and therefore argued to involve XCOMPs under functional control. In this analysis, 'ellipsis' in this context actually arises from the lack of a structural position. A special mechanism (Functional control) is necessary here because of the occurrence of Subject-Object co-reference, a pattern that cannot be produced by the standard mechanisms by which a single verb assigns GFs to its arguments.

SCOMPs, where overt dependent Subjects are possible, were all associated with verbs of speech or cognition, and always had an overt Subject unless it was highly topical. Nine out of a total of 11 SCOMPs were independent SCOMPS with their own tense, aspect or illocutionary force, exceptions were complements in irrealis mood. SCOMPs are possible for a relatively restricted set of verbs, with a common semantic core relating to presenting the propositional content of a thought, desire or speech act, i.e. discourse complements. Ellipsis in this context represents an unfilled position that is present in f-structure, its reference is determined by pragmatic control.

3.1.2.2 Clausal Adjuncts

Causal and temporal Adjuncts are clauses which exhibit the property of allowing only one overt Subject between two clauses, but the Subject may precede either clause. They are examples of adjunction of a modifying clause as sister to the Subject and another VP. The Subject NP stands in the same relationship to both clauses, and so is a shared argument of both clauses. These structures alternate with ones where a sentence is adjoined as modifier to a sentence. That structure is not reduced: it has two overt Subject positions; and is chosen when the Subjects are not coreferent.

3.1.2.3 Apposition, Conjunction and Correlative Constructions

Clauses that stand in semantic relationships of greater independence than causality may also be conjoined, but the Subject always precedes the first clause. They are therefore thought to consist of two VPs conjoined beneath a single VP node. Once again this removes a Subject position, but it does not place one VP in a position to modify the other. These structures may also alternate with sentence conjunctions, but this choice would convey the idea that the two events were in a random relationship, producing a sense of textual incohesion.

3.1.3 Classification of Mandarin Verbs

A classification of Mandarin verbs was proposed based on their ability to take an SCOMP and/or XCOMP argument, the ability to take an XCOMP depending upon the specification by the verb's semantics, that a participant expressed by one of its own arguments is agent in a secondary event. Given the distinctions between transitivity and intransitivity, subject- and Object- control, this classification generates sufficient variability to account for seven different collocational patterns observable in Mandarin.

3.1.3.1 VPs in Object or Copula Complement Position

Even though they sometimes have some nominal trappings (e.g. a numeral and classifier), clauses in Object or Copula complement positions are still dependent clauses. In this analysis they are considered to be XCOMPs with a functionally controlled Subject just as other dependent clauses with obligatorily ellipsed Subjects.

3.2 Pragmatic Control

Zero elements under pragmatic control include heads of relative clauses, Subjects of VPs in Subject position, \emptyset Objects in Topic-comment structures, and arguments in independent clauses and tensed SCOMPs. Since VPs in Subject position can always take an overt Subject, they are thought to be SCOMPs, and \emptyset Subjects are structurally present, but optionally ellipsed. These are essentially all determined by the same mechanisms, as described above: with Subject-Subject co-reference and Object-Topic co-reference, the norm.

3.2.1 Pragmatics and Structural Alternatives

The idea that the appearance of contrastive and optional ellipsis arises as a result of alternations of similar structures explains why reference is more restricted in conjoined sentences with \emptyset elements (i.e. reduced structures) than in comparable conjoined sentences with overt pronouns. It also means that speakers have a choice of structures with which to express very similar ideas, and it is suggested that they make the choice on the basis of an iconic correlation between syntactic interdependence and semantic interdependence.

3.2.2 Ellipsis and the Chain of Reference

It is only where definite ellipsis is regular, that ellipsis can be employed to pragmatic ends; in other types of languages, overt pronouns are usually the least prominent encoding device, and, unless stressed, are unlikely to have a highlighting function. But in languages like Mandarin, the overt pronoun can and clearly does have this role. This may well explain the marked differences in the use of ellipsis and overt pronouns by speakers of English from pro-drop language backgrounds (for instance as reported by White, 1985).

The multi-functional heterogeneous and subtle nature of ellipsis, means the speaker's choice between zero and a pronoun, cannot be adequately explained on the basis of measures of potential ambiguity, and referential distance. Nor can it be explained in terms of topic-continuity, so long as topic is defined simply as Subject or Topic NPs, or as 'what the sentence is about'. To advance our understanding of the interpretation, roles and functions of ellipsis in discourse, there must be clarity as to which constructions admit a choice with regard to lexicalisation, and which prescribe or proscribe it altogether; characterisations of topicality need to take account of the changing relative status of participants from sentence to sentence and clause to clause, throughout a text, and above all, elliptic structures must be considered in the natural contexts in which they occur.

