



# THE UNIVERSITY *of* EDINBURGH

<b>Title</b>	Formal approach to long-distance anaphora : the case of Korean
<b>Author</b>	Gil, Kook-Hee
<b>Qualification</b>	PhD
<b>Year</b>	2001

Thesis scanned from best copy available: may contain faint or blurred text, and/or cropped or missing pages.

## **Digitisation Notes:**

Page XII and 188 are missing in original document

# **A Formal Approach to Long-Distance Anaphora: The Case of Korean**

**Kook-Hee Gil**  
**B.A., M.Sc., M.Sc.**

**A thesis submitted in fulfilment of requirements for the degree of  
Doctor of Philosophy**

**to**

**The Department of Theoretical and Applied Linguistics,  
University of Edinburgh**

**August 2000**





# Abstract

The aim of this thesis is to provide a unified account of Long-Distance Anaphora in Korean. We especially focus on the behaviour of the long-distance anaphor *caki*. We set two objectives in this work. The first objective is empirical and consists in the establishment of the different operative factors licensing *caki*-binding. The second objective is theoretical, and consists in the formulation of a uniform account of the different binding phenomena characteristic of *caki*. The syntactic framework assumed is Head-driven Phrase Structure Grammar (HPSG), where binding relations are determined in a single level of representation.

Long-distance anaphora such as *caki*-binding has received a great deal of attention due to its recalcitrant nature and the difficulty that scholars have encountered in incorporating it into standard binding theory. The accounts provided so far combine elements from different parts of the grammar (syntax, semantics and pragmatics). However, such hybrid approaches cannot explain uniformly and parsimoniously mainstream data patterns such as subject and non-subject long-distance binding of *caki*.

The review of such approaches in this thesis shows that *caki* resists pronoun-like binding in the discourse domain. From a theory internal point of view, on the other hand, recent proposals, within HPSG, to handle long-distance anaphora are shown to be not general enough to cover non-subject binding cases of *caki*.

The objectives set above are realised through a three-pronged strategy. First, a comprehensive investigation of the empirical domain together with a critical review of relevant works; Second, a unified explanation for both subject and non-subject binding of *caki*; Finally, the syntactic formalisation of the proposed account and its incorporation into HPSG together with the necessary revisions to the framework in order to achieve maximal empirical coverage with minimal cost to the elegance and generality of the theory.

The outline of the thesis is as follows: Chapter Two presents a detailed description of the data of *caki*-binding. This will involve the distribution of *caki* in a wide range of syntactic environments in order to establish the right set of properties characteristic of this anaphor and its binding patterns. Chapter Three presents approaches to long-distance anaphora and provides a critical evaluation against the data from *caki*-binding. Chapter Four covers the basics of HPSG and its binding theory including recent developments, which are also evaluated in terms of *caki*-binding. Chapter Five provides a theory for *caki*-binding and its implementation in HPSG. The theory is essentially based on the lexical argument hierarchy (Kiss 1990) where syntactic prominence (external argument) and thematic prominence are lexically represented. Chapter Six questions the status of ARG-ST as the appropriate level for the implementation of our theory of *caki*-binding and offers a revision. This revision involves the altogether dismissal of ARG-ST as a significant theoretical construct and presents arguments in favour of its replacement by DEPS list. It is argued that this is a desirable move as it makes the framework more elegant while avoiding unnecessary redundancy. Chapter Seven explores the case of topic-binding. We show that topic-binding can be neatly reduced to our formal account without involving discourse factors. Topic-binding is also presented in terms of type-hierarchy in Chapter Eight.

# Declaration

I hereby declare that this thesis is of my own composition, and that it contains no material previously submitted for the award of any other degree. The work reported in this thesis has been executed by myself, except where due acknowledgment is made in the text.

A handwritten signature in black ink, appearing to read 'Kook-Hee Gil', written in a cursive style.

Kook-Hee Gil

# Dedication

I dedicate this thesis to my beloved parents,  
Ssang-Suk Gil and Hee-Ja Jung.

이 논문을 사랑하는 부모님께 바칩니다.

# Acknowledgements

First, I would like to thank my first supervisor Ronnie Cann for all the constructive advice and help during the past 4 years of my PhD life. I would like to say thanks to him for all the patience he took during the supervision, especially when I very often took more than 10 minutes to try some judgement over one single Korean sentence and changed my mind within a second afterwards. I remember that I have always come out of the supervision feeling greatly encouraged. He may not believe it, but I have enjoyed every single bit of supervision. I would have not survived my Ph.D days if he had not been my supervisor.

I am greatly grateful to my second supervisor Colin Matthieson for such detailed comments on each part of the draft of the thesis. He always gave me a quick response whenever I had an unexpected theoretical problem and have been always patient in explaining to me all the technical details I asked about. I would never be able to pay back all the coffees he offered me in each supervision even when I unexpectedly popped in with a question or a chunk of draft. He has been a great supervisor.

I am very much indebted to George Tsoulas, Peter Sells, Masayo Iida, Dora Alexopoulou, Claire Grover, Caroline Heycock, Ivan Sag and Sunshim Hong. Firstly, I would like to thank George Tsoulas for all the difficult but profound and constructive questions which made me extremely stressed but which were extremely helpful.

I would not have been so much advanced in linguistics if I had never met him. Secondly, I would like to thank Peter Sells who always gave me such a spontaneous and constructive advice and encouragements. He was one of a few people who got back to me with helpful comments on my papers and questions as he promised. I am also very grateful to Masayo Iida for being patient whenever I asked about her view on *zibun*-binding. It was the most helpful especially when she took so much time answering all my questions in very much detail about *zibun*-binding relations-it was one afternoon over lunch in Stanford I will not forget. My thanks also go to Dora Alexopoulou who is the most wonderful linguist friend. We organised an HPSG workshop just between her and me. We had a lot of chat about HPSG over tons of coffee and when I look back now, I realise that I mastered all the basics of HPSG through this lonely but cosy meeting with her. I also thank Claire Grover who was so willing to offer help whenever I desperately needed some technical answer to my troubled questions. I am very grateful to Caroline Heycock for her great help whenever I just appeared at her office door for abrupt questions, references or even copies of books (I should not forget that I still have to return a few books to her).

There are a few people who always appear in the acknowledgement in every PhD thesis in our department and who deserve it: Ethel Jack, Eddie Dubourg, Irene Macleod and Fiona Paterson. They are the most truly helpful people in the department that no student can survive their PhD days without.

Then, I must say that I had so many wonderful friends around me throughout my PhD days: Ivan Yuen, Dora Alexopoulou, Natasha Mangana, Olga Goubanova, Satu Manninen, Ash Asudeh, Jessie Tseng, Patricia Mabugu. I thank Ivan Yuen for being my best friend throughout my days in Edinburgh, for having been my wonderful bridesmaid and for everything. I also owe a lot of things to Dora for keeping me company even when I thought I was the more depressed person in the world. I am truly grateful to Natasha Mangana who at some point became a sister of mine (a Greek one) and a member of the LumberJacks. She has been really patient when I was the most grumpy towards the end of my PhD days and had to listen to all my ramblings about each sentence in my whole thesis. I now feel



very sorry for asking her to proofread about 200 pages of my thesis within about 20 minutes. There are two friends who are equally wonderful in offering helps in anything I may need: Ash Asudeh and Jessie Tseng, who made me always feel that I am lucky having such a good friends.

My heartiest thanks to George Tsoulas who became my husband yesterday. It is simply impossible describe here all the things I am grateful to him for. I thank him for everything and I am just so lucky.

I would also like to thank my sister Sung-Hee Gil and my brother-in-law, Kyung-Hoon Lee for their consistent support throughout my stay away in Edinburgh. Especially my sister has been the most wonderful person in cheering me up by giving me updates, and funny and happy stories about the families and relatives back in Korea. I have been always looking forward to talking to her. (I hope she will phone me more often.)

Last but not the least, Heeja Jung and Ssangsuk Gill have been always the most supportive and loving parents that one can ever have. For both my education and personal life, they have been always unconditionally supportive and encouraging to my decisions and they have always been there for me whenever I needed them. My stay abroad would have been much more painful without their love. They are just so wonderful and I hope they know that I love them very much.

# Contents

Abstract	i
Declaration	iii
Dedication	iv
Acknowledgements	v
Glossary	viii
Chapter 1 <b>Introduction</b>	1
1.1 Outline of the Thesis . . . . .	4
Chapter 2 <b>Long-Distance Anaphora in Korean</b>	7
2.1 Anaphora in Korean . . . . .	7
2.2 Background of <i>Caki</i> -Binding . . . . .	9
2.3 <i>Caki</i> in Matrix and Embedded Clauses . . . . .	10



2.3.1	<i>Caki</i> in Simple Sentences . . . . .	10
2.3.2	<i>Caki</i> in Embedded Sentences . . . . .	12
2.3.3	<i>Caki</i> in A Complex NP Domain . . . . .	14
2.4	<i>Caki</i> in Adjunct Clauses . . . . .	18
2.5	Genitive <i>Caki</i> . . . . .	21
2.6	Multiple occurrences of <i>Caki</i> . . . . .	23
2.7	<i>Caki</i> without an Overt Antecedent . . . . .	25
2.8	Topic Antecedents . . . . .	27
2.9	<i>Caki</i> vs. Pronouns . . . . .	32
2.10	Summary of the Facts . . . . .	35
Chapter 3	<b>A Critical Review of Approaches to Long-Distance Anaphora (LDA)</b>	38
3.1	Syntactic Approaches . . . . .	39
3.1.1	The Subjecthood Condition . . . . .	39
3.1.2	The Basic C-commanding Requirement . . . . .	42
3.1.3	Approaches in GB . . . . .	45
3.2	Non-Syntactic Approaches . . . . .	51
3.2.1	Discourse Approaches . . . . .	52
3.2.2	Thematic roles . . . . .	55
3.3	Disjunctive Theory of Long-Distance Anaphora . . . . .	60

Chapter 5	<b>Long-Distance Anaphora: The Core Factors</b>	104
5.1	A New Approach . . . . .	104
5.1.1	Revisiting Cases of Non-Subject Binding and The Thematic Hierarchy . . . . .	106
5.1.2	The Thematically Motivated Lexical Argument Hierarchy . .	113
5.1.3	Formal Treatment of <i>Caki</i> -Binding Within HPSG . . . . .	120
	Parametrising the locality of ARG-ST . . . . .	121
	The Lexical Argument Hierarchy in ARG-ST . . . . .	128
	<i>Caki</i> -Binding in Structured ARG-ST . . . . .	136
5.1.4	Summary and Conclusion . . . . .	141
Chapter 6	<b>Theoretical Consequences:</b>	
	<b>Argument Structure, ARG-ST, DEPS and <i>Caki</i>-binding</b>	143
6.1	Problems in ARG-ST in HPSG . . . . .	145
6.1.1	Syntactic Functions in ARG-ST . . . . .	145
6.1.2	Adjuncts in ARG-ST . . . . .	147
6.1.3	ARG-ST as a phrasal sign . . . . .	150
6.2	Is ARG-ST a Binding Structure? . . . . .	152
6.3	Redundancy in BINDING LIST and DEPS-LIST . . . . .	155
6.4	Binding in the DEPS List . . . . .	157
6.4.1	DEPS and Constraints on SLASH . . . . .	158
6.4.2	DEPS as a HEAD Feature? . . . . .	162

6.5	Conclusion . . . . .	165
<b>Chapter 7 Other Components of LDA</b>		<b>167</b>
7.1	Introduction . . . . .	167
7.2	Caki Without an Overt Antecedent . . . . .	168
7.2.1	The Major Factors . . . . .	168
7.2.2	The Basic Facts and Previous Accounts . . . . .	169
7.2.3	The Double Nominative Construction . . . . .	171
7.2.4	<i>Caki</i> -Binding in Double Nominative Constructions . . . . .	173
7.2.5	Topic-Orientedness and the Expression of Topics . . . . .	175
7.2.6	An HPSG Approach . . . . .	177
7.3	Topicalisation and <i>Caki</i> -Binding . . . . .	184
7.3.1	Topic and Word-order . . . . .	184
7.3.2	Data: Fronted Topics and Antecedents of <i>caki</i> . . . . .	186
7.3.3	“Scrambling” vs. Extraction in HPSG . . . . .	189
	Short-Distance Topicalisation . . . . .	189
	Long-distance Topicalisation: Extraction . . . . .	190
7.3.4	Predictions: Topic-Binding and DEPS . . . . .	195
	Topic-Binding in Short-Distance Topicalisation . . . . .	195
	Topic-Binding in Long-Distance Topicalisation . . . . .	198
7.3.5	Parametric Possibilities . . . . .	201

Tsoulas 1999 . . . . .	201
Anaphors and $\phi$ -features . . . . .	203
7.3.6 Conclusion and Open Problems . . . . .	207
<b>Chapter 8 Topic-Binding in The Type-Hierarchy</b>	<b>208</b>
8.1 Type-Hierarchy: Phrases and Clauses . . . . .	209
8.2 A Revised Phrasal Hierarchy: TOPICALITY . . . . .	217
8.3 Summary and Conclusion . . . . .	230
8.4 Conclusion . . . . .	232
<b>Chapter 9 Summary and Final Remarks</b>	<b>233</b>
Appendix A List of Type-Hierarchy and Constraints	237
References	245

# Chapter 1

## Introduction

Anaphora is, as far as we know, a universal phenomenon, but it is only universal inasmuch as we have to admit that probably all languages contain anaphoric items, that is, items that refer back to another element. In general, anaphoric elements can be divided into three main subtypes: NP anaphors, Pronominal anaphors and reflexive anaphors:

- (1.1) a. NP Anaphor  
A man came into the office. The man approached the secretary.
- b. Pronominal Anaphor  
A man came into the office. He approached the secretary.
- c. Reflexive Anaphora  
A man came into the office. He showed a picture of himself to the secretary.

Among the different kinds of anaphoric elements above, it is the third category, reflexive anaphors, that is the centre of the focus in this thesis. In distributional terms, reflexive anaphors exhibit, in large, two subtypes: local reflexive anaphors and long-distance reflexive anaphors (long-distance anaphora, or LDA, from here on). A typical example of reflexive anaphors can be found in English where a reflexive takes a clausemate antecedent as illustrated below:

(1.2) Mary<sub>i</sub> believes that Jennifer<sub>j</sub> criticises herself<sub>\*i/j</sub> too often

Here, only the clausemate subject *Jennifer* is allowed as an antecedent of *herself*, but not the higher subject *Mary*. On the other hand, long-distance anaphors can take an antecedent in a much larger domain. Consider the following example in Korean.

(1.3) Mary<sub>i</sub>-nun [Jennifer<sub>j</sub>-ka caki<sub>i/j</sub>-lul nemwu pipanhanta]-ko mitnunta  
Mary-TOP Jennifer-NOM self-ACC too much criticised-COMP believes  
'Mary<sub>i</sub> believes that Jennifer<sub>j</sub> criticises herself<sub>i/j</sub> too much'

Compared with the English reflexive in (1.2), Korean long-distance anaphor *caki* can take as its antecedent either the higher subject *Mary* or the clausemate subject *Jennifer*.

Even though the analyses of local reflexives have been relatively well established (for example, in Chomsky (1981) and Pollard & Sag (1994)), long-distance anaphora has remained very much a topic of debate and received various analyses using different aspects of grammar such as syntax, semantics and discourse (for example, the Subjecthood Condition (Kuroda 1965, Kuno 1973), cyclic movement (Chomsky 1986), thematic roles (Momoi 1985), discourse conditions (Sells 1987, Iida 1996), etc.).

On the other hand, it appears that the variation in the data is more important in LDA than it is in local anaphora. This fact though is not always recognised. One evident example of this comes from LDA in Korean and Japanese. The two languages have been generally believed to have similar structural properties including anaphoric phenomena. In fact, there has been a tendency to believe that whatever principle is valid for one language should be applicable to the other. Despite the low-level differences though, the more profound question remains: what is the underlying universal principle that licenses all these varieties of Long-distance anaphora?

In order to answer this question, this thesis looks at a subset of long-distance anaphora, *caki*-binding in Korean. Even though Korean is well known to be one of the languages with long-distance anaphora, *caki*-binding itself has been paid relatively less attention due to its seemingly similar properties to *zibun*-binding in Japanese. *Zibun*-binding has been analysed as involving various aspects of grammar such as syntax, semantics and pragmatics. One of the central concerns of this thesis is to investigate *caki*-binding patterns in a wide range of constructions from the point of view of various approaches to date. By way of this, we hope to reveal the core factors which underly the complex binding patterns of *caki*.

The ultimate findings of such investigation will show that in treating *caki*-binding, one can make do without reference to “contextual” antecedents whose binding domains are, almost by definition, rather obscure and unconstrained and thus to reduce the long-distance, context-controlled binding instances to a more formal and syntactically constrained mechanism. Related to this, the extensive observation of *caki*-binding in various syntactic environments will show that the part of *caki*-binding known to be discourse induced (namely, topic-binding) is in fact due to the particular syntactic constructions *caki* occurs in, such as double nominative constructions and different topic constructions.

As for the technical side of this thesis, the theory I adopt and revise to provide a formal syntactic account of *caki*-binding is the theory of binding proposed by Pollard & Sag (1994) for Head-Driven Phrase Structure Grammar (HPSG) and its recently updated version as set out in Manning (1996), Manning & Sag (1998a) and Manning & Sag (1999). The binding theory in HPSG is centrally couched at the level of argument structure (ARG-ST) where elements are ordered by the notion of *obliqueness*. In the course of implementing the core factors to *caki*-binding, I will propose two major and interrelated modifications to the theory. The first one is to adopt the lexical argument hierarchy (Kiss 1991) at ARG-ST. The lexical argument hierarchy arranges arguments in a hierarchical order depending on the lexical item; in this it is distinct from the notion of *obliqueness* or thematic hierarchy. This will not only treat cases of subject binding of *caki*, but also explains cases of irregular

non-subject binding pattern of *caki*. It is on the basis of the lexical argument hierarchy at ARG-ST that we will propose the formal binding principle of *caki*.

The second major modification to the theory is to replace ARG-ST by Dependent Structure (DEPS). There are two reasons for this. First, adjuncts play an important role in explaining *caki*-binding pattern, but ARG-ST as it is does not include adjuncts. Second, the parallel existence of ARG-ST and DEPS induces a redundancy in the theory.

In sum, the major goal of this thesis is firstly to investigate what represents the major factors for *caki*-binding, about which we will subsequently argue that *caki*-binding is not subject to discourse factors, but rather to syntactic constraints. The second goal of this thesis is to provide a syntactic formulation for *caki*-binding within the framework of HPSG.

## 1.1 Outline of the Thesis

This thesis is divided into seven chapters. Chapter Two describes a wide range of *caki*-binding in various structural environments from simple sentences to complex sentences such as N-complement clauses, relative clauses and adjunct clauses. Further, we examine the occurrence of *caki* in topic constructions and also provide a distributional comparison between *caki* and pronouns in Korean. This will provide an extensive overview of the data, which will help in reaching a generalisation of the complex binding patterns of *caki*. The data in this chapter are uncontroversial in judgement unless the text indicates otherwise.

In Chapter Three, we critically review the approaches offered for long-distance anaphora. We will subdivide the literature review into syntactic and non-syntactic approaches, which are followed by the disjunctive and conjunctive theory of long-distance anaphora. In each section of this chapter, we will attempt to apply the approaches in question to *caki*-binding. From a thorough review of the literature from the point of view of *caki*-binding, we argue that the analyses we reviewed fail



to provide a unified account for subject and non-subject instances of *caki*-binding. The conclusion of this chapter roughly describes an ideal theory for *caki*-binding.

In Chapter Four, we present an introduction to the basics of HPSG and its binding theory, focusing on the description of the relevant technical aspects and basic principles of the theory. Though a large portion of this chapter will be devoted to the introduction to the theory, in the third section of the chapter will discuss some recent developments in the binding theory in HPSG, A-binding (Manning 1996, Manning & Sag 1998a) where we will point out some empirical problems for *caki*-binding.

In Chapter Five, on the basis of the overview of the data in Chapter Two, we propose the core factors which underlie *caki*-binding and argue that it is the lexical argument hierarchy (Kiss 1991) that adequately explains both subject and non-subject cases of *caki*-binding. This is followed by the formal implementation of the lexical hierarchy into the binding theory of HPSG, on the basis of which we provide a formal binding principle for *caki*. In doing so though, we will raise some technical problems originating from ARG-ST as a lexical feature, which we will provisionally amend in this chapter. However, the chapter concludes with a brief discussion of such residual problems, which will be addressed in the subsequent chapter.

Chapter Six reexamines ARG-ST with respect to two points: one is whether it is an adequate level for the determination of *caki*-binding relations. The other is to point out its different aspects from the traditional argument structure. The discussion will lead us to conclude that ARG-ST is not an adequate level for explaining *caki*-binding and the provisionally revised list (namely, Binding Structure) causes a unnecessary parallel to DEPS. Therefore, the rest of the chapter attempts to reduce the two lists to one DEPS and show that once DEPS is made a HEAD feature, it can serve as a proper level for binding without disturbing the extraction mechanism which is the original purpose of DEPS.

Chapter Seven deals with the data which have been generally treated by discourse analyses. By looking at the data closely, we will argue that the data in question are not actually subject to discourse factors, but rather related to the common property

among languages with long-distance anaphora: topic-orientedness, and thus we call such case of *caki*-binding *topic-binding*. That is, we will propose that *caki*-binding in the data in question is determined by topic-related constructions such as double nominative constructions or short/long topic constructions. Thus, *caki*-binding is structurally explained without depending on any discourse factors. Whereas *caki*-binding in double nominative constructions and short topic constructions is well accounted for by way of lexical rules, the account in long-distance topic constructions encounters a theory internal problem. We provide two possibilities to resolve the problem and also discuss their residual problems.

Relating to Chapter Seven, Chapter Eight provides a unified account of topic binding based on the type-hierarchy. We will present the topic construction discussed in Chapter Seven in type-hierarchy. We discuss the problem with the current type-hierarchy with respect to the constraint on topic clauses and provide a major extension to it by adding another dimension of type-hierarchy of TOPICALITY in parallel to HEADEDNESS and CLAUSALITY. We account for topic-binding on the basis of the multiple inheritance in the new type-hierarchy. We also discuss the advantages of such a type-hierarchy based account over the ones provided in Chapter Seven.