References:

- Andrews, Avery. In preparation 'Syntax Textbook'
- Aoun, J. 1985. *A Grammar of Anaphora*. Cambridge, MA: MIT Press.
- Battistella, E. 1985. 'On the distribution of PRO in Chinese' *Natural Language and Linguistic Theory* 3:317-340.
- Besnier, Niko 1985. 'The local organisation of zero anaphora in Tuvaluan Conversation.' *Te Reo: Journal of the Linguistic Society of New Zealand* 28:119-147.
- Borer, Hagit, 1989. 'Anaphoric AGR' in: Jaeggli and Safir (eds) *The null Subject parameter*. Dordrecht: Kluwer.69-109.
- Bouchard, Denis. 1983. *On the content of empty categories*. Dordrecht: Foris.
- 1989. 'Null Objects and the theory of Empty Categories; selected papers from the 17th Linguistic Symposium on Romance Languages. Rutgers University, 27-29 March 1987'. In Kirschner and DeCesaris, (eds) *Studies in Romance Linguistics*. Amsterdam: Benjamins.
- Bresnan, J. (ed) 1982. *The Mental Representation of Grammatical Relations*. Cambridge, MA: MIT Press
- Brody, M. 1985. 'On the complementary distribution of Empty Categories. *Linguistic Inquiry* 16 (4):505-546.
- Chao Yuen Ren, 1968. *A Grammar of Spoken Chinese*. Berkeley/ Los Angeles: University of California Press.
- Charters, A.H. to appear, 'A comparison of Learner and Native NP ellipsis in Mandarin Discourse.' *Te Reo: Journal of the Linguistic Society of New Zealand*.
- Chafe, Wallace. 1987. 'Cognitive constraints on information flow.' in R. Tomlin (ed) *Coherence and Grounding in Discourse*. Typological Studies in Language vol. 11: 175-188. Amsterdam/Philadelphia: John Benjamins.
- Chen, Paul, 1990. *The distribution and referential interpretation of empty categories in Chinese*. Dissertation Abstracts International
- Chen Ping, 1984. *A Discourse Analysis of Third Person Zero Anaphora in Chinese*. Bloomington Indiana: Indian University Linguistics Club.
- Chomsky, N. 1980. 'On Binding' *Linguistic Inquiry* 11:1-46
- 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
- 1982. *Some Concepts and Consequences of the theory of Government and Binding*, Cambridge, Mass.; MIT Press.
- Chung, Sandra, 1989. 'On the notion of 'null anaphor' in Chamorro'. In Jaeggli and Safir (eds), 143-184.
- Curnow, T.J. 1992. *Complement clauses in Spanish: A Semantic Approach*. Honours thesis (The Australian National University), Canberra.
- Dowty, D.R. 1980. 'Comments on the Paper by Bach and Partee.' In K.J. Kreiman and A.E. Oteda (eds) *Papers from the Parasession on Pronouns and Anaphora*. Chicago: Chicago Linguistic Society.

- Du Bois, J. 1987. 'The Discourse Basis of Ergativity' *Language* 63 (4): 805-855.
- Enfield, Nick, 1994. *Aspects of Lao Syntax : Theory, Function and Cognition*. Honours thesis, (The Australian National University), Canberra.
- Espino de Valdivia, I. 1989. 'Universal trends in Nominal Zero Anaphora: an Inquiry into Japanese and Spanish Discourse.' *Transactions of the International Conference of Orientalists in Japan* 34:70-83.
- Farmer A.K. and R.M. Harnish, 1987. 'Communicative reference with pronouns. In Verschueren & Bertuccelli-Papa (eds) 547-65.
- Fillmore, Charles J. 1969. 'Types of Lexical Information.' In Ferenc Kiefer (ed) *Studies in Syntax and Semantics*. Dordrecht: Reidel, pp. 109-137.
- 1986. 'Pragmatically controlled zero anaphora'. In Nikiforidou et al. (eds), *Proceedings of the Twelfth Annual Meeting of the Berkeley Linguistics Society*. Berkeley: BLS. 95-107.
- Foley, W.A. and R.D. Van Valin. 1984. *Functional syntax and universal grammar*. Cambridge: Cambridge University Press.
- Ford, M, J. Bresnan and R.M. Kaplan, 1982. 'A competence-based theory of syntactic closure' in Bresnan (ed) 1982. *The Mental Representation of Grammatical Relations*. Cambridge, MA: MIT Press. 727-796.
- Givon, T. 1983. 'Introduction' in T. Givon (ed) *Topic Continuity in Discourse: A Quantitative Cross-Language Study*. Vol 3 of Typological Studies in Language. Amsterdam: Benjamins. pp5-39.
- 1984. 'Universals of Discourse Structure and Second Language Acquisition.' In W. Rutherford (ed) *Language Universals and Second Language Acquisition*. Amsterdam: John Benjamins, pp 109-33.
- 1987. 'Beyond Foreground and Background' in R. Tomlin (ed) *Coherence and Grounding in Discourse*. Typological Studies in Language vol. 11: 175-188. Amsterdam/Philadelphia: John Benjamins.
- Grice, H. P. 1975. 'Logic and conversation.' In P. Cole and J. Morgan (eds) *Syntax and Semantics, vol. 3* 41-58 New York: Academic Press.
- Grima, John A. 1986. 'Discourse factors contributing to the understanding of a zero pronoun in a passage from the Phraraatchawicaan' In Robert J. Bickner, et al. (eds) *Papers from a conference on Thai Studies in honour of William J. Gedney*. Ann Arbor: Centre for South and South-East Asian Studies.
- Haegeman, L. 1991. *Introduction to Government and Binding Theory*. Oxford/Cambridge Mass: Basil Blackwell.
- Hopper, P. and S. Thompson. 1980. 'Transitivity in Grammar and Discourse' *Language* 56: 251-299.
- Horrocks, G. 1987. *Generative Grammar*. London and New York: Longman.
- Hou, John Y. and C. Kitagawa, 1987. 'Null Operators and the Status of Empty Categories in Chinese.' *Linguistic Inquiry* 1987:518-523.

- Huang, C.T. James. 1982. *Logical relations in Chinese and the theory of grammar*. Ph.D. dissertation. MIT.
- 1984. 'On the distribution and Reference of Empty Pronouns', *Linguistic Inquiry* 15(4) 531-44
- 1987. 'Remarks on Empty Categories in Chinese' *Linguistic Inquiry* 18:35-43.
- 1989. 'Pro-drop in Chinese: a Generalised Control Theory' in Jaeggli and Safir (eds) *The Null Subject Parameter*. Dordrecht/Boston/London: Kluwer Academic Publishers.
- 1993. 'Reconstruction and the structure of VP: some theoretical consequences' *Linguistic Inquiry*, 24(1) 103-138.
- Huang Lillian Meei-Jin, 1987. *Aspect, a general system and its manifestation in Mandarin Chinese*. PhD (Rice Universty) Ann Arbor: UMI Dissertation Information Service.
- Huang, Yan. 1992. 'Against Chomsky's typology of Empty Categories.' *Journal of Pragmatics*; 17: 1-29.
- 1994. *The syntax and pragmatics of anaphora*. Cambridge: Cambridge University Press.
- Jaeggli O. and K.J. Safir (eds), 1989. *The null Subject parameter*. Dordrecht: Kluwer.
- Junghare, Indira Y. 1988. 'Topic-prominence and zero NP-anaphora in Marathi and Hindi.' In Jazayery and Winter (eds) *Languages and Cultures: studies in honour of Edgar C. Polome*. Trends in Linguistics series 36. Berlin: Mouton de Gruyter. 309-327.
- Kaplan, Ronald, M., and Joan Bresnan, 1982. 'Lexical-Functional Grammar: A Formal System for Grammatical Representation.' In Bresnan (ed) *The Mental Representation of Grammatical Relations* 174-281.
- Keenan. 1975. 'Towards a Universal Definition of Subject.' In Charles Li (ed) *Subject and Topic*. 303-333.
- Kroeger, Paul. 1993. *Phrase Structure and Grammatical Relations in Tagalog*. Dissertations in Linguistics. Stanford: CSLI Publications.
- Lambrecht, K. 1987. 'Sentence focus, information structure, and thethetic-categorial distinction.' *Proceedings of the 13th Annual Meeting of the Berkeley Linguistics Society*, 366-82.
- La Polla, R. 1993. 'Arguments Against 'Subject' and 'Object' as Viable Concepts in Chinese.' *The Bulletin of the Institute of History and Philology, Academia Sinica*. Vol LXIII Part IV. Taiwan: Academia Sinica.
- 1994. 'Pragmatic relations and word order in Chinese.' in M. Noonan and P. Downing (eds) pp 299-331.
- Lasnik, H. and R. Fiengo, 1989. 'Complement Object Deletion' in Howard Lasnik *Essays on Anaphora Studies in Natural Language and Linguistic Theory*. Dordrecht: Kluwer. 58-89.