## Chapter 2

# Long-Distance Anaphora in Korean

Long-distance anaphora is a complex linguistic phenomenon which requires a careful and detailed description of a range of constructions before theoretical generalisations can be established. The purpose of this chapter is to provide a description of the data concerning long-distance anaphora in Korean, which will then be analysed. The data we will examine involve the Korean anaphor *caki* and its binding patterns in a range of constructions from simple sentences to complex sentences containing complement clauses, relative clauses and adjunct clauses. The data will also include cases involving “psych” verbs and topic constructions, and will provide a comparison between anaphors and pronouns. At this point, the discussion will be made strictly in descriptive terms keeping the theoretical apparatus to a minimum.

### 2.1 Anaphora in Korean

Korean anaphors include two subtypes as in many other languages: reciprocals and reflexives. However, in this thesis, the Korean reciprocal ‘*sero*’ will not be examined and all our attention will be given to reflexives. Korean has two subtypes of reflexives: *caki* and *casin*. Both can be case marked for nominative, genitive and accusative by being suffixed with the appropriate case marker. *Casin* can stand on

its own as a reflexive, and it can be also combined with a pronoun which agrees in person with the antecedent. This is illustrated in (2.1):

- (2.1) a. Ku<sub>i</sub>-ka **casin**<sub>i</sub>/**caki**<sub>i</sub>-ul pipanhayessta  
 He-NOM **self**-ACC criticised  
 ‘He<sub>i</sub> criticised **self**<sub>i</sub>’
- b. Na-nun **na-casin**-ul miwehanta  
 I-NOM **I-self**-ACC hate  
 ‘I hate **myself**’
- c. Ku-nun **ku-casin**-ul miwehanta  
 He-NOM **he-self**-ACC hate  
 ‘He hates **himself**’

*Caki* differs from *casin* in that it cannot be combined with pronouns as shown in (2.2a). *Caki* on its own is specified as third person singular, and accordingly, it should be bound by a third person singular nominal (Sohn 1994), as illustrated in (2.2b), whereas *casin* is free of this restriction.

- (2.2) a. Na<sub>i</sub>/Ne<sub>j</sub>/John<sub>k</sub>-un \***na-caki**<sub>i</sub>/\***ne-caki**<sub>j</sub>/\***ku-caki**<sub>k</sub>-lul miwehanta  
 I/You/John-NOM **I-self/you-self/he-self**-ACC hate  
 ‘I<sub>i</sub>/You<sub>j</sub>/John<sub>k</sub> hate(s) \***myself**/\***yourself**/\***himself**’
- b. Na<sub>i</sub>/Ne<sub>j</sub>/John<sub>k</sub>-un **caki**<sub>\*i/\*j/k</sub>-lul miwehanta  
 I/You/John-NOM **self**-ACC hate  
 ‘I<sub>i</sub>/You<sub>j</sub>/John<sub>k</sub> hate(s) **self**<sub>\*i/\*j/k</sub>’

Apart from this difference between the two reflexives, they behave identically with respect to binding relations (Hong 1985), thus, we will use one of these, *caki*, for our data illustrated in this chapter.

There is, in fact, yet another kind of reflexive in Korean. The reflexives we just mentioned, *caki* and *casin*, can be combined together, which produces a morphologically complex reflexive *cakicasin*. *Cakicasin* has properties more similar to *caki* than to *casin*, in that it can be bound only by a third person singular nominal and

also it is not allowed to combine with a personal pronoun<sup>1</sup>. Further, *cakicasin* must be locally bound<sup>2</sup>.

- (2.3) Mary<sub>j</sub>-ka [John<sub>i</sub>-i **cakicasin**<sub>i/\*j</sub>/**\*ku-cakicasin**-ul miwehanta]-ko malhayssta  
 Mary-NOM John-NOM **selfself/he-self**-ACC hate-COMP said  
 ‘Mary said that John<sub>i</sub> hates **selfself**<sub>i</sub>/**\*himself**’

## 2.2 Background of *Caki*-Binding

The Korean reflexive *caki*, firstly, exhibits properties similar to those of a reflexive pronoun in languages such as English in that it can be bound by a clause-mate antecedent (Sohn 1994) as in (2.4).

- (2.4) John-i **caki**<sub>i/\*j</sub>-ul ttayliessta  
 John-NOM **self**-ACC hit  
 ‘John<sub>i</sub> hit **self**<sub>i/\*j</sub>’

Furthermore, *caki*, along with the Japanese equivalent *zibun*, has been frequently described as a ‘long-distance anaphor’ which takes antecedents outside of the immediate clause as in (2.5).

- (2.5) a. John<sub>i</sub>-i [Bill<sub>j</sub>-i **caki**<sub>i/j</sub>-lul pipanhanta]-ko malhayssta  
 John-NOM Bill-NOM **self**-ACC criticised-COMP said  
 ‘John<sub>i</sub> said that Bill<sub>j</sub> criticised **self**<sub>i/j</sub>’  
 b. Taroo<sub>i</sub>-ga [Hanako<sub>j</sub>-ga **zibun**<sub>i/j</sub>-o hihansita]-koto-o hanasita  
 Taroo-NOM Hanako-NOM **self**-ACC criticised-COMPL-ACC said  
 ‘Taroo<sub>i</sub> said that Hanako<sub>j</sub> criticised **self**<sub>i/j</sub>’

<sup>1</sup>We will discuss relevant aspects of the distribution of *casin* in due course when it becomes relevant in section 7.3.4.

<sup>2</sup>By ‘locally bound’, we mean that an anaphor is bound with its local domain, for example here, within the local clausal boundary.

Given the fact that *caki* allows a much wider binding domain than reflexives in English, in the remaining part of this chapter we will look at a range of constructions in order to highlight both the binding domain and binding properties of *caki*.

## 2.3 *Caki* in Matrix and Embedded Clauses

In this section, we will examine the binding patterns of *caki* in three contexts: a local binding domain, an extended domain and a complex NP domain.

### 2.3.1 *Caki* in Simple Sentences

In a simple clause construction, *caki* can occur in any nominal position except for the subject position. This is shown in the examples in (2.6).

- (2.6)
- a. John<sub>i</sub>-i Bill<sub>j</sub>-ekey *caki*<sub>i/\*j</sub>-etayhayse malhayssta  
 John-NOM Bill-DAT **self**-about talked  
 ‘John<sub>i</sub> talked to Bill<sub>j</sub> about **self**<sub>i/\*j</sub>’
  - b. John<sub>i</sub>-i Bill<sub>j</sub>-etayhayse *caki*<sub>i/j</sub>-ekey malhayssta  
 John-NOM Bill-about **self**-DAT talked  
 ‘John<sub>i</sub> talked about Bill<sub>j</sub> to **self**<sub>i/j</sub>’
  - c. John<sub>i</sub>-i Bill<sub>j</sub>-lopwute *caki*<sub>i/?j</sub>-etayhayse tulessta  
 John-NOM Bill-from **self**-about heard  
 ‘John<sub>i</sub> heard from Bill<sub>j</sub> about **self**<sub>i/?j</sub>’
  - d. **Caki**<sub>\*i/\*j</sub>-ka John<sub>i</sub>-ekey teipu-lul cwuessta  
**self**-NOM John-DAT tape-ACC gave  
 ‘**self**<sub>i/j</sub> gave John<sub>i</sub> a tape’

In (2.6), *caki* appears in the adjunct (*about NP* in (2.6a) and (2.6c)) or in the complement PP (*to NP* in (2.6b)). The occurrence of *caki* in the matrix subject position is ruled out as shown in (2.6d). From (2.6a) to (2.6c), the local subjects are legitimate antecedents of *caki*, a fact well attested for anaphors in other languages. However, we can also observe that non-subject binding is possible. In (2.6c), *caki* can be bound by the NP *Bill* inside the adjunct as well as by the subject, though

the interpretation of the adjunct binding is slightly harder to obtain than that of subject-binding. Interestingly enough, however, other PPs such as *to Bill* fail to bind *caki* as shown in (2.6d) and (2.6a).

While an adjunct phrase like *from NP* can bind *caki*, a direct object may also bind *caki* such as in (2.7) below:

- (2.7) John<sub>i</sub>-i Mary<sub>j</sub>-lul caki<sub>i/j</sub> cip-ulo tolye ponayssta  
 John-NOM Mary-ACC self house-to return sent  
 ‘John<sub>i</sub> returned Mary<sub>j</sub> to self<sub>i/j</sub>’s house’

Here, *caki* may be ambiguously bound by the direct object *Mary* or the subject *John*, indicating that the direct object can be a possible antecedent of *caki*. We will examine the direct object binding in more detail later in section (2.5). We should note however that this possible antecedence can be overridden if the resulting interpretation is pragmatically odd. Consider the following:

- (2.8) John<sub>i</sub>-i Mary<sub>j</sub>-lul caki<sub>i/j</sub> cip-ulo chotayhayssta  
 John-NOM Mary-ACC self house-to invited  
 ‘John<sub>i</sub> invited Mary<sub>j</sub> to self<sub>i/j</sub>’s house’

(2.8) has the same structure as (2.7), but the direct object does not appear as a possible antecedent in the former. In (2.8), if the direct object antecedes *caki*, it yields the odd interpretation that Mary is invited to her own house by John, and this immediately excludes *Mary* from the antecedence of *caki*. Thus, if the sentence is contextualised as below, the direct object binding is allowed:

- (2.9) John<sub>i</sub>-i Bill<sub>j</sub>-lul caki<sub>i/j</sub>-uy cen anay-uy cip-ulo chotayhayssta  
 John-NOM Bill-ACC self-GEN ex wife-GEN house-to invited  
 ‘John<sub>i</sub> invited Bill<sub>j</sub> to self<sub>i/j</sub>’s ex-wife’s house’

The context is such that *Bill* whose ex-wife is *John*’s current partner, is invited by *John* to his ex-wife’s house. Given this context, the direct object *Bill* can bind *caki*.

The observation we made this section shows that even though the subject acts as a prime antecedent of *caki* in the local domain, there are cases where a non-subject may antecede *caki* and this includes a certain category of adjuncts, especially, *from NP* and also a direct object.

Before we close this section, let us briefly consider an example in relation to (2.6c). We saw that the antecedency of *from NP* gives a marginally acceptable interpretation. However, the following example with the same structure as (2.6c) yields a completely acceptable reading when the subject is replaced by a first person pronoun.

- (2.10) Na-nun John<sub>i</sub>-uropwute **caki**<sub>i</sub>-eytayhay tulessta  
 I-TOP John-from **self**-about heard  
 ‘I heard from John<sub>i</sub> about **self**<sub>i</sub>’

In (2.10), as *caki* is third person, it is not bound by the first person subject and *John* in the adjunct phrase is the only possible antecedent. The reading of this binding relation is not as marginal as that in (2.6c). The only speculation at this stage may be that the awkwardness of the non-subject binding disappears when there is only one possible antecedent, and all things being equal subjects are preferred binders.

We now move on to the extended binding domain where *caki* appears in the embedded clause to observe cases of the long-distance binding relations of *caki* and to compare the binding relations in such domains with those in the local domain.

### 2.3.2 *Caki* in Embedded Sentences

While we have seen in (2.6d) that *caki* cannot appear in the subject position in matrix clauses, when it occurs in an embedded clause as in (2.11), it can appear in the lower subject position and can be bound by the higher subject across the clausal boundary.



- (2.11) John<sub>i</sub>-i [caki<sub>i</sub>-ka aputa]-ko malhayssta  
 John-NOM self-NOM is ill-COMPL said  
 ‘John<sub>i</sub> said that self<sub>i</sub> is ill’

*Caki* can also appear as a lower object and in this case it can be bound by either the higher or the lower subject as in (2.12) below. This indicates that the subject binding takes place both in the local and in extended binding domains.

- (2.12) John<sub>i</sub>-un [Bill<sub>j</sub>-i caki<sub>i/j</sub>-lul hyemohanta-nunkes]-ul anta  
 John-NOM Bill-NOM self-ACC hate-COMPL-ACC know  
 ‘John<sub>i</sub> knows that Bill<sub>j</sub> hates self<sub>i/j</sub>’

In fact, long-distance binding is not limited to occurring across only one clausal boundary but can occur across a number of clausal boundaries. One such case is shown as in (2.13) below, where *caki* can be potentially bound by every NP that precedes it, regardless of the number of intervening clausal boundaries.

- (2.13) John<sub>i</sub>-i [Mary<sub>j</sub>-ka [Bill<sub>k</sub>-i [Anna<sub>l</sub>-ka caki<sub>i/j/k/l</sub>-lul cohahanta]-ko  
 John-NOM Mary-NOM Bill-NOM Anna-NOM self-ACC like-COMPL  
 anta]-ko malhayssta]-ko mitnunnta  
 know-COMPL said-COMPL believe-COMPL  
 ‘John<sub>i</sub> believes that Mary<sub>j</sub> said that Bill<sub>k</sub> knows that Anna<sub>l</sub> likes self<sub>i/j/k/l</sub>.’

Subject-binding is preferred just as in the local domain. Now let us examine cases where nominals other than the subject are involved. Consider the following examples:

- (2.14) a. Mary<sub>i</sub>-ka Bill<sub>j</sub>-ekey [caki<sub>i</sub>-ka John-ul coahanta]-ko malhayssta  
 Mary-NOM Bill-to self-NOM John-ACC like-COMPL said  
 ‘Mary<sub>i</sub> said to Bill<sub>j</sub> that self<sub>i</sub> likes John’  
 b. Mary<sub>i</sub>-ka John<sub>j</sub>-uropwute [caki<sub>i/j</sub>-ka Bill<sub>k</sub>-lul hyemohanta]-ko tulessta  
 Mary-NOM John-from self-NOM Bill-ACC hate-COMPL heard  
 ‘Mary<sub>i</sub> heard from John<sub>j</sub> self<sub>i/j</sub> hates Bill<sub>k</sub>’

In the previous section (2.3.1), we have seen cases of non-subject binding in the local domain. For example, the adjunct *from NP* can bind *caki*, whereas other adjuncts cannot. Now, the two examples given above include adjuncts: one is *from NP* and the other is *to NP*. These adjuncts appear in the higher clause in relation to *caki* in the lower clause, and their binding relations exhibit the long-distance pattern. Apart from the subject-binding which we will take as given, the interesting fact is that the binding relation of *caki* with respect to the adjuncts in the extended domain is consistent with that in the local domain. That is to say, the PP *from NP* binds *caki* in the embedded clause, whereas the PP *to NP* does not. At this point, we can conclude that the non-subject binding patterns are consistent in the local domain and in the extended domain. From this set of data, we can also assume that the antecedency of *from NP* unlike other adjuncts may not be accidental. There may be factors that determine the antecedency of a specific adjunct, which one hopes is reducible to the same principle as subject binding.

### 2.3.3 *Caki* in A Complex NP Domain

In this section, we will examine a set of examples which involve relative clause constructions and N-complement clause constructions. Both constructions have in common that a clause is attached to a head noun phrase to form a complex NP. What we are interested in in these constructions is whether the binding pattern of *caki* in embedded clauses in the type of sentences that we saw in the previous section will differ from the binding relations of *caki* in clauses embedded in an NP.

When *caki* is embedded within a relative clause, it consistently allows the subject-binding patterns.

- (2.15) Mary<sub>i</sub>-ka [NP [RC John<sub>j</sub>-i **caki**<sub>i/j</sub>-eytayhay ssu-n] chayk-ul] ilkessta  
 Mary-NOM John-NOM **self**-about wrote-REL book-ACC read  
 ‘Mary<sub>i</sub> read the book which John<sub>j</sub> wrote about **self**<sub>i/j</sub>’

Here, the example has a relative clause modifying the NP which is the direct object. *Caki* within the relative clause is shown to take either the higher subject *Mary* or the lower subject *John* as its antecedent. This shows that subject-binding can be licensed across the complex NP domain. Now let us turn our attention to cases of non-subject binding. As we have examined the contrast in the antecedency of the adjuncts, *from NP* and *to NP*, we will consider the same contrastive set of examples in the context of relative clauses. Consider the following examples:

- (2.16) a. Mary<sub>i</sub>-ka John<sub>j</sub>-ekey [<sub>NP</sub> [<sub>RC</sub> Bill<sub>k</sub>-i **caki**<sub>i/\*j/k</sub>-etayhay ssu-n]  
 Mary-NOM John-DAT Bill-NOM **self**-about wrote-REL  
 chayk]-etayhayse malhayssta  
 book-about talked  
 ‘Mary<sub>i</sub> talked to John<sub>j</sub> about the book which Bill<sub>k</sub> wrote about  
**self**<sub>i/\*j/k</sub>’
- b. Mary<sub>i</sub>-ka John<sub>j</sub>-uropwute [<sub>NP</sub> [<sub>RC</sub> Bill<sub>k</sub>-i **caki**<sub>i/j/k</sub>-etayhay ssu-n]  
 Mary-NOM John-from Bill-NOM **self**-about wrote-REL  
 chayk]-etayhayse tulessta  
 book-about heard  
 ‘Mary<sub>i</sub> heard from John<sub>j</sub> about the book which Bill<sub>k</sub> wrote about  
**self**<sub>i/j/k</sub>’

These two examples have *to NP* and *from NP* respectively in the matrix clause. The relative clauses are also situated in the adjunct position *about NP*, and *caki* is embedded in the relative clauses. Apart from the local subject *Bill*, we can observe that *caki* can be bound by the antecedent *Mary* which is not only outside the clause boundary (which we have seen so far), but also outside the complex NP domain. Further, the very interesting contrast in these two examples shows that even in the complex NP domain, the adjunct *from NP* still acts as an appropriate antecedent, whereas the PP *to NP* does not.

As this contrastive non-subject binding seems to be consistent over the local, embedded, and the complex NP domain, we will now consider examples where these constructions are embedded within each other, as in (2.17).

- (2.17) a. Mary<sub>i</sub>-nun John<sub>j</sub>-ekey [[**caki**<sub>i/\*j</sub>-ka ip-un] os-i messi  
 Mary-NOM John-DAT **self**-NOM wear-REL cloth-NOM smart  
 issta]-ko malhayssta  
 is-COMPL said  
 ‘Mary<sub>i</sub> said to John<sub>j</sub> that the clothes that **self**<sub>i/\*j</sub> is wearing are  
 smart’
- b. Mary<sub>i</sub>-nun John<sub>j</sub>-uropwute [[**caki**<sub>i/j</sub>-ka ip-un] os-i messi  
 Mary-NOM John-from **self**-NOM wear-REL cloth-NOM be  
 issta]-ko tulessta  
 smart-COMPL heard  
 ‘Mary<sub>i</sub> heard from John<sub>j</sub> that the clothes that **self**<sub>i/j</sub> is wearing are  
 smart’

(2.17) is constructed as a result of the combination of a complement clause and a relative clause. That is, the relative clause is embedded within a complement clause. This means that, as for the long-distance binding, *caki* has to look for its antecedent, firstly, across the complex NP domain and then again across the complement clause. In spite of a number of different clausal boundaries, as the indexing shows, the higher subject *Mary* can bind *caki* without any problem. Furthermore, the contrast in non-subject binding between *from NP* and *to NP* is still maintained in this structurally complex environment.

So far, we have seen examples of the complex NP domain in relative clause constructions. However, the complex NP domain does not only comprise relative clause constructions, but also some others, for example, N-complement clause constructions (e.g. *the fact that...*, *the rumour that...*). The following are examples of such clause constructions embedding *caki*.

- (2.18) a. John<sub>i</sub>-i Mary<sub>j</sub>-lopwute [Tom<sub>k</sub>-i **caki**<sub>i/j/k</sub>-lul silhehanta]-nun  
 John-NOM Mary-from Tom-NOM **self**-ACC hate-COMPL  
 somwun-ul tulessta  
 rumour-ACC heard  
 ‘John<sub>i</sub> heard from Mary<sub>j</sub> the rumour that Tom<sub>k</sub> hate **self**<sub>i/j/k</sub>’

- b. John<sub>i</sub>-i Mary<sub>j</sub>-ekey [Tom<sub>k</sub>-i **caki**<sub>i/\*j/k</sub>-lul silhehanta]-nun  
 John-NOM Mary-DAT **self**-ACC hate-COMPL  
 somwun-ul malhayssta  
 rumour-ACC told  
 ‘John<sub>i</sub> told Mary<sub>j</sub> the rumour that Tom<sub>k</sub> hate **self**<sub>i/\*j/k</sub>’

As these examples of N-complement clause constructions illustrate, the same binding patterns obtain in these constructions. That is to say, *caki* embedded in the N-complement clause is bound by the higher subject *John* or the lower subject *Tom* as instances of subject-binding. Furthermore, in (2.18a), *Mary* in the *from* adjunct can bind *caki*, whereas *Mary* in (2.18b) cannot be an antecedent of *caki*.

So far, we have observed the binding patterns of *caki* in a range from the local domain to the extended domain where *caki* occurs in a deeply embedded clause. The embedded clauses include a plain complement clause as well as a clause with a complex NP. Throughout the set of data in these constructions, the subject acts as a primary antecedent of *caki*. However, we have seen cases of non-subject binding. This includes a specific adjunct phrase like *from NP* in contrast with adjunct phrases like *to NP*. Such non-subject binding patterns remain consistent from the local to the extended domain we have explored so far. We also have seen that the direct object can be the antecedent of *caki* in the local domain. It is not yet very clear whether the direct object-binding can extend to the extended domain since it is hard to find an appropriate example to illustrate this binding pattern. However, the following example shows the kind of configuration we have in mind:

- (2.19) John<sub>i</sub>-i Mary<sub>j</sub>-lul {caki<sub>i/j</sub>-uy emma-ka/ caki<sub>i/j</sub>-ka} sa-n cip-uro  
 John-NOM Mary-ACC self-GEN mother-NOM self-NOM buy-REL house-to  
 tolleypnayssta  
 returned  
 ‘John<sub>i</sub> returned Mary<sub>j</sub> to the house that {self<sub>i/j</sub>’s mother/self<sub>i/j</sub>} had just  
 bought’

We have also seen that there is a regular pattern in the non-subject antecedents and this may provide a significant base to hypothesise a binding principle, which may also cover subject-binding.

## 2.4 *Caki* in Adjunct Clauses

We now turn to slightly different examples from the previous ones. So far, *caki* has appeared in an embedded clause: a complement clause, a N-complement clause or a relative clause. However, the sentences in this section deal with *caki* in adjunct clauses. Adjunct clauses are different from the clauses we have seen so far in that the latter can be an argument of a verb, whereas the former are sentential adjuncts which means that they modify a complete sentence with all arguments saturated. At this stage, we can make a rough generalisation that *caki* in adjunct clauses may well exhibit different behaviour from that in embedded clauses as the two types of clause are distinct from each other in structure. Firstly, let us have a look at the following examples:

- (2.20) a. Mary<sub>i</sub>-ka [caki<sub>i/\*j</sub>-ka nara-lul ttenan-hwuey] John<sub>j</sub>-ekey ku  
 Mary-NOM self-NOM country-ACC left-after John-to the  
 sosik-ul malhayssta  
 news-ACC told  
 ‘Mary<sub>i</sub> told John<sub>j</sub> the news after self<sub>i/\*j</sub> left the country’
- b. Mary<sub>i</sub>-ka [caki<sub>i/\*j</sub>-ka nara-lul ttenan-hue] Bill<sub>j</sub>-lopwute  
 Mary-NOM self-NOM country-ACC left-after Bill-from  
 ku sosik-ul tulessta  
 the news-ACC heard  
 ‘Mary<sub>i</sub> heard the news from Bill<sub>j</sub> after self<sub>i/\*j</sub> left the country’

In Korean, the adjunct clause can appear anywhere except post-verbally (note that Korean is a head final language). In (2.20), the adjunct clause appears immediately after the subject and *caki* occurs in the adjunct clause. The given binding relations show that *caki* in these cases is only bound by the subjects, which was the case

with *caki*-binding in complement clauses. However, the point to note is that the adjunct NP does not bind *caki* in (2.20b), unlike the examples with complement clauses. For some linguists, for example Chung (1996), *caki* cannot be bound by *Bill* because the antecedent does not precede the anaphor<sup>3</sup>.

In order to assess whether linear precedence affects the binding patterns, in the following examples the adjunct phrases precede *caki*:

- (2.21) a. ‘Mary<sub>i</sub>-ka John<sub>j</sub>-ekey [**caki**<sub>i/\*j</sub> nara-lul ttenan-hwuey]  
 Mary-NOM John-to **self**-NOM country-ACC left-after  
 ku sosik-ul malhayssta  
 the news-ACC told  
 ‘Mary<sub>i</sub> told the news to John<sub>j</sub> after **self**<sub>i/\*j</sub> left the country’
- b. Mary<sub>i</sub>-ka Bill<sub>j</sub>-lopwute [**caki**<sub>i/\*j</sub>-ka nara-lul ttenan-hue] ku  
 Mary-NOM Bill-from **self**-NOM country-ACC left-after the  
 sosik-ul tulessta  
 news-ACC heard  
 ‘Mary<sub>i</sub> heard the news from Bill<sub>j</sub> after **self**<sub>i/\*j</sub> left the country’

However, even after relocating the adjunct phrases, the binding patterns still remain the same. Thus, we can assume that precedence in the surface order is not a factor here, although it may be in other situations. All we can generalise so far is that binding with adjunct clauses seems to allow subject-binding, but non-subject binding appears different between complement clauses and adjunct clauses.

The next thing we should look at in this section is different types of adjunct clause. In Japanese, different binding relations seem to be possible depending on different types of adjunct clauses as observed in Iida (1996):

---

<sup>3</sup>Chung (1996) proposes that anaphors like *caki* have to be preceded by their antecedents. This proposal is based on the Principle of Information Flow (Givón 1975, Kim 1985) which states that the constituents in a sentence tend to be sequentially ordered in such a way that a constituent expressing given information comes first, a constituent expressing new information next, and a constituent expressing unpredictable information last. This is similar to saying that other things being equal, a more prominent constituent (given information) linearly precedes a less prominent constituent. Chung (1996) argues that as for anaphor binding, the principle of information flow can account for the fact that the relative linear order of an anaphor and its antecedent affects binding possibilities, assuming that the antecedent needs to be more prominent than its anaphor.



- (2.22) a. \*Taroo<sub>i</sub>-ga [Hanako-ga **zibun**<sub>i</sub>-o hihansita toki] damatteita  
 Taroo-NOM Hanako-NOM **self**-ACC criticised when kept silence  
 ‘Taroo<sub>i</sub> said nothing when Hanako criticised **self**<sub>i</sub>’
- b. Taroo<sub>i</sub>-ga [Hanako-ga **zibun**<sub>i</sub>-o hihansita noni] damatteita  
 Taroo-NOM Hanako-NOM **self**-ACC criticised though kept silence  
 ‘Taroo<sub>i</sub> said nothing though Hanako criticised **self**<sub>i</sub>’

(2.22a) includes a temporal adjunct clause whereas (2.22b) has a concessive clause. The Japanese long-distance anaphor *zibun* in the temporal adjunct clause (2.22a) cannot be bound by the subject *Taroo*. However, *zibun* in the concessive clause (2.22b) can be bound by the subject *Taroo*. Thus we can see that there is a clear difference between the temporal and concessive clauses. However, the different types of adjunct clause do not seem to affect the binding of *caki* in Korean as shown in the following examples<sup>4</sup>:

- (2.23) a. Mary<sub>i</sub>-ka [John-i **caki**<sub>i</sub>-lul pipanhal-ttay] chimmuk-ul cikyessta  
 Mary-NOM John-NOM **self**-ACC criticised-when silence-ACC kept  
 ‘Mary<sub>i</sub> kept silence when John criticised **self**<sub>i</sub>’
- b. Mary<sub>i</sub>-ka [John-i **caki**<sub>i</sub>-lul pipanha-nunteto] chimmuk-ul cikyessta  
 Mary-NOM John-NOM **self**-ACC criticised-though silence-ACC kept  
 ‘Mary<sub>i</sub> kept silence though John criticised **self**<sub>i</sub>’

Korean thus differs from Japanese in that it makes no distinction between different types of adjunct clause in these circumstances.

To sum up, the adjunct clause behaves differently from other kind of complement clauses with respect to *caki*-binding. The first major finding is that non subject-binding breaks down. Secondly, all the binding relations we have seen in this section are the same irrespective of the type of adjunct clauses, which is not the case in Japanese.

<sup>4</sup>The local subject may also bind *caki*, although this is irrelevant here.



## 2.5 Genitive *Caki*

In introducing Korean reflexives, it was briefly mentioned that *caki* can also be case-marked with the genitive marker ‘-uy’. Genitive *caki* is just as frequent as nominative and accusative *caki* in Korean, and so its binding properties need to be described. Firstly, consider the following example:

- (2.24) John<sub>i</sub>-i    **caki<sub>i</sub>-uy** kabang-ul ilhepelyessta  
 John-NOM **self-GEN** bag-ACC lost  
 ‘John<sub>i</sub> lost **self<sub>i</sub>**’s bag’

In (2.24), genitive *caki* embedded in the direct object NP is bound by the local subject *John*. Also, genitive *caki* can take an antecedent outside the clausal domain as follows:

- (2.25) a. John<sub>i</sub>-i    Mary<sub>j</sub>-lopwute [**caki<sub>i/j</sub>-uy** uysa-ka    kapelyessta]-ko  
 John-NOM Mary-from    **self-GEN**    doctor-NOM has gone-COMPL  
 tulessta  
 heard  
 ‘John<sub>i</sub> heard from Mary<sub>j</sub> that **self<sub>i/j</sub>**’s doctor has gone’
- b. John<sub>i</sub>-i    Mary<sub>j</sub>-ekey [**caki<sub>i/\*j</sub>-uy** uysa-ka    kapelyessta]-ko  
 John-NOM Mary-DAT **self-GEN**    doctor-NOM has gone-COMPL  
 malhayssta  
 said  
 ‘John<sub>i</sub> said to Mary<sub>j</sub> that **self<sub>i/\*j</sub>**’s doctor has gone’

In (2.25), *caki* appears in the complement clause and may be bound by the higher subject *John*. The same non-subject binding that we have observed before pertains for genitive *caki*; that is, in (2.25a), the adjunct *from Mary* in the matrix clause can bind *caki* in the lower clause, whereas the complement PP *to Mary* in (2.25b) cannot. This suggests that genitive *caki* has the same binding pattern as both nominative and accusative *caki*.

There is also a pair of examples which are interesting as in (2.26) below:

- (2.26) a. [John<sub>i</sub>-uy mitum]-un [caki<sub>i</sub>-ka yongkamhata]-nunkes ita  
 John-GEN belief-NOM **self**-NOM brave-COMPL is  
 ‘John<sub>i</sub>’s belief is that **self<sub>i</sub>** is brave’
- b. \*[Caki<sub>i</sub>-uy mitum]-un [John<sub>i</sub>-i yongkamhata]-nunkes ita  
**self**-GEN belief-NOM John-NOM brave-COMPL is  
 ‘**self<sub>i</sub>**’s belief is that John<sub>i</sub> is brave’

These again are sentences involving a complement clause: *caki* is placed either in the subject position of the complement clause (2.26a) or in the matrix subject (2.26b). The only possible binding relation allowed is (2.26a) where *John* binds *caki* over into the complement clause, but not vice versa as in (2.26b). One may explain the impossible binding in (2.26b) by simply saying that the antecedent should precede *caki* or that *caki* is not allowed to occur as a part of the matrix subject.

However, matters become slightly more complicated when genitive *caki* appears in various types of adjunct phrases. Consider the following simple sentences:

- (2.27) a. Mary<sub>i</sub>-ka John<sub>j</sub>-ul [caki<sub>i/j</sub>-uy cip-ulo] tollyeponayssta  
 Mary-NOM John-ACC self-GEN house-to sent back  
 ‘Mary<sub>i</sub> sent John<sub>j</sub> back to **self<sub>i/j</sub>**’s house’
- b. Mary<sub>i</sub>-ka John<sub>j</sub>-ul [caki<sub>i/\*j</sub>-uy kapang-ulo] ttalyessta  
 Mary-NOM John-ACC self-ACC bag-with hit  
 ‘Mary<sub>i</sub> hit John<sub>j</sub> **with self<sub>i/\*j</sub>**’s bag’

We have already seen (2.27a) briefly in section (2.3.1) as an instance of non-subject binding. Both sentences in (2.27) have similar structures, and genitive *caki* is included in the adjunct. However, in (2.27a) *caki* is bound by either the subject or the direct object. In the other (2.27b), *caki* is bound by the subject, but not the direct object. The only conceivable difference between the two examples is that they have adjunct phrases that have different thematic roles. This factor may be also related to the discrepancy in the binding patterns like those we have seen in section (2.3).

The last set of examples we will consider in this section has to do with the problem of the locality of *caki* binding. Firstly, let us look at the following example:

- (2.28) John<sub>i</sub>-i [nay-ka **caki**<sub>i</sub>-uy apeci-la]-nunkes-ul an-i?  
 John-NOM I-NOM **self**-GEN father-be-COMPL-ACC know-Q  
 ‘Does John<sub>i</sub> know that I am self<sub>i</sub>’s father?’

In (2.28), *John*, the matrix subject is the only possible antecedent of *caki* as the lower subject is a first person pronoun, which does not agree with *caki* in  $\phi$ -features. However, even if the lower subject was a third person nominal, the local binding would still not be allowed as shown below<sup>5</sup>.

- (2.29) Mary<sub>i</sub>-ka [John<sub>j</sub>-i **caki**<sub>i/\*j</sub>-uy apeci-la]-nunkes-ul an-i?  
 John-NOM Mary-NOM **self**-GEN father-be-COMPL-ACC know-Q  
 ‘Does Mary<sub>i</sub> know that John<sub>j</sub> is [self<sub>i</sub>’s father]<sub>j</sub>?’  
 ‘Does Mary<sub>i</sub> know that John<sub>j</sub> is [self<sub>j</sub>’s father]<sub>j</sub>?’

Structurally, both the matrix subject *Mary* and the lower subject *John* are possible antecedents. However, the antecedency of *John* would result in a pragmatically incoherent interpretation (e.g., one cannot be one’s own father). Therefore, it seems that in such cases, semantics/pragmatics overrides the syntactic conditions.

In this section, it has been shown that the genitive *caki* behaves as *caki* in other cases: it has the same subject and non-subject binding patterns both in the local and extended domain.

## 2.6 Multiple occurrences of *Caki*

*Caki* can appear more than once in a sentence as below:

---

<sup>5</sup>(2.29) may be explained in terms of the *i-within-i* condition (Chomsky 1981). However, I am not sure how this could be dealt with within a lexicalist framework.

- (2.30) Mary-ka [John-i **caki-uy** pang-eyse **caki-uy** il-ul hayssta]-ko  
 Mary-NOM John-NOM **self-GEN** room-in **self-GEN** work-ACC did-COMPL  
 malhayssta  
 said  
 ‘Mary said John did **self’s** work in **self’s** room’

(Modified from Howard & Neyekawa-Howard (1976))

Here, *caki* appears once as a genitive reflexive of the direct object and once as a genitive reflexive in the locative PP; there are also two possible antecedents, *Mary* and *John*. We have seen already that *caki* can be bound by either the matrix subject *Mary* or the lower subject *John*. From this, then, there should be four ambiguous readings as follows:

- (2.31) a. Mary said that John did his work in his room  
 b. Mary said that John did her work in her room  
 c. Mary said that John did his work in her room  
 d. Mary said that John did her work in his room

In the first two readings (2.31a-b), each sentence has two *cakis* coindexed with the same antecedent, whereas in the other two (2.31c-d), each sentence has two *cakis* coindexed with different antecedents from each other. The latter case is not allowed in Korean. The same applies to simple sentences with multiple occurrences of *caki*.

- (2.32) John-i Mary-lul caki-uy cip-ulo caki-uy cha-lo tolye ponayssta  
 John-NOM Mary-ACC self-GEN house-to self-GEN car-by return sent  
 ‘John sent Mary to his house by his car’  
 ‘John sent Mary to her home by her car’  
 ‘\*John sent Mary to his house by her car’  
 ‘\*John sent Mary to her house by his car’

In (2.32), the subject as well as the direct object are possible antecedents as we have already seen in (2.27a). As can be seen in the readings in (2.32), the two

occurrences of *caki* have to be bound either by *John*, or by *Mary*. Therefore, there seems to be a constraint that multiple *caki*'s in a sentence should take the same antecedent.

## 2.7 *Caki* without an Overt Antecedent

So far all the examples we have looked at have an antecedent in the sentence and we have observed the relations between the antecedent and *caki*. Quite unlike such examples, however, there are instances where there is no overt antecedent as illustrated below:

- (2.33) a. **Caki**<sub>*i*</sub>-uy il-un      **caki**<sub>*i*</sub>-ka hay-ya      hanta  
**Self**'s      work-TOP self-NOM do-if only do  
'Everybody should do his own work'
- b. **Caki**<sub>*i*</sub>-uy kotong-un      **caki**<sub>*i*</sub>-ka anta  
**Self**'s      suffering-TOP self-NOM know  
'One knows one's own suffering'
- c. **Caki**<sub>*i*</sub>-lul cikhi-l      salam-un      **caki**<sub>*i*</sub>-ppwun ita  
**self**-ACC protect-REL person-TOP **self**-only      is  
(lit.) 'One who must protect oneself is only himself'  
'One has to protect oneself'

In (2.33) above, *caki* on its own or with multiple occurrences appears without any possible antecedent within the sentence, which is grammatical. This type of sentence raises the question of whether they should be described as cases of what is usually called discourse binding, in other words, the case where a reflexive element is used as a simple pronoun. The cases at hand turn out to be rather different. As the glosses on the examples indicate, *caki* in these sentences is not a referential pronoun.

Instead, it is rather a generic bound pronoun corresponding to English *one*. Observe further that this use of *caki* is restricted to sentences with present tense<sup>6</sup>:

- (2.34) \***Caki**<sub>i</sub>-uy kotong-un    **caki**<sub>i</sub>-ka al-ass-ta  
**Self**'s    suffering-TOP self-NOM know-PAST-DC  
 ‘\*One knew one’s own suffering’

Nonetheless, there are some examples which include *caki* without any overt antecedent within the sentence even when they are not used generically.

- (2.35) Ku kes    cengmal **caki** son-ulo    mantul-ess-ni?  
 the thing really    **self** hand-with make-PAST-Q  
 ‘Did (he) make it with **self**’s hands?’

In the example above, one may well assume that an antecedent does exist in the prior discourse or may be known by inference as has been traditionally believed in many studies of long-distance anaphora. In (2.35), however, given the fact that Korean is a pro-drop language, a possible explanation is that the subject of the sentence is understood to be present, but simply dropped. Therefore (2.35) can be treated just as other instances where *caki* is bound by the subject.

Taking the two instances above (the generic use of *caki* and the dropped antecedent) into consideration together, the following sentence should be considered carefully.

- (2.36) **Caki** chaykim-ul    nam-ekey tollici ma-ra  
**Self** responsibility-ACC other-to shift not-IMP  
 ‘Don’t shift **self**’s responsibility to somebody else!’

This is an example of an imperative, and the subject of the imperative is normally not expressed, which may look like a case of pro-drop. However, if we consider (2.36) a normal imperative sentence, it should be ungrammatical. The dropped

---

<sup>6</sup>We will not deal with those instances of *caki* in this thesis. However, we believe that they could be treated along the lines of the proposal made in Kural & Tsoulas (1999) concerning English generic pronouns, that is, they are unselectively bound by a generic operator along with the temporal variable of the sentence which explains the fact that in general generic sentences are not acceptable in the past tense; we will leave this issue open, however.

subjects in singular imperatives are generally second person and *caki* cannot take a second person antecedent. The grammaticality of the sentence implies that *caki* is used for generic purpose and that *caki* is used non-referentially in an imperative sentence. If the sentence was indeed used directly referring to the hearer, then *caki* should be replaced by a second person pronoun shown below:

- (2.37) Ne/\*caki chaykim-ul nam-ekey tolli-ci mara  
 You/\*self responsibility-ACC other-to shift-NOM stop  
 ‘Don’t shift your/\*self’s responsibility to somebody else!’

If *caki* were indeed bound by any discourse antecedent (as suggested by many discourse approaches), then it should apply to sentences where all the argument positions are filled (as opposed to (2.35)). One of the evident examples is as follows:

- (2.38) \*Caki-ka John-ekey teipu-lul chu-ess-ta  
 Caki-NOM John-DAT tape-ACC give-PAST-DC  
 ‘Self gave John a tape’

Here, *caki* is not used for generic purposes, nor is there a dropped noun as (2.38) is a fully saturated sentence. If *caki* was able to refer to an entity in the discourse level, the sentence should be grammatical, but it is not. This implies that *caki* has to have an antecedent which exists somewhere in the sentence level.

We have dealt with a set of examples where *caki* does not appear with any overt antecedent. One instance of such a case is when *caki* is used as a generic anaphor. Another is when an antecedent is not expressed as a case of pro-drop. Other than these two instances, *caki* does not seem to be allowed to appear without an overt antecedent.

## 2.8 Topic Antecedents

In this section, we will examine whether topics can be possible antecedents of *caki*. While many subjects function as topic, there are topics which are not subjects.

These non-subject topics may also serve as antecedents, as illustrated below:

- (2.39) Mary<sub>i</sub>-**nun** caki<sub>i</sub>-ka ceyil silhehanta  
Mary-TOP **self**-NOM most dislike  
'As for Mary<sub>i</sub>, self<sub>i</sub> dislikes (herself<sub>i</sub>) most'

The example above has *caki* in the subject position bound by the topic *Mary*. Here the most plausible status of *Mary* would be as a topicalised direct object, which has moved to the beginning of the sentence and been topic marked with '-*nun*'. The evidence for topicalisation movement is that if we assume that the topic *Mary* is an independent constituent of the sentence and that the direct object of the verb is simply dropped (like a dislocation structure), this *should* make the following examples grammatical.

- (2.40) \*Mary<sub>i</sub>-**nun** caki<sub>i</sub>-ka kunye<sub>i</sub>-lul ceyil silhehanta  
Mary-TOP **self**-NOM *her* most dislike  
'As for Mary, self<sub>i</sub> dislikes her<sub>i</sub> most'

What we did in this example is just recover the potentially dropped pronoun in the direct object position. However, this turns out to be ungrammatical. The direct object *her* is ungrammatically bound by the local subject *caki* as well as the topic *Mary*.

The following sentence truly shows an example where *caki* is bound by the topic which stands independently of the sentence. That is, in (2.41), the topic is not derived by movement since all argument positions of the verb are already saturated and we will call this type of topicalisation *Gapless topic constructions*.

- (2.41) Mary<sub>i</sub>-**nun** tongsaying-i caki<sub>i</sub>-pota te ttokttokhata  
Mary-TOP sibling-NOM **self**-than more smart  
'As for Mary<sub>i</sub>, her younger sibling is smarter than **self**<sub>i</sub>'

Here, we have a full sentence *Her younger sibling is smarter than self (caki)* and *caki*, the object of the adjectival phrase, is bound by the sentential topic *Mary*.



In the English interpretation of this example, what the topic binds is generally expressed by a pronoun as follows:

(2.42) ‘As for Mary, her younger sibling is smarter than **her**’

In Korean, on the other hand, the bindee of the topic should be always expressed by reflexive *caki* and a pronoun is not allowed:

(2.43) Mary<sub>i</sub>-**nun** tongsay<sub>i</sub>-i (\***kunye<sub>i</sub>/caki<sub>i</sub>**)-pota te ttokttokhata  
 Mary-TOP sibling-NOM **her/self**-than more smart  
 ‘As for Mary<sub>i</sub>, her younger sibling is smarter than (\***her<sub>i</sub>/self<sub>i</sub>**)’

- (2.44) a. \*Ku cangkwaun<sub>i</sub>-**un** \***ku<sub>i</sub>/caki<sub>i</sub>**-ka momso cakcen-ul  
 The general-TOP **he/self**-NOM personally operation-ACC  
 cihwihayssta  
 commanded  
 ‘As for the general<sub>i</sub>, \***he<sub>i</sub>/self<sub>i</sub>** commanded the operation in person’  
 (or, ‘The general himself commanded the operation in person’)
- b. Sensayingnim<sub>i</sub>-**un** \***ku<sub>i</sub>/caki<sub>i</sub>**-ka motun chaykim-ul  
 Teacher-TOP **he/self**-NOM all responsibility-ACC  
 ttemathassta  
 took over  
 ‘As for the teacher<sub>i</sub>, \***he<sub>i</sub>/self<sub>i</sub>** took over all the responsibility’

Considering both gapped and gapless topic constructions together, it may be worth looking at whether *caki* itself can be topicalised. Firstly consider the following pair of examples, where (2.45b) is a gapped topic constructions:

- (2.45) a. Mary<sub>i</sub>-ka [John<sub>j</sub>-i caki<sub>i/j</sub>-lul ceyil cohahanta]-ko sayngkakhanta  
 Mary-NOM John-NOM self-ACC best like-COMPL thinks  
 ‘Mary<sub>i</sub> thinks that John<sub>j</sub> likes self<sub>i/j</sub> the best’
- b. Caki<sub>i/\*j</sub>-**nun** Mary<sub>i</sub>-ka [John<sub>j</sub>-i ceyil cohahanta]-ko sayngkakhanta  
 self-TOP Mary-NOM John-NOM best like-COMPL thinks  
 ‘It is herself that Mary thinks that John likes the best’

In (2.45a), *caki* is the lower object in the complement clause. When *caki* is topicalised by movement, which results in (2.45b), not only does it show that *caki* itself can be licensed as a topic, but it also shows that only the long distance binding of *caki* is possible, that is, *caki* is not bound by the local subject *John*. On the other hand, this is not quite the same in gapless topic constructions. Consider the following:

- (2.46) a. Mary<sub>i</sub>-un ayin<sub>j</sub>-i caki<sub>i/j</sub>-uy cip-eyse cenyek-ul cwunpihayssta  
 Mary-TOP lover-NOM self-GEN house-at dinner-ACC prepared  
 ‘Mary<sub>i</sub>’s lover<sub>j</sub> prepared dinner at self<sub>i/j</sub>’s house’
- b. \*Caki-nun ayin-i John-uy cip-eyse cenyek-ul cwunpihayssta  
 self-TOP lover-NOM John-GEN house-at dinner-ACC prepared  
 ’\*Self’s lover prepared dinner at John’s house’

(2.46a) is an instance of a gapless topic construction and genitive *caki* occurs in the adjunct phrase and is bound by either the topic *Mary* or the subject *ayin*. (2.46b) has the same structure as (2.46a) but (2.46b) has *caki* in the topic position. Having *caki* itself as a topic in (2.46b) makes the sentence ungrammatical. One may argue that this is probably because *caki* cannot appear in sentence initial position. However, we have seen many instances where topic-marked *caki* can occur at the beginning of the sentence, such as (2.45b).