- Levinson S.C. 1983. *Pragmatics*. Cambridge: Cambridge University Press.
- 1987a. 'Minimisation and conversational inference'. In Verschueren and Bertuccelli-Papi (eds) 61-129.
- 1987b. 'Pragmatics and the grammar of anaphora: a partial pragmatic reduction of Binding and Control phenomena.' *Journal of Linguistics* 23, 379-434.
- 1991. 'Pragmatic reduction of the Binding Conditions revisited.' *Journal of Linguistics* 27. 107-61.
- Li, Y.A. 1990. *Order and constituency in Mandarin Chinese*. Dordrecht: Kluwer
- Li, Charles, N. 1976. *Subject and Topic* London: Academic Press.
- Li, Charles, N. and Sandra Thompson, 1979. 'Third-Person Pronouns and Zero-Anaphora in Chinese Discourse' in Talmy Givon (ed) *Syntax and Semantics* Vol 12. Discourse and Syntax. New York and London: Academic Press
- 1981. *Mandarin Chinese, a functional reference*. USA: Berkeley, University of California Press.
- Lust, Barbara. 1986. (ed) *Studies in the Acquisition of Anaphora* 2 Vols. Dordrecht: D. Reidel.
- Manzini, Maria Rita, 1983. 'On control and control theory.' *Linguistic Inquiry* 14: 554-561.
- Mohanan, K.P. 1981. *Grammatical and Anaphoric Control*. Ms Department of Linguistics and Philosophy, MIT.
- Mosel, U. 1991. 'Towards a typology of Valency.' in *Partizipation. Das sprachliche Erfassen von Sachverhalten*. Seiler and Premper (eds) .Tübingen: Gunter Narr Verlag.
- Noonan, M. and P. Downing (eds), 1994. *Word Order in Discourse* (Typological Studies in Language 30). Amsterdam/Philadelphia: John Benjamins.
- Pinker, Steven, 1982. 'A Theory of the Acquisition of Lexical Interpretive Grammars' in Bresnan (ed) "*The Mental Representation of Grammatical Relations*" 655 - 726.
- Reinhart, Tanya, 1976. *The Syntactic Domain of Anaphora* (PhD dissertation, MIT, Cambridge, Mass.)
- 1983. *Anaphora and Semantic Interpretation*. London and Canberra: Croom Helm.
- Rizzi, L. 1986. 'Null Objects in Italian and the theory of *pro*', *Linguistic Inquiry*, 17, 501-58.
- Rosenbaum, P.S. 1967. *The Grammar of English Predicate Complement Constructions*. Cambridge, Mass: MIT Press.
- Safir, K. 1986. 'Relative clauses in a theory of binding and levels', *Linguistic Inquiry*, 17, 4, 663-90.

- Shopen T., 1972. 'A Generative Theory of Ellipsis: A Consideration of the Linguistic Use of Silence.' Reproduced by the Indiana University Linguistics Club.
- 1973 'Ellipsis as Grammatical Indeterminacy' in *Foundations of Language* 10:65-77.
- Smith, S. Doug, N. B. Chin, K. Louie, and C. Mckeras, 1993. *Unlocking Australia's Language Potential. Profiles of 9 key languages in Australia. Vol 2. Chinese.* Canberra: National Languages and Literacy Institute of Australia Ltd, and DEET.
- Sun, Chao-Fen and Talmy Givon, 1985. 'On the so-called SOV word order in Mandarin Chinese: a quantified text study and its implications.' In *Language* 61:329-377.
- Tan Chieh-Fu, 1991. *Notion of Subject in Chinese* PhD (Stanford University) Ann Arbor: UMI Dissertation Information Service.
- Tomlin, Russell S. and Ming-Ming Pu 1991. 'The Management of Reference in Mandarin Discourse' *Cognitive Linguistics* 2:65-95.
- Verschueren, J. and M. Bertuccelli-Papi (eds), 1987. *The Pragmatic Perspective.* Amsterdam: John Benjamins.
- White, L., 1985. 'The 'pro-drop' parameter in adult second language acquisition' *Language Learning* 35: 47-61.
- Wierzbicka, Anna, 1988. *The Semantics of Grammar.* Studies in language Companion Series 18. Amsterdam/ Philadelphia: John Benjamins.
- 1992. *Culture and Cognition. Universal concepts in culture specific configurations.* New York / Oxford: Oxford University Press.
- In press. *Semantics. Primes and Universals.* Oxford: Oxford University Press.
- Williams, E. 1980. 'Predication', *Linguistic Inquiry*, 11, 1, 203-38.
- Williams, J., 1989. 'Pronoun copies, pronominal and zero anaphora.' In S. Gass et al (eds) *Variation in Second Language Acquisition: Discourse and Pragmatics.* 153-189 Avon: Multilingual Matters.
- Wu Guo. 1992. 'Thethetic-categorical distinction as expressed by Subject-predicate sentences in Chinese.' *Australian Journal of Linguistics* 12 (2) 271:296.