The other interesting aspect of a topic phrase in relation to *caki* is that the topic seems to behave as an optimal antecedent of *caki* no less than the subject. One such example can be taken from the cases where the nominal in the complement PP (i.e., *to NP*) could not bind *caki*:

- (2.47) Mary<sub>i</sub>-ka Bill<sub>j</sub>-ekey [John<sub>k</sub>-i caki<sub>i/k</sub>-lul coahanta]-ko malhayssta  
 Mary-NOM Bill-to John-NOM self-ACC like-COMPL said  
 ‘Mary<sub>i</sub> said to Bill<sub>j</sub> that John<sub>k</sub> likes self<sub>i/k</sub>’

In (2.47), *Bill* does not bind *caki*. However, once it is topicalised, it becomes a legitimate antecedent of *caki* as below.

(2.48) Bill<sub>j</sub>-ekey-**nun** Mary<sub>i</sub>-ka [John<sub>k</sub>-i **caki**<sub>i/j/k</sub>-lul coahanta]-ko malhayssta  
 Bill-to-**top** Mary-NOM John-NOM **self**-ACC like-COMPL said

‘ To Bill<sub>j</sub>, Mary<sub>i</sub> said that John<sub>k</sub> likes **self**<sub>i/j/k</sub>’

Here, the dative *Bill* is moved to the initial position of the sentence and topic-marked with *nun* and is now able to bind *caki*. However, if an object is topicalised by being moved further away, the topic does not seem to antecede *caki* any more. Compare the following examples with (2.48):

(2.49) a. Anna<sub>i</sub>-ka [Mary<sub>j</sub>-ka Bill<sub>k</sub>-ekey [John<sub>l</sub>-i **caki**<sub>i/j/l</sub>-lul coahanta]-ko  
 Anna-NOM Mary-NOM Bill-to John-NOM **self**-ACC like-COMPL  
 malhayssta]-ko mitnunta

said-COMPL believes

‘Anna<sub>i</sub> believes that Mary<sub>j</sub> said to Bill<sub>k</sub> that John<sub>l</sub> likes **self**<sub>i/j/l</sub>’

b. Bill<sub>k</sub>-ekey-**nun** Anna<sub>i</sub>-ka [Mary<sub>j</sub>-ka [John<sub>l</sub>-i **caki**<sub>i/j/l</sub>-lul coahanta]-ko  
 Bill-to-**TOP** Anna-NOM Mary-NOM John-NOM **self**-ACC like-COMPL  
 malhayssta]-ko mitnunta

said-COMPL believes

‘ To Bill<sub>k</sub>, Anna<sub>i</sub> believes that Mary<sub>j</sub> said that John<sub>l</sub> likes **self**<sub>i/j/l</sub>’

In (2.49a), the dative *Bill* is now within the complement clause instead of the matrix clause. The dative *Bill* here does not antecede *caki* in the second complement clause. As we saw that topicalisation of the dative *Bill* will allow it to bind *caki*, we topicalise *Bill* in (2.49b) by moving it to the beginning of the sentence across the complement clause rather than within its own clause boundary as in (2.48). While this movement seems to be allowed, such topicalisation does not seem to make any difference to the binding relation. That is, even though *Bill* is topicalised, it still cannot bind *caki*. If we look at the movement for topicalisation and the binding relations both in (2.48) and (2.49b), we can note the clear relation between movement and binding such that when topicalisation takes place locally, the topicalised object can antecede *caki* as in (2.48). On the other hand, when topicalisation involves long distance movement (across a clausal boundary), then the topicalised object is not able to bind *caki*.

In this section, we have observed topic-hood with respect to *caki*-binding. Firstly, two different topic constructions were introduced: one is a gapped topic construction resulting from extraction and the other is a gapless topic construction which does not involve extraction, thus does not leave any gap. We have also found that in the former, *caki* itself can be topicalised and in this case, only the long distance of *caki*-binding is licensed, but not the local binding. As for the latter, on the other hand, *caki* does not seem to function as a topic at all, and this makes the sentence ungrammatical. Another peculiar fact about the gapped topic constructions is that even if an object does not bind *caki*, once it is topicalised, it becomes able to bind *caki*. However such cases are allowed only when the topicalised object remains within the clause. A topic element which is extracted across a clausal boundary cannot bind *caki* even if it could do so before topicalisation took place.

## 2.9 *Caki* vs. Pronouns

There has been some general discussion of whether a long-distance anaphor is not actually a reflexive and should be treated as a pronoun. This question originates in the fact that a long-distance reflexive such as *caki* in Korean can be bound by an antecedent outside the immediate clause boundary. Therefore, there can be two stories about Korean *caki*. That is to say, there are two kinds of *cakis*: one of which should be categorised as a pure reflexive when it is bound within a syntactic local domain, and the other is that long-distance anaphor ‘caki’ should be categorised as a pronoun in that it is bound by an antecedent outside a local domain and remains free in that local domain.

One example of such an idea is shown in Sohn (1994), who suggests the following:

‘As we have seen Korean reflexives occur in direct or indirect position, as well as in possessive and other case positions, Korean reflexive forms may also occur as the subject of a clause. The reason for this is obviously that the major condition on the use of a reflexive pronoun is determined at the discourse level as well as at the sentence level in a discourse-oriented

language like Korean. When used as a possessor or a subject, a Korean reflexive pronoun is rendered as a personal, not reflexive, pronoun in English'

(Sohn 1994, p.153)

Following this suggestion, let us first consider the example below:

- (2.50) a. John<sub>i</sub>-un [nay-ka Tom<sub>j</sub>-eykey **caki<sub>i</sub>/ku<sub>i/j/k</sub>**-uy sacin-ul  
 John-NOM I-NOM Tom-DAT **self/he**-GEN picture-ACC  
 poyecwue-yahanta]-ko malhayssta  
 show-should-COMPL said  
 'John<sub>i</sub> said that I should show **self<sub>i</sub>'s/his<sub>i/j/k</sub>** picture to Tom<sub>j</sub>'
- b. John<sub>i</sub>-un [**caki<sub>i</sub>/ku<sub>i/j</sub>**-ka salang-ey ppacyessta]-ko malhayssta  
 John-NOM **self/he**-NOM love-in fell-COMPL said  
 'John<sub>i</sub> said that **self<sub>i</sub>/he<sub>i/j</sub>** fell in love'

In (2.50a), genitive *caki* appears in a lower object position and is bound by the matrix subject *John* and this is exactly what happens when the pronoun occurs in place of the reflexive *caki*, though the antecedent of the pronoun *ku* can also be the indirect object or some referent in discourse apart from the matrix subject whereas the reflexive *caki* should have *John* as the only antecedent. The same applies to (2.50b) as well, where both *caki* and the pronoun occur in the lower object position; both can be bound by the matrix subject but the pronoun can also be bound by some referent not mentioned in the sentence (indicated by *j*). However, there are quite a few facts which suggest that the above approach may not be correct. Firstly, subject *caki* should behave in the same way as a pronoun. However, the following examples show that this is not the case.

- (2.51) a. \***Caki**-(ka/nun) kal-kessita  
**Self**-NOM/TOP go-will  
 'Self will go'

- b. **Ku**-(ka/nun) kal-kessita  
**He**-NOM/TOP go-will  
 ‘**He** will go’

(2.51a), which has the subject *caki* in a simple clause, is ruled out, whereas the counterpart with the pronoun subject is grammatical in (2.51b). The examples also show that when they are structurally the same, the pronoun can be topic-marked whereas *caki* cannot. One suggestion may be that Sohn’s analysis applies only to a sentence which has an embedded clause, as below:

- (2.52) a. \*Na-nun [**caki**-ka ka-l]-kelako malhayssta  
 I-NOM **self**-NOM go-will-COMPL said  
 ‘I said that **self** will go’
- b. Na-nun [**ku**-ka ka-l]-kelako malhayssta  
 I-NOM **he**-NOM go-will-COMPL said  
 ‘I said that **he** will go’

(2.52a) is again ruled out where *caki* appears in the lower subject position, whereas the same sentence with a pronoun is fine. Clearly, both *caki* and *ku* (*he*) look for the antecedent in some discourse domain, and *ku* is successfully bound, but *caki* is not. The same distinction also applies to the genitive reflexive vs. pronoun as shown below:

- (2.53) a. \*Na-nun [ney-ka **caki-uy** il-ul cemkemhayssta]-nunkes-ul  
 I-NOM [you-NOM **self**-GEN work-ACC examined-COMPL-ACC  
 anta  
 know  
 ‘I know that you examined **self**’s work’
- b. Na-nun [ney-ka **ku-uy** il-ul cemkemhayssta]-nunkes-ul anta  
 I-NOM [you-NOM **he**-GEN work-ACC examined-COMPL-ACC know  
 ‘I know that you examined **his** work’

What we can conclude from the examples is that the long-distance anaphor *caki* in Korean should not be categorised as a pronoun simply because it is bound by an antecedent outside the local domain and remains locally free. This would ignore the significant point that *caki* is somehow constrained in that its antecedent should be present within the sentence, whereas this restriction does not apply to pronouns.

Yet, there seem to be more things to be said about this restriction on *caki* other than the constraint that an antecedent must be present in the sentence. This need for more elaborate constraints on *caki* can be seen from examples such as the following:

- (2.54) a. \*Nay-ka [John<sub>i</sub>-i ilhepin kapang]-ul **caki<sub>i</sub>-ekey** chaca cwuessta  
 I-NOM John-NOM lost bag-ACC **self-DAT** find gave  
 ‘I found and gave **self<sub>i</sub>** the bag that John<sub>i</sub> lost’
- b. Nay-ka [John<sub>i</sub>-i ilhepin kapang]-ul **ku<sub>i</sub>-ekey** chaca cwuessta  
 I-NOM John-NOM lost bag-ACC **he-DAT** find gave  
 ‘I found and gave **him<sub>i</sub>** the bag that John<sub>i</sub> lost’

In (2.54a) and (2.54b), the sentences have the same structure, but (2.54a) includes the dative *caki* and (2.54b) the dative pronoun. In (2.54b), the pronoun has the antecedent *John* which is inside the relative clause. In (2.54a), the reflexive *caki* is not coindexed with anything (either inside or outside the sentence). This clearly shows that the long-distance anaphor *caki* should not be categorised as a pronoun, and also that further factors influence binding phenomena with *caki*.

## 2.10 Summary of the Facts

We have seen the behaviour of *caki* in various contexts. The following summarises the observations:

1. *Caki* is third person singular so only takes a third person singular antecedent.
2. *Caki* can take an antecedent in a local domain as well as outside the local clausal boundary: domains outside the local clausal boundary include complement clauses, relative clauses and N-complement clauses. Both in the



local domain and in the extended domain, there are regular binding patterns: firstly, subject binding seems to be possible in all environments. Secondly, as for the non-subject binding, certain adjuncts (e.g., *from NP* but not *to NP*) as well as the direct object may bind *caki* as long as the binding does not produce any pragmatic incoherence.

3. *Caki* in an adjunct clause exhibits different binding relations from *caki* in complement clauses. As for the subject-binding, *caki* can be bound by subjects. However, when it comes to non-subject binding, *caki* seems to resist binding to a non-subject. For example, *from NP* cannot antecede *caki* in adjunct clauses, unlike in complement clauses. In other words, the binding pattern of *caki* in adjunct clauses is distinct from that in complement clauses. We also observed that although in Japanese different types of adjunct clauses seem to show different patterns of subject-binding, *caki*-binding does not seem to be affected by different types of adjunct clauses.
4. Genitive *caki* seems to behave in the same way as nominative and accusative *caki*.
5. *Caki* can occur more than once in a sentence. In such cases, multiple *cakis* must be bound by the same antecedent.
6. In cases where the antecedent does not appear in the sentence, there are two possible ways to recover the antecedent. One is when *caki* is used to create a generic interpretation in the same way as English pronouns can function. The other way is that *caki* may well be bound by a dropped nominal (*pro*), which is a case of sentence-internal binding. Other than these two cases, the antecedent should be overtly expressed in the sentence.
7. A topic can act as an antecedent of *caki*. In gapped topic constructions, even if a nominal is not a possible antecedent, once it is topicalised, it becomes an legitimate antecedent of *caki*. However, this is limited to constructions in which the topicalised expression remains within the same clause as the reflexive. Any topic moved beyond the local clause cannot bind *caki*. Also, in this topic construction, *caki* can be topicalised itself, allowing only the long-distance binding case of *caki*.



8. We also have observed the distribution of *caki* and pronouns. The only conclusion that it is possible to draw is that *caki* cannot be categorised as a pronoun.

So far, we have observed the data on *caki* with most structural environments. A number of excellent works have attempted to provide an account for all or part of the data that we have considered. The accounts provided so far combine elements from different parts of the grammar (syntax, semantics and pragmatics). With such a great deal of attention in the linguistic literature on this type of phenomenon, it is quite a daunting task to take up the general topic without running the risk of merely recapitulating earlier scholarship. However, armed with an objective and comprehensive review of the empirical domain, we will try to focus on the binding patterns in different structural environments at the same time. Our aim will be to understand the mechanisms at play in Korean grammar, and possibly universally.

## Chapter 3

# A Critical Review of Approaches to Long-Distance Anaphora (LDA)

When long-distance anaphora started receiving linguists' attention in the early 60's, even though its behaviour was found to deviate from what we know about reflexive pronouns, the initial explanation looked rather simple, such as the subjecthood condition which we will see in the next section. Although this account is widely accepted today, once it is investigated more extensively the issue turns out to be more complicated. Just as the phenomenon of long-distance anaphora is complex, so the treatments that it has received are probably among the most intricate, involving every aspect of grammar: syntax, semantics and pragmatics. In a nutshell, the topic is an old one, but the answer to the phenomenon is not yet resolved.

There are two main groups of approaches to long-distance anaphora: syntactic and non-syntactic approaches. The former include the subjecthood condition as well as the c-command condition. The latter, on the other hand, involve semantic notions such as a thematic hierarchy as well as purely discourse notions like point of view or logophoricity, among many others. In this chapter, we will overview the two groups of approaches with regard to *caki*-binding and see whether they are applicable and adequate in accounting for the facts of *caki*-binding which we

observed in the last chapter. We will then proceed to consider views according to which a combination of syntactic and non-syntactic factors must be taken into consideration in the account of long-distance anaphora. There are two main currents within this type of approach. According to the first one long-distance anaphora can be licensed if either the relevant syntactic condition is satisfied or else if the relevant non-syntactic condition is satisfied. This has been dubbed the “disjunctive theory”. On the other hand, what is usually known as the “conjunctive theory” claims that both syntactic and non-syntactic conditions must be satisfied in order for long-distance anaphora to be licensed.

### 3.1 Syntactic Approaches

#### 3.1.1 The Subjecthood Condition

Kuroda (1965) first investigated Japanese reflexivization and concluded that the antecedent of the reflexive must be the subject of a sentence. For the Japanese long-distance anaphor, *zibun*, the subjecthood condition amounts to the requirement that the reflexive refers back to the subject in the same simplex sentence or the subject in any higher sentence. This condition is supported by Akatsuka (1971) and Kuno (1973) among many others. It has been widely claimed that Korean is also subject to this condition (Fiengo & Kim 1981, Yang 1983).

- (3.1) a. John<sub>i</sub>-i Bill<sub>j</sub>-ekey caki<sub>i/\*j</sub>-uy sacin-ul poyecwuessta  
 John-NOM Bill-DAT self-GEN picture-ACC showed  
 ‘John<sub>i</sub> showed Bill<sub>j</sub> a picture of self<sub>i/\*j</sub>’
- b. John<sub>i</sub>-i Bill<sub>j</sub>-ekey [Mary<sub>k</sub>-ka caki<sub>i/\*j/k</sub>-lul choahanta]-ko malhayssta  
 John-NOM Bill-DAT Mary-NOM self-ACC like-COMPL said  
 ‘John<sub>i</sub> said to Bill<sub>j</sub> that Mary<sub>k</sub> liked self<sub>i/\*j/k</sub>’

In both sentences above, the subjects *John* and *Mary* can serve as antecedent of *caki* unlike the non-subject constituent, the dative *Bill*. Furthermore, as pointed out by Kuroda, the domain of Japanese reflexivization is not confined to the simplex

sentence. It acts like English simple pronominalization, extending to relative clauses and sentences in apposition. This is also true for *caki*-binding as we have seen already. An example of a noun complement clause is illustrated below:

- (3.2) John<sub>i</sub>-i [caki<sub>i</sub>-ka Mary-lul cwuki]-nun kkwum-lul kwuessta  
 John-NOM selfNOM Mary-ACC is killing-REL dream-ACC dreamt  
 ‘John<sub>i</sub> dreamt the dream that self<sub>i</sub> was killing Mary’

It has been noted, however, that there are many exceptions to the subjecthood condition. McCawley (1976) provides a list of cases of non-subject binding in Japanese. Causative constructions belong to such cases, as the following illustrates:

- (3.3) Hiroshi<sub>i</sub> wa Michiko<sub>j</sub> o zibun<sub>i/j</sub> no heya de benkyoo-sase-ta  
 Hiroshi TOP Michiko ACC self GEN room in study-make-PAST  
 ‘Hiroshi<sub>i</sub> made Michiko<sub>j</sub> study in self<sub>i/j</sub>’s room’

(McCawley 1976, p54, ex.17)

In (3.3), *zibun* can refer both to the subject *Hiroshi* and the direct object *Michiko*. The binding relation of the latter obviously violates the subjecthood condition. While Korean causative constructions exhibit the same binding relations as in (3.4a), such non-subject binding relations in other constructions in Korean are also observed (3.4b):

- (3.4) a. John<sub>i</sub>-i Mary<sub>j</sub>-lul caki<sub>i/j</sub>-uy chayk-ul ilk-hi-essta  
 John-NOM Mary-ACC self-GEN book-ACC read-make-PAST  
 ‘John<sub>i</sub> made Mary<sub>j</sub> read self<sub>i/j</sub> book’  
 b. John<sub>i</sub>-i Mary<sub>j</sub>-ekey caki<sub>i/j</sub>-uy pang-lul chiwulako malhayssta  
 John-NOM Mary-DAT self-GEN room-ACC to clean told  
 ‘John<sub>i</sub> told Mary<sub>j</sub> to clean self<sub>i/j</sub>’s room’

In the causative construction (3.4a), *caki* is ambiguously bound either by the subject *John* or the direct object *Mary*. A non-causative construction like (3.4b) shows the same binding pattern. McCawley (1976) attempts to reduce these cases of non-subject binding to the subjecthood condition. Noting that sentences like (3.4a)

and (3.4b) share the same complex deep structure, she provides a constraint for such main verbs which states that the subject of the complement clause must be coreferential with the indirect object. For example, she assumes that the following deep structure underlies (3.4b):

(3.5) 'John<sub>i</sub> told Mary<sub>j</sub> (Mary<sub>j</sub> clean self<sub>i/j</sub>'s room)'

In this deep structure representation, the subjecthood condition is satisfied given that *Mary* appears in the subject position of the complement clause<sup>1</sup>. Although it seems to capture the generalisation of the subjecthood condition, it does not provide a good answer to the case of non-subject binding we saw in the last chapter. We have already seen that when *caki* is bound outside the local domain, *from* adjuncts can antecede *caki* whereas dative NPs cannot. Consider the following:

(3.6) Mary<sub>i</sub>-ka John<sub>j</sub>-uropwute [**caki**<sub>i/j</sub>-ka Bill<sub>k</sub>-lul hyemohanta]-ko tulessta  
 Mary-NOM John-from self-NOM Bill-ACC hate-COMPL heard  
 'Mary<sub>i</sub> heard from John<sub>j</sub> that self<sub>i/j</sub> hates Bill<sub>k</sub>'

(As (2.14b) in section 2.3.2)

As McCawley proposed the deep structure constraint for specific verbs such as causative verbs or verbs like 'tell' or 'order', (3.6) is not within the domain of such a constraint. The equivalent of (3.6) in Japanese is also grammatical. As for (3.5), moreover, the credibility of this account is undermined by the same type of argument that was initially used against the treatment of control constructions by Equi-NP deletion<sup>2</sup>. Unless the deep structure constraint is extended to capture some cognitive or discourse notion, it is hard to see how sentences like (3.6) can be explained in terms of the subjecthood condition.

<sup>1</sup>Surface structure 'John told Mary to clean self's room' is derived from (3.3) by the successive application of reflexivization (in the lower clause), Equi-NP deletion, deleting the lower clause subject, and finally complementizer insertion (McCawley 1976, pp55-57).

<sup>2</sup>This argument is based on the fact that the meaning of 'everyone wants to go' is not what would be expected under an Equi-NP deletion analysis. This type of fact led to the positing of an empty element PRO as the subject of infinitivals. Notice that under this type of account, the subjecthood condition is still satisfied. Interestingly though it is only satisfied at S-structure rather than at Deep structure, or in any event after the control rule has applied.

### 3.1.2 The Basic C-commanding Requirement

When the subjecthood condition was widely discussed, some notion of ‘command’ already played a role in reflexivization, such that an antecedent must command a reflexive (known as the antecedent-command condition). The notion of ‘command’ states that a node A ‘commands’ B if (1) neither A nor B dominates the other and (2) the S node that most immediately dominates A also dominates B (Langacker 1969). Subsequently, Reinhart (1976) provides a more general definition of ‘command’, namely, ‘c-command’ which is stated as follows:

- (3.7) X c-commands Y iff neither X nor Y dominates the other and the first branching node dominating X also dominates Y.

Hong (1985) highlights the role of c-command in accounting for *caki*-binding and claims that in Korean *any* c-commanding NP can be a binder of a reflexive. In terms of the notion of c-command, she also emphasises the significance of the pure structural notion for Korean anaphoric binding over the grammatical relation. That is, for her, the c-command condition replaces the subjecthood condition. This is illustrated below:

- (3.8) a. John<sub>i</sub>-i Bill<sub>j</sub>-ekey [caki<sub>i/\*j</sub>-ka toktokhata]-ko malhayssta  
John-NOM Bill-DAT self-NOM is clever-COMPL said  
‘John<sub>i</sub> told Bill<sub>j</sub> that self<sub>i/\*j</sub> is smart’  
b. John<sub>i</sub>-i Bill<sub>j</sub>-ekey caki<sub>i/\*j</sub>-uy erinsicel-ul iyakihayssta  
John-NOM Bill-DAT self-GEN childhood-ACC told  
‘John<sub>i</sub> told Bill<sub>j</sub> of self<sub>i/\*j</sub>’s childhood’

Hong (1985, p.19, ex.13a,b)

Both sentences are familiar in that the dative *Bill* does not antecede *caki*. Hong suggests that *Bill* does not c-command *caki* as it is not an immediate constituent of VP, but is inside a PP. Thus, c-command seems to explain *caki*-binding without relying on the subjecthood condition. Furthermore, the condition seems to work also for the case of direct object binding we observed in section 2.3.1:

- (3.9) John<sub>i</sub>-i Mary<sub>j</sub>-lul caki<sub>i/j</sub> cip-ulo tolye ponayssta  
 John-NOM Mary-ACC self house-to return sent  
 ‘John<sub>i</sub> returned Mary<sub>j</sub> to self<sub>i/j</sub>’s house’

(As (2.7) in section 2.3.1)

Here, *caki* is inside the PP *caki cip-ulo* (self’s house) and this PP is subcategorised by the verb, thus inside the VP. Within this VP environment, the direct object *Mary* does c-command and antecede *caki* without any problem.<sup>3</sup> The same applies to the subject *John*. However, this condition faces problems. Firstly, as noted by Hong, the c-command condition alone cannot account for backward reflexivization. Consider the following example:

- (3.10) [Caki<sub>i</sub>-ka Mary-lul ttaylyessta]-nunkes-i John<sub>i</sub>-ul sulpukey hayssta  
 self-NOM Mary-ACC hit-COMPL John-ACC sad made  
 ‘That self<sub>i</sub> hit Mary made John<sub>i</sub> sad’

Most cases of *caki*-binding are licensed when the antecedent precedes *caki* except for special cases such as a topic or focus (see section 2.4). Sentences like (3.10) belong to such special cases, namely, backward reflexivization. The structure of (3.10) is as follows:

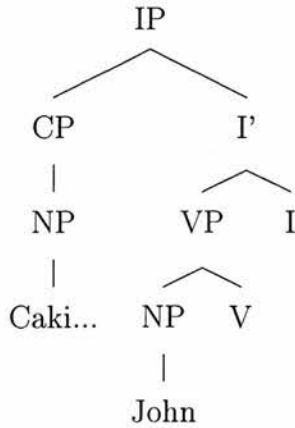
---

<sup>3</sup>When the PP is not subcategorised by the verb as below, the object cannot antecede *caki* because the object does not c-command it. We assume that this VP is right adjoined to the VP.

i John<sub>i</sub>-i Bill<sub>j</sub>-ul caki<sub>i/\*j</sub>-uy cip-ese mannassta  
 John-NOM Bill-ACC self-GEN house-at met  
 ‘John<sub>i</sub> met Bill<sub>j</sub> at self<sub>i/\*j</sub>’s house’

See Hong (1985) for further discussion of this kind of data.

(3.11)



In (3.11), *John* does not c-command *caki*, however the intended coreferential reading is possible and the sentence is grammatical. The case of *caki* being bound by a non c-commanding antecedent can be also found in the examples we saw as cases of non-subject binding in the last section, repeated below:

(3.12) Mary<sub>i</sub>-ka John<sub>j</sub>-uropwute [**caki**<sub>i/j</sub>-ka Bill<sub>k</sub>-lul hyemohanta]-ko tulessta  
Mary-NOM John-from self-NOM Bill-ACC hate-COMPL heard  
'Mary<sub>i</sub> heard from John<sub>j</sub> self<sub>i/j</sub> hates Bill<sub>k</sub>'

(As (3.6) in section 3.1.1)

Here, the nominal inside the PP adjunct, *from John*, can bind *caki* even though it is neither the subject nor c-commands *caki*. Therefore, it seems that the notion of c-command itself is not sufficient to explain *caki*-binding. It may be argued at this point that a different command relation may be used to explain this data<sup>4</sup>. One such relation is 'm-command' which can be stated as follows:

(3.13) A m-commands B if A does not dominate B and every maximal projection that dominates A also dominates B.

---

<sup>4</sup>though m-command works no better than c-command



Strictly speaking, however, even this notion will not be of any help in cases where the antecedent is inside a PP given that the first maximal projection dominating the antecedent (which is an NP) will be the PP as in (3.12). This definition will not help either in cases of backward reflexivization given that VP does not dominate the subject<sup>5</sup>. It seems then that it is not the definition of the relevant command relation that is at stake here. Rather, at least for the cases of the antecedent within PPs, it is the status of the prepositions and their phrases that should be reconsidered. This was the intuition of Reinhart (1976) that in essence, PP nodes do not count for the calculation of c-command domains, which is empirically sound but theoretically ad hoc.

### 3.1.3 Approaches in GB

The starting point of most current discussion of anaphora in Government and Binding theory (GB) is developed by Chomsky in a series of works from the early 80's (Chomsky 1980, Chomsky 1981, Chomsky 1982). The theory is based on the classification of noun phrases using the two binary feature [a] (anaphor) and [p] (pronominal). The features [a] and [p] indicate whether elements should be bound or must not be bound, in a certain local domain.

- (3.14)
- a. An *anaphor* is [a+], [p-]
  - b. A *pronominal* is [a-], [p+]
  - c. An *R-expression* is [a-], [p-]

A definition of what it means to be bound follows:

- (3.15) An NP A is bound if and only if there is an NP B such that both of the following conditions are satisfied:
- (a) A and B are coindexed

---

<sup>5</sup>This difficulty could, of course, be overcome if we adopted the VP internal subject hypothesis and we also applied it to clausal subjects. There is, however, little empirical evidence, independent of the assumption that  $\theta$ -marking takes place under government which in turn requires m-command, to suggest that the subject is internal to the VP. This is even less clear for clausal subjects.

(b) B c-commands A

And the binding domain is defined as the *governing category* as in (3.16):

(3.16) The governing category for  $\alpha$  is the minimal clause or NP containing  $\alpha$ , a governor for  $\alpha$ , and a subject

Now, each type of NP has different binding conditions in terms of the governing category:

- (3.17)
- Principle A: An *anaphor* is bound in its governing category
  - Principle B: A *pronominal* is free (= not bound) in its governing category
  - Principle C: An *R-expression* is free.

Binding Principle A given in (3.17) accounts for a wide range of phenomena with respect to the distribution of reflexives. In English, for example, consider the following:

- (3.18)
- John<sub>i</sub> overpraised himself<sub>i</sub>
  - \*Himself<sub>i</sub> left
  - \*They<sub>i</sub> said that themselves<sub>i</sub> were clever
  - \*John<sub>i</sub> thought that Mary liked himself<sub>i</sub>

In (3.18a), the governor of *himself* is the verb, and the whole clause is the governing category for the anaphor, therefore, *himself* is bound by *John* and correctly allowed. As for (3.18b), AGR is a governor, and the governing category for *himself* is the whole sentence and this anaphor is free in its governing category. Therefore, principle A will rule out (3.18b). By the same token, (3.18c) and (3.18d) are further cases where an anaphor is not bound in its governing category, here, the embedded clause (Chomsky 1981).

While the principle works for the reflexive in English, it has been noted that for anaphors in many languages, principle A does not hold as stated. This is in particular true for long-distance anaphors. Concerns of this type have been raised in the literature. First, some problems are attributed to the inadequate typology of NPs. For example, Bredekamp (1996) argues that most of the problems are a consequence of the original theory being developed with reference almost exclusively to English, which has a comparatively impoverished anaphoric system (i.e. basically one reflexive, one reciprocal and one pronoun). Compared with the English NP typology, for example, Norwegian has three different anaphors *seg*, *seg selv* and *ham selv* and they each exhibit different binding behaviours (Dalrymple 1993; and see Reinhart and Reuland 1993 for Dutch). On the other hand, Korean *caki* and Japanese *zibun* are categorised as reflexives, but their binding patterns are in part like a reflexive and in part like a pronominal.

Secondly, relevant to the restricted NP typology, a theory with a single definition of governing category cannot describe languages where there is more than one item in each of the categories of *anaphor* like Norwegian and Dutch. Furthermore, many languages like Korean, Japanese and Chinese have been shown to have long-distance reflexives which are syntactically bound in relatively large domains. For these languages, the governing category cannot be the same as stated in principle A.

Thirdly, as the binding domain is defined in principle A and B, pronouns and anaphors are in complementary distribution. That is, a pronoun is disjoint in reference from any antecedent in a certain domain, whilst an anaphor is an element which is bound in the same domain. However, research into languages other than English has shown that this is not a tenable claim. Reinhart & Reuland (1993), for example, shows that in Dutch, the morphologically simplex anaphor *zich* resists local binding, but it is bound rather in a domain similar to that for pronouns.

- (3.19) Jan sag [jou achter zich/hem staan]  
 Jan saw [you behind SE/him stand  
 ‘Jan saw you stand behind SE/him’

In fact, even in English, there are contexts where complementarity breaks down such as *Mary saw a book near herself/her*. The anaphor/pronominal distinction, it seems, is therefore little more than a generalisation which holds for some, but not all, languages; namely that the domain of binding for pronouns is generally less restricted than that for anaphors.

These inadequacies of the theory have motivated numerous proposals for long-distance anaphora. One of the main ones is to try to maintain the locality of the binding domain but to explain the larger domain for long-distance anaphora in terms of movement (Chomsky 1986, Pica 1987, Katada 1990, Yang 1989, Cole *et al.* 1990, Cole & Sung 1994). These works assume that long-distance anaphors undergo LF movement from INFL to INFL. That is, if long-distance anaphors undergo successive cyclic movement just like *wh*-movement (Chomsky 1986), we can account for the fact that they may be bound from beyond the so called local binding domain (Yang 1989). This applies to the LF representation and implies no changes to the binding theory itself as the choice of antecedents is merely determined by the extent to which reconstruction applies (within the head chain headed by anaphor). This is illustrated in the following example:

- (3.20) John<sub>i</sub>-i [Mary<sub>j</sub>-ka caki<sub>i/j</sub>-lul choahanta]-ko sayngkakhanta  
 John-NOM Mary-NOM self-ACC like-COMPL think  
 ‘John<sub>i</sub> thinks that Mary<sub>j</sub> likes self<sub>i/j</sub>’

Two possible LF representations of English translations of (3.20) are shown as follows:

- (3.21) a. [John [<sub>VP</sub> thinks [<sub>CP</sub> that [Mary INFL(*caki*<sub>i</sub>) [<sub>VP</sub> likes *t*<sub>i</sub>]]]]]]]  
 b. [John INFL(*caki*<sub>i</sub>) [<sub>VP</sub> thinks [<sub>CP</sub> that [ Mary INFL*t*<sub>i</sub> [<sub>VP</sub> likes *t*<sub>i</sub>]]]]]]

In (3.21a), *caki* moves to the INFL of the embedded (local) clause where it can be locally bound only by the subject of that clause. In (3.21b), the anaphor moves out of its own clause to the matrix INFL, where it can be locally bound only by the

matrix subject. This accounts for why certain anaphors are long-distance bound and why long-distance bound anaphors are subject oriented. The problem of this approach, as one might have noted already, is that it does not surmount the problems we saw with the subjecthood condition and the c-commanding requirement. That is, the approach cannot account for the cases of non-subject binding as, at each intermediate trace position, the subject, but not the object, will satisfy the binding theory requirements as the latter will be located lower than I. As for cases of backward reflexivization, the movement of the anaphor from I in the subject clause to I in the matrix clause would create a chain where the moved element does not c-command its trace (it does m-command it though) but despite that, unless the antecedent is also moved to a position within the immediate domain of IP (its inner specifier, for example, which is highly unlikely) a proper binding configuration cannot be created as the antecedent still does not c-command the anaphor (see section 3.1.1 and 3.1.2 for the examples).

As such problems show that the theory is inadequate, it has been claimed that long-distance anaphors do not categorially belong to reflexives, but rather pronominals or bound pronominals. The former claim has been made by Bouchard (1984), Kameyama (1984) and Maling (1984). What these works have in common is that despite the identical morphological form, there are two distinct nominal expressions, a long- and a short-distance anaphor, and that the long-distance anaphor is actually a variant of a Pronominal. Even though this claim may sound desirable for theoretical reasons, it has many drawbacks on empirical grounds. In section 2.9, we have seen many differences between a pronoun and *caki* in terms of the binding domains as well as the distribution. For example, while a pronoun can have an antecedent freely outside the sentence, the antecedent of *caki* seems to appear inside the sentence domain. Furthermore, the following example illustrates the fact that *caki* cannot always appear wherever a pronoun does.

- (3.22) a. John<sub>i</sub>-i Bill<sub>j</sub>-ekey [ku<sub>i/j</sub>/caki<sub>i/\*j</sub>-ka sihem-ey hapkyekhayssta]-ko  
 John-NOM Bill-DAT he/self-NOM exam-in passed-COMPL  
 malhayssta  
 told  
 ‘John<sub>i</sub> told Bill<sub>j</sub> that he<sub>i/j</sub>/self<sub>i/\*j</sub> passed the exam’

Further, as pointed out by Huang (1994), given this approach, the vital question of when the long-distance anaphor is predicted to be a true anaphor and when it is predicted to be a pronominal (to be bound outside the local domain) remains unanswered. Unless there is independent evidence to suggest otherwise, to argue that a long-distance reflexive is a pronominal simply because it violates binding condition A is circular.

On the other hand, there is a claim that the long-distance anaphor should be treated as a bound pronominal (Fukui 1984, Ueda 1984). Ueda (1984), for example, looks at a number of problems which arise from viewing *zibun* in Japanese as an anaphor. He suggests that *zibun* is not an anaphor, but rather is an inherently bound pronominal, a pronominal with only a bound interpretation. He provides a case of topic-binding as a piece of evidence and the same case is available in Korean as follows:

- (3.23) John<sub>i</sub>-un caki<sub>i</sub>uy ayin-i Seoul-ey kapulyessta  
 John-TOP self-GEN lover-NOM Seoul-to has gone  
 ‘As for John<sub>i</sub>, (his) lover<sub>i</sub> has gone to Seoul’

Although the approach seems to work in Korean, it has a serious inherent problem. Given that a bound Pronominal is a Pronominal, we would expect that it falls under binding principle B. Thus the approach completely ignores the fact that *caki* may also be locally bound.

Within the Minimalist Program (Chomsky 1993, Chomsky 1995), although the status of binding theory remains quite unclear, Chomsky proposes to account for principle A effects including long-distance anaphora along the lines outlined above. Specifically he dubs the LF head operation CLITICIZATION<sub>LF</sub>. In general, however,

we must point out that binding theory in its totality is not a popular domain for researchers working within the Minimalist Program. The reason for this is that minimalist syntax is a strongly derivational theory and binding theory is a module of the theory of grammar whose principles are stated over representations. To my knowledge, there has not been a serious attempt to develop a derivational binding theory. In fact, Kural & Tsoulas (1999) suggest that it may be impossible or at least highly undesirable to construct a derivational theory for principle B effects.

In this section, we have looked at some syntactic approaches to long-distance anaphora. Even if the syntactic approaches have gone through an extensive modification, it seems that they face a similar and constant set of problems. Most approaches discussed so far do not offer any way to treat the case of non-subject binding let alone the case of backward reflexivization. The attempt to redefine long-distance anaphora (i.e. to be bound outside the local domain) as other NP categories fails to provide a unified account for *caki*-binding. As this suggests that syntax may not be sufficient in specifying the domain or the set of possible antecedents for long-distance reflexives, many researchers believe that a pure syntactic account alone cannot truly capture the nature of long-distance anaphora. This has led to a number of non-syntactic treatments for long-distance anaphora, which we will discuss in the next section.

## 3.2 Non-Syntactic Approaches

In this section, we will examine non-syntactic approaches; the discussion will be divided into two subsections. The first subsection will look at some discourse approaches and the second section will discuss semantic approaches that take advantage of thematic roles. In the former, as an example of discourse approaches, we will look at “logophoricity” whose effects have been investigated in many other languages. In the latter, we will discuss the approaches that exploit the semantic thematic hierarchy to account for the phenomena.



### 3.2.1 Discourse Approaches

Logophoricity refers to the phenomenon whereby the ‘point of view’ of an internal protagonist of a discourse, as opposed to that of the current, external speaker, is reported.<sup>6</sup> It describes the behaviour of certain lexical items normally considered anaphoric, that is, bound in some syntactic domain, or which are bound under certain circumstances outside that domain. The phenomenon was first identified in certain West African Languages where there is a class of so called ‘logophoric pronouns’ which is used to refer to an antecedent ‘whose speech, thoughts, feelings, or general state of consciousness are reported’ (cf. Clements (1986)). Crosslinguistically, however, logophoricity is also expressed in the mechanisms of long-distance reflexives. In Icelandic, for example, Maling (1984) argues that the antecedent of a non-clause bound reflexive must be both ‘logophoric’ (=notion of ‘source’) and a grammatical subject at the same time. Consider the following:

- (3.24) a. Hann<sub>i</sub> sagdu [ad sig<sub>i</sub> vabtadi haefileika]  
he said [that self lacked ability]  
‘He<sub>i</sub> said that he<sub>i</sub> lacked ability’
- b. \*Honum<sub>i</sub> var sagt [ad sig<sub>i</sub> vabtadi haefileika]  
he was told [that self lacked ability]  
‘He<sub>i</sub> is told that he<sub>i</sub> lacked ability’

In (3.24a), *Hann(He)* can be a legitimate antecedent by being logophoric (as a source of the information) as well as a grammatical subject, thus correctly binding *sig*. However, in the passive counterpart (3.24b), even though *Honum(he)* is a grammatical subject, because it is not a source, it cannot be an antecedent of *sig*, which makes the indicated binding relation ungrammatical. Kuno(1987) also adopts the notion of logophoricity to account for *zibun-binding*. For him, the logophoric use of *zibun* is when *zibun* is used in the subordinate clause of a verb which denotes speech,

---

<sup>6</sup>There are indeed various approaches grouped under the discourse approaches. In this section, we will look at a few of them. For a wide range of approaches, see Kuno (1972) and Kuno (1973) for direct-discourse analysis, Kuno (1976) for empathy and Banfield (1982) and Zribi-Hertz (1989) for point of view.



thought, or consciousness, in which cases *zibun* refers to the speaker or experiencer of that speech, thought or consciousness. Consider the following example:

- (3.25) Taroo<sub>i</sub> wa zibun<sub>i</sub> ga boku yori benkyoo ga dekinai koto o  
 Taroo TOP self NOM me than study ACC cannot COMP ACC  
 zuibun kinisiteiru nasii  
 very-much bother seem  
 ‘Taroo<sub>i</sub> seems to be bothered by the fact that he<sub>i</sub> is not academically  
 better than me.’

(Kuno 1978)

In (3.25), *zibun* is bound by *Taroo* because the complement which contains *zibun* in these sentence is interpreted as a logophoric domain, presenting *Taroo*’s mental state.

Among approaches relevant to Logophoricity, the definition of logophoric anaphora differs and is rather loosely based on concepts such as consciousness, point of view, mental states etc. However Sells (1987) defines the concept of logophoric anaphor in terms of three primitive discourse-semantics notions: *Source*, *Self* and *Pivot*. He also provides four discourse environments defined by the discourse-semantics roles: *Direct Speech*, *Third-person point of view*, *Psych-verb* and *Logophoric-verb*. These two contexts (constituted by the discourse-semantics roles and discourse environments) decide which role is assigned to a logophoric anaphor between the external speaker and the internal protagonist. To exemplify this, consider the following example with *caki*:

- (3.26) a. John<sub>i</sub>-un Mary-ka macimakulo caki<sub>i</sub>-lul po-la-wassul-ttay aphassta  
 John-TOP Mary-NOM last self-ACC see-to-came-when was sick  
 ‘John<sub>i</sub> was sick when Mary came to see him<sub>i</sub> last’

- b. ??John<sub>i</sub>-un Mary-ka macimakulo caki<sub>i</sub>-lul po-la-kassul-ttay aphassta  
 John-TOP Mary-NOM last self-ACC see-to-went-when was sick  
 ‘??John<sub>i</sub> was sick when Mary<sub>j</sub> went to see him<sub>i</sub> last’

(Kuno 1978)

(3.26) represents a logophoric setting for what Sells calls ‘third-person point of view’, where only the pivot role is predicated of the internal protagonist. The use of *wassul* (come) in the (3.26a) sentence makes clear that what is reported is from the space location of *John*, therefore *John* is the pivot, and so *caki*-binding is licensed. On the other hand, the use of *kassul* (went) in (3.26b) indicates movement away from rather than towards *Mary*, therefore *John* is not the pivot, and so *caki*-binding is only very marginally acceptable.

When logophoric approaches are applied to *caki*-binding such as in (3.26), the notion seems to predict the correct binding relation. Against this background, consider the following:

- (3.27) Mary<sub>i</sub>-ka aki<sub>j</sub>-lul caki<sub>i/j</sub> chimday-e nwuphiessta  
 Mary-NOM baby-ACC self bed-in put  
 ‘Mary<sub>i</sub> put (her) baby<sub>j</sub> in self<sub>i/j</sub>’s bed’

The context of ‘putting a baby in bed’ is neutral in the sense that it does not involve any deictic notion. Moreover, there is no possible interpretation of ‘consciousness’, ‘mental-state’, ‘source’ ‘pivot’ etc. In such a case, the logophoric notion cannot apply and the anaphors in (3.27) can have nonlogophoric antecedents without any problem. This shows that the approach based on the notion of logophoricity cannot on its own explain the data. It needs to be supplemented by a different theory for cases of non-logophoric contexts.

Most logophoric binding approaches also do not require that the antecedent is structurally represented in the sentence, so the anaphoric element is allowed to be free

(cf. (Thráinsson 1991)). However, this is not necessarily the case since there are examples where the anaphor cannot be left unbound. Consider the following examples in Korean.

- (3.28) A. Mary<sub>i</sub>-ka ku pati-e kass-ni anim tarun salam-i taysin kass-ni?  
 Mary-NOM the party-to went-Q or other person-NOM instead go-Q  
 ‘Is it Mary<sub>i</sub> who went to the party or somebody else instead?’
- B. Ani, \*caki<sub>i</sub>-nun kasse  
 No, self<sub>i</sub>-TOP went  
 ‘No, \*self<sub>i</sub> went’

The ungrammaticality of *caki*-binding in (3.28) is not improved even if some appropriate context is provided. This indicates that *caki*-binding cannot be free from any syntactic condition. We will return to discourse approaches in section 3.4.

### 3.2.2 Thematic roles

In parallel with the discourse approaches above, some researchers provide evidence that certain conditions on anaphors have to be stated in terms of the *thematic roles* of binder and bindee and argue that the theory of reflexivization in terms of semantic notions would be more parsimonious, and thus more highly valued, than one that relies on the syntactic order or hierarchical arrangement of constituents (Jackendoff 1972, Hellan 1988, Wilkins 1988, Kiss 1991, Momoi 1985). Much of the work stems from the the notion of thematic roles defined by Jackendoff (1972), who describes a number of syntactic phenomena that require explanation with reference to semantic structure. He also emphasises that the syntactic impact of thematic roles can be further seen in the more widespread effect of the *Thematic Hierarchy Condition*, which reflexivization is also sensitive to. The motivation for such a hierarchy effect on reflexivization comes from the following data.

- (3.29) a. I talked to John about himself  
 b. ??I talked about John to himself

For both sentences in (3.29), syntactically, *to John* in (3.29a) is a complement PP, and so is *about John* in (3.29b). Despite this, the different binding patterns in (3.29) prevent syntax from offering any explanation why *John* can be the antecedent for the reflexive in (3.29a) but not in (3.29b). Let us also consider the following examples:

- (3.30) a. I talked to myself about myself  
b. I talked about myself to myself

Why is the inversion of the two prepositional phrases allowed, with reflexivization maintained, when the semantic content of the arguments changes to *speaker (myself)*? For the explanation of examples like (3.29) and (3.30), Jackendoff (1972) proposes a hierarchy based on the content of thematic roles and a thematic hierarchy condition as follows:

- (3.31) Agent > Location, Source, Goal > Theme

- (3.32) Condition: A reflexive may not be higher on the thematic hierarchy than its antecedent

With respect to the condition above, (3.29b) is bad, because the reflexive (Goal) is higher on the hierarchy than the antecedent (Theme). (3.29a) is grammatical because the reflexive is lower than the antecedent in the hierarchy. As for (3.30), the prepositional phrase alternation with reflexives is allowed because the antecedent is the agent, *I*, not the theme or the goal. The agent is higher on the thematic hierarchy than the other two arguments, theme and goal. Thus the ranks of these other two are irrelevant to the choice of antecedent for the reflexives. However, as pointed by Wilkins (1988), the thematic hierarchy condition seems to make some wrong predictions:

- (3.33) a. With that new kind of kryptonite lock, you have to lock the bike to  
*itself*  
b. The therapist introduced the new Mary to *herself*

(Wilkins 1988, p.206, ex.(32a,b))

The antecedents *the bike* and *Mary* in (3.33) are Patients. A Patient is a Theme that is directly affected and there is an agent or instrument. At the same time, the reflexives are goals. In order to license the binding, the patient role should be higher than the goal in the hierarchy, which means that the hierarchy as stated needs revising to incorporate the role of patient.

The notion of thematic hierarchy has been widely used to explain long-distance anaphora, often invoking a parametrisation of the hierarchy. For example, in treating *zibun*-binding, Momoi (1985) claims that the semantic representation of NPs can provide a uniform account for subject and nonsubject *zibun*-binding. A revised thematic hierarchy to predict the correct antecedents of *zibun* is given below:

(3.34) Experiencer > Agent > Objective (Patient) > Others (Goal, Source, etc.)

Momoi (1985) suggests that the experiencer and agent roles are equal in strength to be an antecedent, and these are followed by the objective.<sup>7</sup> This is illustrated in the following examples cited by Iida (1996).

- (3.35) a. John<sub>i</sub> wa Mary<sub>j</sub> ni zibun<sub>i/?j</sub> no hon o miseta  
 John TOP Mary DAT self GEN book ACC showed  
 ‘John<sub>i</sub> showed Mary<sub>j</sub> his/?her book’
- b. John<sub>i</sub> wa Steve<sub>j</sub> ni [zibun<sub>i/j</sub> no tomodati ga ikkai ni kiteiru]  
 John TOP Steve DAT self GEN friend NOM 1st floor to came  
 koto o osieta  
 COMP ACC told  
 ‘John<sub>i</sub> told Steve<sub>j</sub> that his<sub>i/j</sub> friend had come to the 1st floor’

According to Momoi (1985), his thematic hierarchy is advantageous in particular for the different acceptability of the dative antecedent of *zibun* (the case of non-subject binding) in (3.35). In (3.35), *John* is the agent, and thus binds *zibun*. The different acceptability of the dative *Mary* and *Steve* being the antecedent in both sentences, as Momoi argues, depends on whether it gets interpreted as goal or experiencer: If

---

<sup>7</sup>Momoi refers to the objective as meaning the patient role.

*Mary* is the goal, it cannot be the antecedent of *zibun* because it is at the bottom of the hierarchy. On the other hand, if it is interpreted as the experiencer, it can antecede *zibun* since it outranks the rest of the thematic roles.

Even though this approach seems attractive in that it accounts for the flexible non-subject binding, it has at least two inherent problems which are noted by Iida (1996). In the light of Korean *caki*-binding, I also agree with her arguments against the approach.

Firstly, in this approach, it is not yet very clearly defined what the thematic role is assigned to. If it should be assigned to an NP in an argument position, then a possible antecedent in a non-argument position should not be allowed. However, this is not borne out:

- (3.36) John<sub>i</sub>-uy silswu-nun [caki<sub>j</sub>-ka nemwu cengcikhata]-nunkes ita  
 John-GEN mistake-TOP self-NOM too is honest-COMPL is  
 ‘John<sub>i</sub>’s mistake is that self<sub>j</sub> is too honest’

Even though the antecedent *John* is not in an argument position, the binding is still licensed. Given that the antecedent receives no thematic role from the main predicate, Momoi’s (1985) theory is not applicable to these data. We will return to the significance of this type of example in more detail in chapter 5.

Secondly, Iida (1996) also claims that there are cases where a legitimate antecedent cannot be defined in terms of a thematic role hierarchy. For example, in cases of backward reflexivization as below, it is not clear what kind of role the antecedent NP *George* bears:

- (3.37) [Caki<sub>i</sub>-ka palmyengha]-n kikey-ka George<sub>i</sub>-uy kun pwu-lul  
 self-NOM invented-REL machine-NOM George-GEN huge wealth-ACC  
 kaceywassta  
 brought  
 ‘The machine that self<sub>i</sub> invented brought about George<sub>i</sub>’s big fortune’

Here the verb cannot assign the role Experiencer or Agent to the potential antecedent *George*. Therefore, it is not clear how the thematic roles will be available on the NPs in a sentence like this.

Empirically, also, the hierarchy does not seem to provide an adequate basis for predicting the correct binding patterns as illustrated below:

- (3.38) John<sub>i</sub>-i Bill<sub>j</sub>-lopwute [caki<sub>i/j</sub>-ka sihem-e hapkeyhayssta]-ko tulessta  
John-NOM Bill-from self-NOM exam-in passed-COMPL heard  
'John<sub>i</sub> heard from Bill<sub>j</sub> that self<sub>i/j</sub> passed the exam'

In (3.38), non-subject binding is possible with *Bill* binding *caki* and there is no difference in acceptability as with the dative NP in Japanese as in (3.35b). *John* is the source in this case. Then, given the hierarchy in (3.34), there is no way to explain why *caki* is bound by the source, *Bill*, which is at the bottom of the hierarchy.

Having looked at the approaches that make use of a thematic hierarchy, it is worth considering whether a thematic hierarchy can be parametrised. According to Culicover and Wilkins (1984), thematic roles have a dual status. One is an integral part of the definition of particular acts and states and as such are components of the *mental* representation of objects and concepts. This is independent of the theory of grammar, as is the cognitive capacity for individuating acts or states. The other role is a component of a level of grammatical interpretation. Even if a thematic role is realised in a different grammatical form in different languages, the defining function of roles remains constant. For example, the agent in one language should have the same role in other languages, though a subject in one language may be realised in another grammatical relation in other languages. Given this, it is difficult to assume that the thematic hierarchy with a pure notion of thematic roles can vary cross-linguistically unless grammatical or lexical reasons are incorporated.

In this section, we have looked at a few approaches that make use of the pure notion of thematic roles and the thematic hierarchy which are relevant for the explanation



of long-distance anaphora. We have seen that the approaches with a pure notion of thematic roles turn out to be useful in explaining some of the non-subject binding cases. However, there were cases where the given hierarchy cannot predict the correct binding relation. In order to account for the problems, parameterising the hierarchy is linguistically not very desirable. Furthermore, it is not clear what the predictions of the thematic hierarchy based theory are for cases where a potential antecedent is not a coargument of *caki*.

So far, in the last two sections, we have seen the syntactic approaches and non-syntactic approaches that are relevant to the explanation of long-distance anaphora. It seems that neither of the approaches is satisfactory. In the meantime, however, there are a number of analyses that claim that the phenomenon may be best explained by a combination of the syntactic and non-syntactic factors. This kind of analysis is divided methodologically into two branches: disjunctive and conjunctive theories, depending on how the two kinds of factors are combined to account for long-distance anaphora, which we will deal with in the subsequent sections. The two theories we will have a look at are mostly proposed to account for long-distance anaphora in Japanese. In the two following sections, we will briefly look at what mechanisms they offer for the explanation for the phenomenon in Japanese and examine whether *caki*-binding can be suitably explained by those accounts.

### 3.3 Disjunctive Theory of Long-Distance Anaphora

In the disjunctive theory, any instance of long-distance anaphora has to satisfy *either* the subjecthood condition *or* the discourse condition. Apart from the subjecthood condition, different people adopt different discourse approaches. For example, Gunji (1987) explains *zibun-binding* in terms of the Foot Feature Principle in the framework of Generalised Phrase Structure Grammar (as in Gazdar *et al.* (1985)). We will not go into the technical details. Basically, however, he distinguishes two cases of *zibun*-binding: subject-binding and non-subject binding. For him, *zibun* which



is left free (i.e., when the SUBJ value is uninstantiated), is free to be bound by any pragmatic or discourse controller, which is, typically, the speaker who is expressing his/her view by uttering the sentence.

- (3.39) Naomi<sub>i</sub>-ga Ken<sub>j</sub>-ni **zibun**<sub>i/k(sp*k*)</sub>-no kuruma-o kasita  
 Naomi-NOM Ken-DAT self-GEN car-ACC lent  
 ‘Naomi<sub>i</sub> lent Kent<sub>j</sub> **self**<sub>i/k(sp*k*)</sub>’s car

(Gunji 1987)

However, it is not explicitly stated how we can predict the correct binding relations among discourse controllers (i.e., non-subjects). If we recall the case of non-subject binding discussed by Momoi (1985), it was shown that the dative NP may serve as antecedent given a context as follows:

- (3.40) John<sub>i</sub> wa Steve<sub>j</sub> ni [zibun<sub>i/j</sub> no tomodati ga ikkai ni kiteiru]  
 John TOP Steve DAT self GEN friend NOM 1st floor to came  
 koto o osieta  
 COMP ACC told  
 ‘John<sub>i</sub> told Steve<sub>j</sub> that his<sub>i/j</sub> friend had come to the 1st floor’

(As (3.35b) in section 3.2.2)

In this case, it is not clear why in (3.39) the dative *Ken* is excluded from being an antecedent and why in (3.40) the dative *Steve* can serve as an antecedent. In Korean, however, examples like the above would not have any problem as the only legitimate antecedent is the subject. However, the problem still persists in the sentence below:

- (3.41) Mary<sub>i</sub>-ka John<sub>j</sub>-uopwute [**caki**<sub>i/j</sub>-ka Bill<sub>k</sub>-lul hyemohanta]-ko tulessta  
 Mary-NOM John-from self-NOM Bill-ACC hate-COMPL heard  
 ‘Mary<sub>i</sub> heard from John<sub>j</sub> **self**<sub>i/j</sub> hates Bill<sub>k</sub>’

When the adjunct *John* binds *caki*, it is not clear what determines such a binding relation unless discourse factors are provided to allow the adjunct *John* as an antecedent, but disallow the dative NP.

Furthermore, we saw that the discourse controller such as the speaker can play a role in Japanese. In Korean however, only the subject binding is allowed but not the speaker, thus the discourse condition is not a sufficient condition as illustrated below:

- (3.42) John<sub>i</sub>-i Mary<sub>j</sub>-ekey **caki**<sub>i/\*spk</sub>-uy cha-lul pilye-chwuessta  
 John-NOM Mary-DAT self-GEN car-ACC lent-gave  
 ‘John<sub>i</sub> lent Mary<sub>j</sub> **self**<sub>i/\*spk</sub>’s car

Kameyama (1984) also provides an approach within the disjunctive theory. Apart from the subjecthood condition as a syntactic factor, she employs logophoricity as a determining factor for non-subject binding. We have already discussed Logophoricity in section 3.2.1, thus we will not repeat the details here. One of the general problems of the disjunctive theory is that, as Iida (1996) argues, in Japanese, long-distance binding can be found in the cases of multiple occurrence of *zibun*. As we have already seen in section 2.6, multiple occurrences of *caki* are also allowed. If there are two independent mechanisms, namely syntactic binding and discourse binding, and they operate disjunctively, nothing blocks the two *zibuns* from taking different antecedents, licensed by the different binding mechanisms. Consider the Korean example in (3.43) below:

- (3.43) John-i Mary-lul caki-uy cip-ulo caki-uy cha-lo tolye ponayssta  
 John-NOM Mary-ACC self-GEN house-to self-GEN car-by return sent  
 ‘John sent Mary to his house by his car’  
 ‘John sent Mary to her home by her car’  
 ‘John sent Mary to his house by her car’  
 ‘John sent Mary to her house by his car’

In this example, it is possible that the subjecthood condition licenses the binding of *John* to *caki* in one of the possessive (genitive) arguments, while the discourse

condition, say a logophoric condition, licenses the binding of *Mary* to *caki* in the other possessive argument. Two instances of *caki* in the same clause, however, must have the same antecedent. That is, only the first two readings are possible.

The crucial reason why the discourse condition fails to apply to *caki*-binding is that Korean does not seem to allow any discourse antecedents, which means that Korean does not follow the discourse condition at all. Let us have a look at Japanese examples which show that in Japanese, *zibun* may be bound by a discourse antecedent.

- (3.44) a. **Zibun**<sub>*j*(*spk*)</sub>-ga Mikisan<sub>*i*</sub>-ni teepu-wo agemashita  
 self-NOM Miki-DAT tape-ACC gave  
 ‘self<sub>*j*(*spk*)</sub> gave Miki<sub>*i*</sub> a tape’
- b. **Zibun**<sub>*j*(*lst*)</sub> Mikisan<sub>*i*</sub>-ni teepu-wo agatandeshoo  
 self-NOM Miki-DAT tape-ACC gave  
 ‘Self<sub>*j*(*lst*)</sub> gave Miki<sub>*i*</sub> a tape (didn’t you?)’

In (3.44a), *zibun* is not bound by any syntactic argument antecedent but by the speaker in the discourse. Similarly in (3.44b), *zibun* is bound by the listener in the discourse. Thus both binding relations satisfy the discourse conditions and satisfy the disjunctive theory. However, the equivalent sentence in Korean in (3.45) shows that *caki* cannot be bound by a discourse antecedent or by any syntactic argument when it occupies the matrix subject positions. Since it was noted that *caki* is specified as third person, the binding of *caki* by a speaker or a listener may be straightforwardly ruled out without any reference to discourse binding being necessary. However, the binding possibility by a discourse antecedent *k* is ruled out, which shows that *caki*-binding does not allow discourse binding. This was also shown clearly in the example earlier in (3.28) even with a given context.

- (3.45) \***Caki**<sub>*i*(*spk*)/*j*(*lst*)/*k*</sub>-ka John<sub>*k*</sub>-ekey teipu-lul cwuessta  
 self-NOM John-DAT tape-ACC gave  
 ‘Self<sub>*i*(*spk*)/*j*(*lst*)/*k*</sub> gave John<sub>*k*</sub> a tape’

One may argue that examples such as (3.45) just show that such cases do not obey the disjunctive theory by not satisfying either the subjecthood condition or the

discourse condition. However, examples such as (3.46) below can be provided to oppose such an argument.

- (3.46) Mary<sub>i</sub>-nun caki<sub>i</sub>-ka John-ekey teipu-lul cwuessta  
Mary-TOP self-NOM John-DAT tape-ACC gave  
'As for Mary, self<sub>i</sub> gave John a tape'

In (3.46), the topic *Mary* binds *caki*. According to the disjunctive theory, it violates the subjecthood condition, thus has to satisfy the discourse condition. However, it is not clear on what basis of the discourse condition *Mary* can be decided as a legitimate antecedent. It looks like it does not satisfy any discourse condition, for example, logophoricity, in the context as in (3.46). However, the sentence is grammatical.

### 3.4 Conjunctive Theory of Long-Distance Anaphora

Unlike the disjunctive theory, in the conjunctive theory, any instance of long-distance anaphora should satisfy both a syntactic condition and a discourse condition as proposed by Maling (1984), Iida (1996) and Sells (1987).

For Icelandic, we already have seen in section 3.2.1 (ex. 3.24) that the antecedent of a non-clause bounded reflexive must be both logophoric and a grammatical subject at the same time. As for Japanese (Iida 1996), both a syntactic condition and a discourse condition have to be satisfied for *zibun*-binding. However, the respective roles of the syntactic and discourse conditions differ from those in the disjunctive theory. Iida (1996) proposes her conjunctive theory in the framework of Head-driven Phrase Structure Grammar and defines the syntactic condition (see section 4.2) as follows:

- **Syntactic Constraint:** *zibun* may not o-command its antecedent.

(Iida 1996, p.117)

This syntactic constraint is minimally involved in *zibun*-binding and states that *zibun* cannot be bound by an antecedent that is realized in a lower coargument position than itself, and this applies to any instance of *zibun*-binding. The following example illustrates the syntactic constraint.

- (3.47) \***Zibun**<sub>i</sub>-ga Taroo<sub>i</sub>-o kurusimeteiru  
 self-NOM Taroo-NOM annoy  
 ‘Self<sub>i</sub> annoys Taroo<sub>i</sub>’

(Iida 1996, p.117, ex.71)

In (3.47), *zibun* is bound by *Taroo* which is its lower coargument, which makes the sentence ungrammatical. Conjunctively with this syntactic constraint, any instance of *zibun*-binding should obey the following discourse constraint.

- **Discourse Constraint:** an antecedent of *zibun* is the perspective chosen by the speaker in describing the situation in question.

The subject is normally used to represent the speaker’s perspective and thus subject binding is a special case of ‘discourse binding’. Thus, subject binding is attributed to discourse factors, in particular, *deictic perspective*: *zibun* is perspective sensitive, and it takes a *sentence internal* antecedent which has perspective. To see how this discourse constraint can be interpreted, let us have a look at the following English examples:

- (3.48) a. John will introduce Mary to the man sitting on the right of the table  
 b. John will introduce me to the man on the right of the table  
 c. (Hanako is very angry) Taroo seemed to tease Hanako in front of her friends.

(3.48a) shows the case where the deictic expression *on the right of the table* may be determined from the subject’s (*John*) perspective. When the object is replaced by the first person pronoun as in (3.48b), then it is interpreted as the speaker taking his/her perspective of the deictic expression. Furthermore, in (3.48c), providing

background information to set the speaker's perspective on Hanako, the speaker interprets the situation from his/her own perspective. Given this notion of *perspective*, let us consider example (3.47) above which is grammatical and in which *zibun* is not bound by *Taroo*, but by the speaker:

- (3.49) **Zibun**<sub>spk</sub>-ga Taroo-o kurusimeteiru  
 self-NOM Taroo-ACC annoy  
 'Self<sub>spk</sub> annoys Taroo'

(3.49) is indicated as grammatical, which means that it must satisfy both the syntactic constraint and the discourse constraint illustrated so far. Firstly, (3.49) obeys the syntactic constraint in that *zibun* is not bound by any lower coargument. Moreover, *zibun* takes the speaker's perspective, that is, it is bound by the speaker, thus satisfies the discourse constraint. Thus, *zibun*-binding is licensed.

Having looked at the conjunctive theory more closely, we now present some points where the theory fails to determine a correct *zibun*-binding and further observe that Korean *caki*-binding cannot be described by the conjunctive theory provided.

First, the discourse constraints in terms of perspective do not seem to provide a completely accurate binding relation for *zibun*. There are instances where both subject and non-subject can be equally an antecedent for *zibun* in terms of deictic perspective, but only one of them can legitimately antecede *zibun*. Consider the following example:

- (3.50) David-wa George no tameni **zibun**-no haizara-wo teiburu-no migi-ni  
 David-TOP George for self-GEN ashtray-ACC table-GEN right-in  
 oita  
 put  
 David<sub>i</sub> put self<sub>i</sub>'s ashtray on the right of the table for George.

In (3.50), both *David* and *George* can be viewed as having a perspective. For example, let us imagine the context where everybody knows that *George* is right handed or is singlehanded and has only a right hand. Then from the interpretation

in that context, the position of the ashtray may well be either *David's* right (if *David* is mean) or *George's* right. If this contextual assumption is right, *zibun* should be able to be bound by either *David* or *George* in terms of perspective. However, the indicated binding does not show this. The subject *David* can antecede *zibun* while *George* cannot. Let us now carefully consider each case when *David* and *George* take perspective respectively.

First, when *David* has perspective, it is less oblique than *zibun*, thus both syntactic and discourse conditions meet and the binding is licensed. While *David* takes perspective, *George* does not, so it does not satisfy the discourse condition. Also *George* is more oblique than *zibun*, thus violating the syntactic condition. Therefore, this case violates both conditions at the same time and so *George* cannot be the antecedent.

The next case is when *George* takes perspective; because it is more oblique than *zibun* it violates the syntactic condition. Thus, it does not satisfy both conditions at the same time and it cannot antecede *zibun*. In the meantime, while *David* does not take perspective, it satisfies the syntactic condition by being less oblique than *zibun*. However, since both conditions are not met at the same time, the binding is not supposed to be allowed. Despite this fact, *David* can still antecede *zibun*. This is where the conjunctive theory breaks down. Thus, the possible candidates for antecedency in terms of the perspective assumption do not accord with the notion of antecedency in the conjunctive theory.

Secondly, Iida (1996) argues that the antecedent of *zibun* is a sentence internal NP which HAS perspective (Iida 1996, p.181). Such constraints on sentence internal antecedents may cover some of the data, but cannot explain cases where the antecedent of *zibun* does not exist sentence internally but in some discourse level as illustrated below:

- (3.51) a. **Zibun**<sub>*j*(*spk*)</sub>-ga Mikisan<sub>*i*</sub>-ni teepu-wo agemashita  
 self-NOM Miki-DAT tape-ACC gave  
 ‘self<sub>*j*(*spk*)</sub> gave Miki<sub>*i*</sub> a tape’



- b. **Zibun**<sub>*j*(*lst*)</sub>-ga Mikisan<sub>*i*</sub>-ni teepu-wo agatandeshoo  
 self-NOM Miki-DAT tape-ACC gave  
 ‘Self<sub>*j*(*lst*)</sub> gave Miki<sub>*i*</sub> a tape (,didn’t you?)’

Further, even if *zibun* was indeed bound sentence internally, the account of its binding relations can be reduced just to a structural relationship, rather than employing the discourse notion of perspective.

Thirdly, the discourse constraint that holds for *zibun*-binding is not applicable to Korean long-distance binding in that *caki* is not bound in the discourse domain. There are no instances where *caki* is not sentence-internally bound. The following would be an example of this type.

- (3.52) Na<sub>*i*</sub>-nun Bill<sub>*j*</sub>-ekey [caki<sub>*\*i*/*\*j*/*\*k*/*\*l*</sub>-ka Mary<sub>*k*</sub>-lul mannassta]-ko malhayssta  
 I-TOP Bill-DAT self-NOM Mary-ACC met-COMP said  
 I<sub>*i*</sub> said to Bill<sub>*j*</sub> that self<sub>*\*i*/*\*j*/*\*k*/*\*l*</sub> met Mary<sub>*k*</sub>’

*Caki* embedded in the complement clause is not bound by either *I* (because it does not agree with *caki* in person), *Bill* (as we have observed) or *Mary* (because it is the lower argument than *caki*). Further it is not bound by an antecedent outside the sentential domain, that is, a discourse antecedent indicated by the indexing ‘*l*’. Given that binding by the discourse antecedent is not an option, a discourse condition is redundant.

Another reason that the conjunctive theory does not apply to *caki*-binding comes from the notion of perspective as a discourse constraint. One of the crucial pieces of evidence Iida (1996) suggests for the role of perspective in *zibun*-binding is based on the following examples (which we also saw in chapter 2):

- (3.53) a. \*Taroo<sub>*i*</sub>-wa [Hanako-ga **zibun**<sub>*i*</sub>-o hihansita toki] damatteita  
 Taroo-TOP Hanako-NOM self-ACC criticised when kept silence  
 ‘Taroo<sub>*i*</sub> said nothing when Hanako criticised self<sub>*i*</sub>’



- b. Taroo<sub>i</sub>-wa [Hanako-ga **zibun**<sub>i</sub>-o hihansita noni damatteita  
 Taroo-TOP Hanako-NOM self-ACC criticised though kept-silence  
 ‘Taroo<sub>i</sub> said nothing though Hanako criticised **self**<sub>i</sub>’

*Zibun* in the temporal clause as in (3.53a) cannot be bound to the matrix subject *Taroo*, while *zibun* in the concessive clause as in (3.53b) can. The use of the conjunction *noni* (‘though’) in (3.53b) indicates that the speaker is making a report based on her own knowledge, evaluation or reasoning about Taroo: the speaker’s interest in *Taroo* may be understood as the speaker’s identification with *Taroo*. However, the speaker’s perspective on the matrix subject does not seem to say much about *caki*-binding. In instances with both temporal and concessive clauses, *caki* is bound by the matrix subject without any difference in the binding relation between the two different kinds of conjunctions, unlike *zibun*-binding relations as in (3.53).

- (3.54) a. Mary<sub>i</sub>-nun [John-i **caki**<sub>i</sub>-lul pipanhal-ttay] chimmwuk-ul cikyessta  
 Mary-TOP John-NOM self-ACC criticized-when silence-ACC kept  
 ‘Mary<sub>i</sub> said nothing when Jhon criticized **self**<sub>i</sub>’
- b. Mary<sub>i</sub>-un [John-i **caki**<sub>i</sub>-lul pipanha-nunteto chimmwuk-ul cikyessta  
 Mary-TOP John-NOM self-ACC criticized-though silence-ACC kept  
 ‘Mary<sub>i</sub> said nothing though John criticized **self**<sub>i</sub>’

The consistent binding patterns in (3.54) show that the different perspective captured in the different adjunct clauses as in *zibun*-binding does not exist in Korean.

So far, we have proposed three points to suggest that the conjunctive theory is not applicable to *caki*-binding. First, we have seen that the conjunctive mechanism of the syntactic condition and the discourse condition do not predict the correct binding relations for *zibun* as illustrated in (3.50). Secondly, the theory proposes the constraint that *zibun* should be bound by a sentence internal antecedent, but there are some clear instances where the antecedents are not found within the sentence (3.51). Thirdly, we have seen *zibun*-binding in the context of different adjunct clauses to demonstrate the effect of the discourse factor, perspective. However,

*caki*-binding does not seem to be sensitive to such discourse factors. Therefore, even though the conjunctive theory explains a wider range of *zibun*-binding both for subject and non-subject binding, the crucial reason that it may not be applicable to *caki* is that the discourse factors seem redundant for *caki*-binding. We have seen many cases where *caki* cannot take a discourse antecedent: either speaker or listener or any referent outside the sentence. Given this, an account of *caki* should be able to do without any discourse factors. and may be reduced to a structural one.

### 3.5 Summary and Conclusion

So far, we have looked at two main groups of approaches to long-distance anaphora: the syntactic approaches and non-syntactic approaches. As for the former, we started with the basic proposal, the subject condition, which was followed by the c-commanding requirement and the approaches within GB. The common problems in these approaches comprise instances of non-subject binding and backward reflexivization. Among the non-syntactic approaches, we have discussed discourse accounts and semantic accounts. Logophoricity was examined as an example of a discourse account. While it seems to handle some mysteries of non-subject binding, we found that this account requires a very specific context, such as verbs of consciousness, feeling and thoughts, etc. In the examples where there is no such context involved, it is not clear how the correct binder should be determined. On the other hand, the thematic hierarchy approaches seem to have resolved some difficult areas for accounts that rely on the syntactic order or hierarchical arrangement of constituents. However, leaving aside the fact that the way the hierarchy is applied remains unclear, proposing a parametrically variant thematic hierarchy does not seem very desirable because of the cognitive status of thematic roles.

The overview of syntactic vs. non-syntactic approaches shows that the pure syntactic or the pure discourse approaches alone are not sufficient to account for long-distance anaphora. Additionally, where one approach accounts for data that the other fails to, then conversely, it seems that the other fails to explain what the first

could cover. It is this complementarity in the contribution of syntactic and non-syntactic approaches that motivates the disjunctive and conjunctive theories. As for the disjunctive theory, when the syntactic condition (the subjecthood condition) is not met, then the discourse condition is solely responsible. This can be problematic for *caki*-binding. Leaving all other instances aside, the discourse factor proves to be difficult for *caki* where *caki* cannot be bound by the speaker or the listener or even any referent in the discourse domain. Furthermore, we have seen that the disjunctive mechanisms of the theory cannot account for multiple occurrences of *caki*. The conjunctive theory we have seen contains a more formalised form of the syntactic condition (o-binding) and discourse condition (deictic perspective). This has proved to be more adequate in accounting for *zibun*-binding in many ways. However, though more formalised, this theory shares some similar problems with the disjunctive theory in *caki*-binding. The main problem stems from the fact that *caki* is rarely bound by any discourse entity (including the speaker and the listener), which we have seen repeatedly throughout the discussion of discourse factors.

Having presented problems in different approaches to *caki*-binding so far, an ideal theory for *caki*-binding we would like to provide in this thesis would be that it should be able to provide a *unified* account for both subject and non-subject binding of *caki*. As determining the binding relations of *caki* does not depend on the discourse level, *caki*-binding should receive a more syntactic and formal account. Our proposal will be presented within the framework of HPSG, where such a formal account will be accommodated in one level of presentation. Before directly moving on to the HPSG account of *caki*-binding, in the next chapter we will first introduce the basic background of HPSG and also discuss its Binding Theory in terms of *caki*-binding.

# Chapter 4

## HPSG and Binding Theory

In this chapter, we will first look at the basics of the HPSG framework (section 4.1), and examine the Binding theory including the classification of NPs, the notion of command and binding principles (section 4.2). The last section will focus on the binding principle for anaphors and its further developments and discuss whether it can adequately explain the binding relations of *caki* (section 4.3).

### 4.1 Head-driven Phrase Structure Grammar

HPSG, first introduced extensively in *Information-Based Syntax and Semantics: Fundamentals* by Pollard & Sag (1987), is a theory of natural language syntax and semantics which takes an *information-based* approach. The theory has been further elaborated in *Head-driven Phrase Structure Grammar* (Pollard & Sag 1994) to provide an unified account of many linguistic structures (e.g., Control Theory and Binding theory), and the theory continues to undergo extensive modification.<sup>1</sup>

---

<sup>1</sup>The theory in Pollard & Sag (1994) is divided into two parts. The fundamental ideas of the theory are presented in chapter 1-8, whereas chapter 9 titled as *Reflections and Revisions* illustrates several modifications to the theory, for example, introducing the valence features SUBJ, COMPS, SPR. Since the publication of the book, even further modification and elaboration have been put forward in various areas of grammar, for example, unbounded dependency constructions (Sag 1997, Bouma *et al.* 1998), Linking Theory (Wechsler 1995, Davis 1997), semantic representation (Copestake *et al.* 1997), argument realization (Sag & Miller To appear).

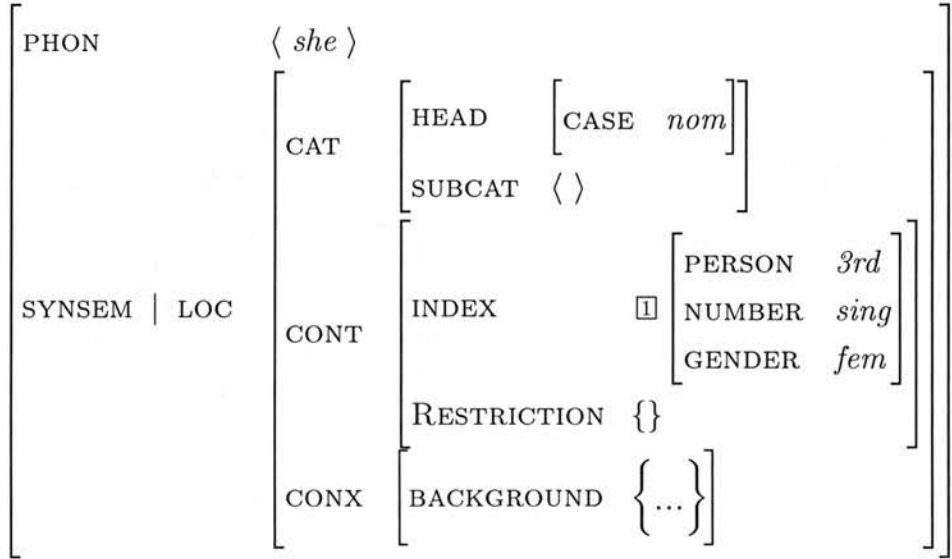
Like Categorical Grammar (CG), Generalised Phrase Structure Grammar (GPSG) and Lexical Functional Grammar (LFG), HPSG is a nontransformational and non-derivational theory. Therefore, it differs fundamentally from Government and Binding Theory, in which distinct levels of syntactic structure are derived sequentially by means of transformational operations. HPSG employs *structure-sharing* to relate parts of linguistic structure, unlike the involvement of *movement* in GB. Structure sharing, here, denotes the token identity between substructures of a given structure in accordance with lexical specifications for grammatical principles. Furthermore, HPSG deals with a comprehensive description of any *sign* of phrases and words in terms of phonological, syntactic, semantics, discourse and phrase-structure information. Despite the difference between HPSG and GB, as we will see in the following sections, many of attributes of the sign have more or less direct counterparts in GB.

This section is divided into two main subsections. In the first subsection, we will present an introduction to HPSG representations, which will help to explain the formalisation of the theory throughout the remaining part of the thesis. This subsection also briefly introduces the major modification in the theory, that is, between the theory given in chapter 1-8 in Pollard & Sag (1994) and the alterations suggested in chapter 9 in Pollard & Sag (1994), in which we will also see the changes in the formal representation of the theory. In the second subsection, we will look at the main mechanisms employed by HPSG. We will firstly introduce Immediate Dominance (ID) Schemata and two universal principles: the Head-Feature Principle (HFP) and the Valence Principle (ValP), which is followed by an example to show the operation of the Schemata and principles together. In the same subsection, we will also sketch two other important tools in the current theory: the Type-Hierarchy and Lexical Rules. This will briefly show how HPSG (as a strict lexicalist and constructionist theory) organises the rich lexical information and phrasal type system in terms of multiple inheritance hierarchies and lexical rules.

### 4.1.1 Architecture of the Framework

HPSG provides a description/representation of a linguistic object in terms of feature structures by specifying information about its **attributes**. Feature structures are more conveniently described using an **attribute-value matrix (AVM)**. A lexical entry for the pronoun *she* is presented in the form of an AVM as follows<sup>2</sup>:

(4.1) THE AVM DESCRIPTION OF THE LEXICAL ENTRY OF *she*



(Simplified from Pollard & Sag (1994, p.20, ex4))

As shown in (4.1), a *sign* is assumed to have the attributes PHONOLOGY(PHON), SYNTAX-SEMANTICS (SYNSEM), and in the case of phrases DAUGHTERS (DTRS). PHON takes a list of phoneme strings as its value to provide phonological and phonetic interpretation. The SYNSEM attribute does not correspond directly to any one level of GB syntactic structure. However, a complex of linguistic information which is included in the SYNSEM is more or less analogous to the information that is distributed between the levels of D-structure and LF in transformational models. SYNSEM also contains three attributes CATEGORY (CAT), CONTENT (CONT), and CONTEXT (CONX). CATEGORY is concerned with syntactic information about the

<sup>2</sup>Each row is an attribute-value pair, and attribute or feature names are written in SMALL CAPS. The order of the attributes is unimportant.

‘underlying’ structure and this contains two attributes HEAD and SUBCAT. The HEAD value of a sign is normally what is termed *a part of speech* and some of HEAD values have attributes of their own such as CASE for *noun* and VFORM for *verb*. SUBCAT is a specification of what other signs the sign in question must combine with in order to become *saturated*. One significant point in a SUBCAT list is that the ordering of the elements in the list relies on the traditional *obliqueness* hierarchy, rather than corresponding to surface order. That is, subjects appear first, followed by other complements (if any) in the order of primary object, secondary object, oblique PP and so on. This list plays an important part in defining HPSG binding principles, which we will see in more detail later in section (4.2). CONTENT bears information about semantic interpretation. It is the level representing the context-independent ‘meaning’ of a word or phrase. CONTEXT, on the other hand, contains the context-dependent information for semantic interpretation such as indexicality, presupposition and conventional implicature.

As we saw, the SUBCAT value of a sign is in essence the sign’s valence. That is, it specifies what arguments must be combined with the sign in order for a head to become a well-formed maximum projection. Here ‘arguments’ mean subcategorised-for elements including subjects of Ss and specifiers of NPs. In (4.1), the value of SUBCAT is  $\langle \rangle$  since the pronoun does not subcategorise for any argument, which can be contrasted with the lexical entry of a verb. The lexical entry of the verb *loves* is shown below:

(4.2) THE AVM DESCRIPTION OF THE LEXICAL ENTRY OF *loves*

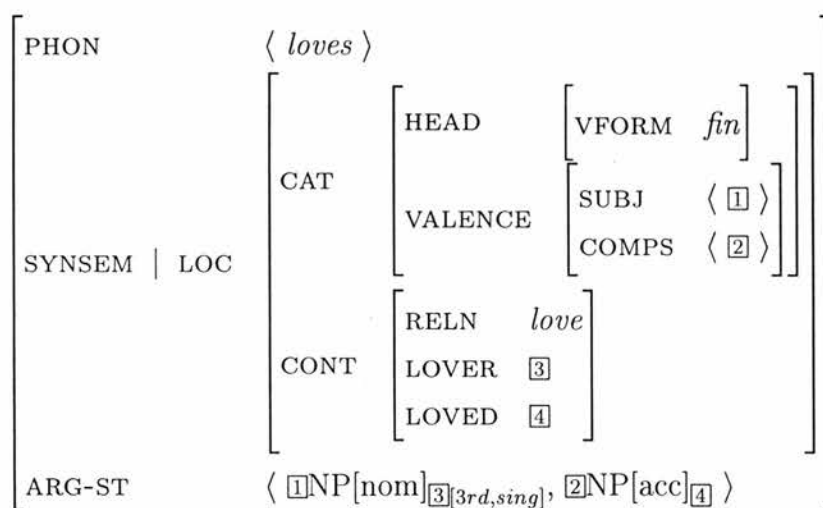
$$\left[ \begin{array}{l} \text{PHON} \\ \text{SYNSEM} \mid \text{LOC} \end{array} \left[ \begin{array}{l} \langle \textit{loves} \rangle \\ \text{CAT} \left[ \begin{array}{l} \text{HEAD} \left[ \text{VFORM} \textit{fin} \right] \\ \text{SUBCAT} \langle \text{NP}[\text{nom}]_{\boxed{1}[\textit{3rd,sing}]}, \text{NP}[\text{acc}]_{\boxed{2}} \rangle \end{array} \right] \\ \text{CONT} \left[ \begin{array}{l} \text{RELN} \textit{love} \\ \text{LOVER} \boxed{1} \\ \text{LOVED} \boxed{2} \end{array} \right] \end{array} \right] \right]$$



The SUBCAT specification indicates that this subcategorizes for a nominative subject and an accusative object. That is, the value of SUBCAT is a list of the argument(s) which the lexical head requires in order to be projected to the phrase level. An argument which is combined with the head is, therefore, cashed out from the list: the SUBCAT value of the S node is an empty list, SUBCAT $\langle \rangle$ , while that of the VP node contains only one element, SUBCAT $\langle \text{NP} \rangle$ , namely a subject. Furthermore, like all tensed verbs in English, this verb bears the VFORM value *fin(ite)*, and like all third-singular verbs, it requires that the index on the subject be [PERSON:3rd] [NUMBER:singular], which is abbreviated as [3rd, sing] in (4.2).

In Chapter 9 of Pollard & Sag (1994), following Borsley (1987)'s argument, a revised version of HPSG is proposed in which subjects and complements are distinguished by distinct corresponding features SUBJ and COMPS. SUBJ and COMPS are in the attribute VALENCE which replaces SUBCAT. CATEGORY, therefore, contains the attributes HEAD and VALENCE. The attribute ARG-ST is also added to represent the argument structure of the given word. The lexical entry of the verb *loves* now has the following representation<sup>3</sup>.

(4.3) THE REVISED VERSION OF THE LEXICAL ENTRY OF *loves*



<sup>3</sup>A preceding box indicates a SYNSEM value like  $\boxed{1}$  and  $\boxed{2}$  in (4.3). A subscripted box represents the value of CONTENT | INDEX like  $\boxed{3}$  and  $\boxed{4}$  in (4.3).



The theory assumes that only words have the ARG-ST feature, which is thus defined for lexical signs. Even though, unlike SUBCAT in the ‘standard’ version, the element(s) in the ARG-ST are not removed when the lexical head is projected into a phrasal level, it takes over many of the functions of the SUBCAT list. As we will see in section 4.2.2, it is the ARG-ST list that is argued to be the locus of binding theory and unbounded dependency constructions<sup>4</sup>.

## 4.1.2 Mechanisms and Principles

### *ID Schemata*

The combinatory possibilities of signs (feature structures) in HPSG are in principle unconstrained. Any feature structure which can unify (i.e., is compatible) with another can combine with it. To restrict this situation, constraints called Immediate Dominance (ID) schemata are stated which effectively define the legal combinations of structures. In order to construct well-formed phrases, HPSG uses seven ID schemata as presented below:

#### (4.4) ID SCHEMATA

- **Schema1:** Head-Subject Schema
- **Schema2:** Head-Complement Schema
- **Schema3:** Head-Subject-Complement Schema
- **Schema4:** Head-Specifier Schema (For combining nouns with determiners and possessives, prepositions with degree adverbials)
- **Schema5:** Head-Marker Schema (For combining VP heads with complementizers)
- **Schema6:** Head-Adjunct Schema (For combining heads with modifiers such as adjectives and adverbs)
- **Schema7:** Head-Filler Schema (Unbounded dependency constructions such as topicalisation, etc.)

---

<sup>4</sup>There still remains the debate of the adequacy of ARG-ST in the framework. More detailed discussion of ARG-ST will be given in chapter 6.

## *Universal Principles*

At the highest level of the combination which satisfies the ID schemata we saw in the last section, signs are defined as being of type *phrase*. The type *phrase* is declared as having a feature DAUGHTERS (DTRS), which encodes the structure of its constituents. The DTRS feature takes as its value a structure containing a HEAD-DTR and a list of NON-HEAD-DTRS. There are three major principles operating on HEAD-DTR: the HEAD-FEATURE PRINCIPLE (HFP), the VALENCE PRINCIPLE (ValP) and the SEMANTICS PRINCIPLE<sup>5</sup>.

(4.5) HEAD-FEATURE PRINCIPLE

The value of HEAD features of a phrase is token identical with the HEAD feature of the HEAD-DTR.

(4.6) VALENCE PRINCIPLE<sup>6</sup> (ValP, previously the Subcategorization Principle)

In a headed phrase, for each valence feature F, the F value of the HEAD-DTR is the concatenation of the phrase's F value with the list of SYNSEM value of the F-DTRS value. (Note that F ranges over the 'valence features' SUBJ, SPR and COMPS).

HPSG draws on the notion of 'subcategorization by cancellation'. The ValP checks off the subcategorization requirements of the lexical head as they become satisfied by the subject and complement daughters of its phrasal projections. Also the HFP guarantees that headed phrases are 'projections' of their heads. The percolation of the CONT value is controlled by the SEMANTICS PRINCIPLE. This principle basically works in the same way as the HFP. That is, the CONT value is passed up from the HD-DTR of a phrase with the exception of the head-adjunct schema. This is formally stated as follows:

---

<sup>5</sup>There are five more principles: the SPEC Principle, Quantifier Inheritance Principle (QIP), MARKING principle (MarkP), Semantics Principle and NON-LOCAL FEATURE Principle (NLFP). See Pollard & Sag (1994) for a detailed description of the principles.

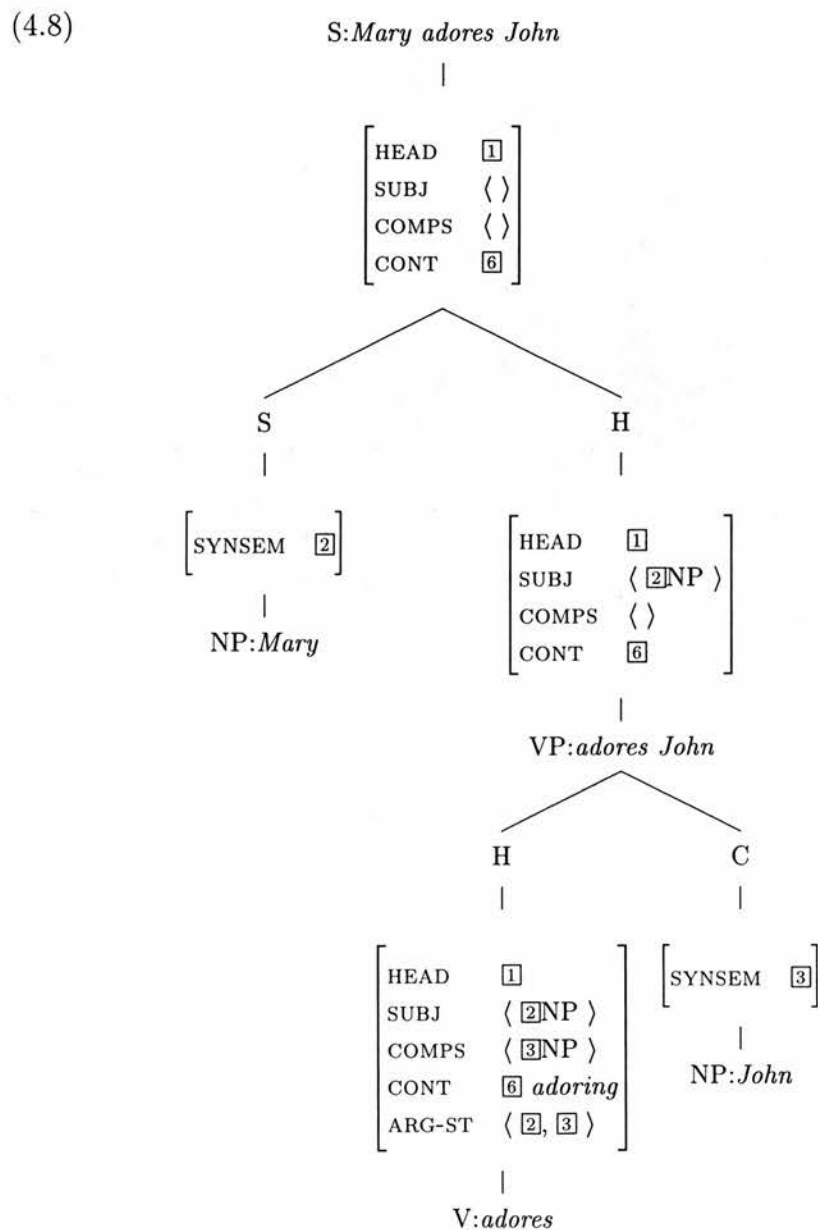
<sup>6</sup>SUBCATEGORIZATION PRINCIPLE (a simplified version from Pollard & Sag (1994, p34))

In a headed phrase, the SUBCAT value of the head daughter is the concatenation of the phrase's SUBCAT list with the list of SYNSEM values of the complement daughters.

(4.7) SEMANTICS PRINCIPLE

The value of CONTENT of a phrase is token-identical with that of the ADJUNCT-DTR if it exists, otherwise that of the HEAD-DTR.

The following illustrates the operation of the HFP, ValP, and two ID SCHEMATA: *head-comp schema* and *head-subj schema*:



Before observing the operation of the schemata and the principles, let us see how the notion of structure-sharing is represented in the AVM notation. Sharing of structure (or token-identity) is indicated by boxed numbers and letters. A box

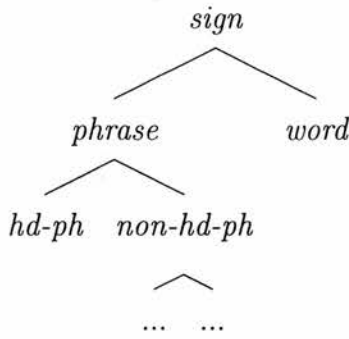
corresponds to the value of an attribute, and it appears exactly where the value would appear in the AVM. For example, at the bottom of (4.8), the value of the SYNSEM of *John* is indicated as ③ and this is structure-shared with the value of the COMPS of the head *adores*. If the value itself is also specified, it comes after the box such as ③NP in (4.8).

Now seen from the bottom level of (4.8), the sentence *Mary adores John* is completed by combining *adores* and *John* which satisfies the *head-complement schema* (the value of the attribute COMPS in the HD-DTR should be structure shared with the SYNSEM value of the COMP-DTR). Then the resulting phrase is again combined with the SUBJ-DTR satisfying the *head-subject schema* (the value of the attribute SUBJ in the HD-DTR should be structure shared with the SYNSEM value of the SUBJ-DTR). At the same time, the HFP and ValP are both satisfied: The HFP ensures that the HEAD value is passed up through the HD-DTRS, indicated by ①. Also the ValP discharges the required argument of the lexical head (*adores* in this case) in the process of combining the complement and the subject. At the bottom of the tree in (4.8), the head *adores* looks for one subject (②) and one complement (③). When the verb *adores* is combined with the complement *John*, the required complement is cancelled and the value of COMPS will appear empty at the higher node VP. The same applies to the subject requirement of the head and the SUBJ value will be empty in the top node of the tree. Note that ARG-ST is a lexical feature and remains on the lexical entry of the lexical head, that is, it is not subject to phrasal projection.

### ***Type Hierarchy***

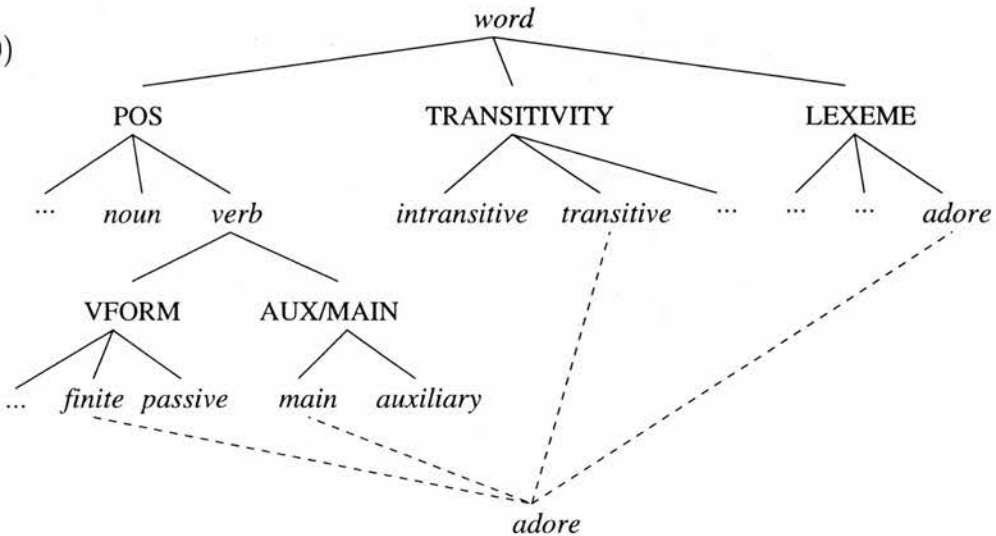
In Sag (1997), *signs* are organised into a type hierarchy, with more general super-types classified into more specific subtypes as below:

(4.9)



The types which contain more information are lower in the hierarchy, while the more general ones are higher up. These hierarchies can get more complex when the subtypes themselves have further subtypes. A subtype inherits all the features of its supertypes. In a **multiple inheritance** hierarchy, a type can be partitioned along more than one dimension. Let us have a look at the multiple inheritance hierarchy for the lexical verb *adore*:

(4.10)



Each cross-cutting classification is identified by a label indicated by capital letters. These labels are not type names; they simply provide a convenient way to refer to the different partitions. From the multiple inheritance hierarchy, the word *adore* falls into several categories; word, verb, finite verb, main verb, transitive verb. Each of these categories can be defined by particular constraints on features. For example, the constraints imposed on the verb *adore* here are as follows

(4.11) a. Verbs: The head feature should be *verb*;

- b. Finite verbs: they have subjects; subjects have nominative case;
- c. Main verbs: they do not invert in English;
- d. Transitives: direct objects have accusative case.

If we organise these categories in the multiple inheritance hierarchy and allow *adore* to inherit all of these constraints, then the specific, truly idiosyncratic information that the lexical entry for *adores* must contribute is minimised as follows:

$$(4.12) \quad \begin{array}{l} \text{word} \wedge \text{verb} \wedge \text{fin-vb} \wedge \text{main-vb} \wedge \text{trans} \wedge \\ \left[ \begin{array}{l} \text{LEXEME} \quad \textit{adore} \\ \text{SUBJ} \quad \left\langle \left[ \text{INDEX} \quad \boxed{1} \right] \right\rangle \\ \text{COMPS} \quad \left\langle \left\{ \text{INDEX} \quad \boxed{2} \right\} \right\rangle \\ \text{CONT} \quad \left[ \begin{array}{l} \textit{adoring} \\ \text{ADORER} \quad \boxed{1} \\ \text{ADORED} \quad \boxed{2} \end{array} \right] \end{array} \right] \end{array}$$

The only information specific to *adore* is the lexeme, and the semantic roles it assigns to its subject and direct object. In fact, since this information is shared by all the inflected forms of *adore* (*adores*, *adored*, *adoring*), we can identify a **lexeme type** for *adore*, defined by the AVM in (4.12). The analysis of many syntactic constructions benefit from the application of the multiple inheritance hierarchy such as English Relative Clause Constructions (Sag 1997) and Causative constructions (Davis 1997).

### *Lexical Rules*

As we have seen, the lexicon is organised into subtypes, or groups of lexical items. There are many cases where groups of words seem to be systematically related, for example, singular vs. plural nouns, or present vs. past tense verbs. Lexical rules are one way to capture the relationship between these groups of lexical entries. Procedurally, a lexical rule takes a lexical entry as input, and gives another lexical entry as output. For example, we could define a pluralisation lexical rule that added

plural morphology and modified the semantics of a singular noun to produce the plural form.

Even for syntactic changes, since HPSG is a lexicalist theory, any operations that change lexical entries must be accounted for in the lexicon (not in the syntax). For example, passivization can be analysed using lexical rules that change VALENCE list values.

(4.13) THE PASSIVE LEXICAL RULE<sup>7</sup>

$$\left[ \begin{array}{l} \textit{active-vb} \\ \text{VALENCE} \left[ \begin{array}{l} \text{SUBJ} \quad \langle \text{NP}_{\boxed{2}} \rangle \\ \text{COMPS} \quad \text{append}(\langle \boxed{3} \rangle, \boxed{4}list) \end{array} \right] \end{array} \right] \Rightarrow \left[ \begin{array}{l} \textit{pas-vb} \\ \text{HEAD} \mid \text{VFORM} \quad \textit{pas} \\ \text{VALENCE} \left[ \begin{array}{l} \text{SUBJ} \quad \langle \boxed{3} \rangle \\ \text{COMPS} \quad \text{append}(\boxed{4}, \langle (\text{PP}[\textit{by}])_{\boxed{2}} \rangle) \end{array} \right] \end{array} \right]$$

The lexical entry for a passivized verb (output) indicates that the first member of the COMPS list is promoted to the SUBJ list, while the subject is demoted to an (optional) *by*-phrase complement. Notice that the CONTENT of the verb is unchanged. The following illustrates the passivization of the verb *adore*, which can be input to the Passive Lexical Rule:

$$(4.14) \quad \left[ \begin{array}{l} \text{VAL} \left[ \begin{array}{l} \text{SUBJ} \quad \langle \text{NP}_{\boxed{2}} \rangle \\ \text{COMPS} \quad \langle \boxed{3}\text{NP}_{\boxed{5}} \rangle \end{array} \right] \\ \text{CONT} \left[ \begin{array}{l} \textit{adoring} \\ \boxed{6}\text{ADORER} \quad \boxed{2} \\ \text{ADORED} \quad \boxed{5} \end{array} \right] \end{array} \right] \Rightarrow \left[ \begin{array}{l} \text{VAL} \left[ \begin{array}{l} \text{SUBJ} \quad \langle \boxed{3}\text{NP}_{\boxed{5}} \rangle \\ \text{COMPS} \quad \langle (\text{PP}_{\boxed{2}}) \rangle \end{array} \right] \\ \text{CONT} \left[ \begin{array}{l} \textit{adoring} \\ \boxed{6}\text{ADORER} \quad \boxed{2} \\ \text{ADORED} \quad \boxed{5} \end{array} \right] \end{array} \right]$$

<sup>7</sup>'append' (often indicated by  $\oplus$ ) is one of the functional or relational symbols employed in AVM descriptions. It means 'combining one list with another without altering the order of elements in each list (as opposed to 'shuffle', or 'o', which allows the order of the elements in the list to be altered (Bouma *et al.* 1998, fn.5).)



Similar kinds of lexical rules are used for other syntactic constructions such as dative alternation and middle constructions in English. They also play a central role in providing a traceless and lexically based account of unbounded dependency constructions (Pollard & Sag 1994, Sag & Fodor 1994, Bouma *et al.* 1998), an issue we will reconsider in section 6.4.

Although Pollard & Sag (1987) claim that lexical rules are ‘neutral between the declarative and procedural interpretations’ (p.209, fn.11), Pollard & Sag (1994) admit that ‘we lack as yet any satisfactory declarative formalisation’ (p.395, fn.1). This is clearly a problem for HPSG, and for other non-derivational frameworks that use lexical rules (e.g., LFG). It is likely that further investigation into the formal properties of lexical rules will secure their place in HPSG. Much recent work, however, has focused on alternatives to lexical rules (see Kathol (1994) on passivization, and Eynde (1995) on extraposition<sup>8</sup>).

## 4.2 HPSG Binding Principles

In this section, we will look at Binding Theory within HPSG. We will first start with the NP typology, which will be followed by the illustration of the notion of *obliqueness*, in terms of which the generalisations about constraints on the binding are stated. At the same time, the notion of ARG-ST will be briefly recapitulated because it is argued to be the level at which binding relations are determined. Then we will move to the notion of *o-command* which is analogous to *c-command* in GB. Lastly, we will present the formal Binding Principles in HPSG. As HPSG Binding theory relies on the notion of the relative hierarchical order (obliqueness), it sharply contrasts with the GB Binding Theory which is based on configurational superiority (*c-command*). Even though the contrast between the two different Binding theories is intriguing to discuss, having looked at GB Binding Theory in section 3.1.2 and 3.1.3, we will leave the readers to refer to chapter 6 in Pollard & Sag (1994) which

---

<sup>8</sup>Many of the generalisations captured by lexical rules can be also be represented directly in the lexical hierarchy. See Koenig & Jurafsky (1995) and Krieger & Nerbonne (1993) for examples of this approach.

contains a detailed discussion of the distinction and different predictions of the binding principles in the two theories.

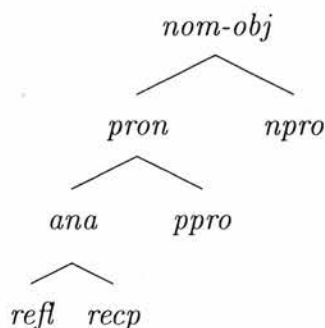
### 4.2.1 Types of NPs

To begin with, let us recall that the semantics of a sign (word or phrase) is encoded as the value of its *CONTENT*. And the *CONTENT* value of every NP is an object of the sort of *nominal-object* whose internal structure is given below:

$$(4.15) \quad \left[ \begin{array}{c} \text{CONTENT} \\ \left[ \begin{array}{c} \text{INDEX} \\ \text{RESTR} \end{array} \right] \end{array} \right] \left[ \begin{array}{c} \text{nom-obj} \\ \left[ \begin{array}{c} \text{index} \\ \text{PER} \\ \text{NUM} \\ \text{GEND} \end{array} \right] \\ \{ \dots \} \end{array} \right] \left[ \begin{array}{c} \\ \textit{per} \\ \textit{num} \\ \textit{gend} \end{array} \right]$$

The classification of nominal objects into sorts is on the basis of the referential properties of the NPs that bear them as below:

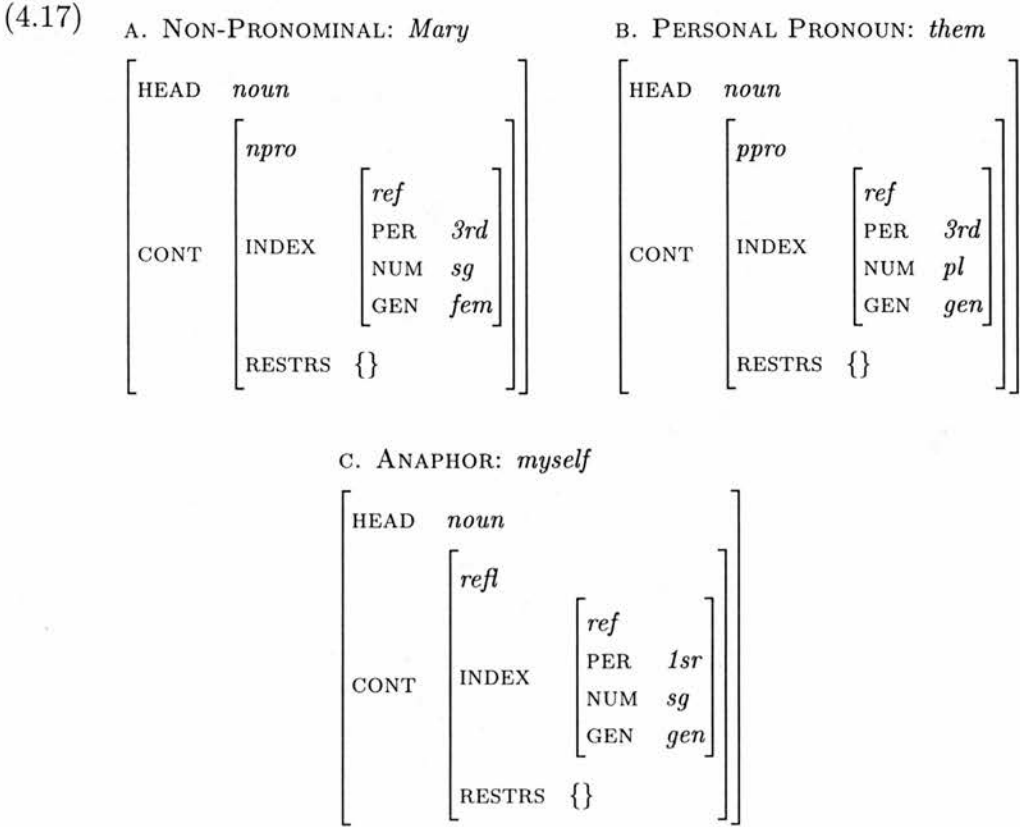
(4.16) Subtypes of *nom-obj*:



Overt nominals are divided into three types of nominal objects: *npro*, *ana* and *ppro*, which correspond to NPs in the GB framework as R-expressions, anaphors, and pronominals, respectively<sup>9</sup>. In HPSG, as the theory tries to avoid any postulation

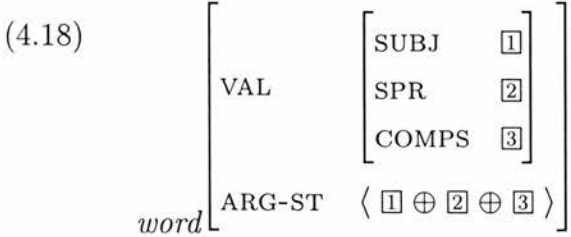
<sup>9</sup>Recall that in GB, NPs are classified by the binary features [+/- anaphoric] and [+/- pronominal]. Pollard & Sag (1994) questions and presents evidence against this way of classifying NPs. See Section 6.3. in Pollard & Sag (1994)

of an empty category, there are no types corresponding to empty NP like PRO or NP variables. The following illustrates the internal structure of each NP type:



### 4.2.2 Obliqueness and ARG-ST

In HPSG, Binding Theory is defined as constraints on the ARG-ST lists of lexical heads. So far we have assumed that the ARG-ST list is simply the concatenation of the valence lists:



In principle, however, the elements in ARG-ST are not necessarily identical to the elements on the valence lists, and furthermore, they may appear in a different order. The ordering of elements on ARG-ST is assumed to be based on obliqueness. HPSG

adopts the assumption that grammatical relations (such as subject, direct object, indirect object) are ordered hierarchically. The subject is considered the least oblique element, followed by the direct object, then the second (indirect) object, and finally other PP objects and sentential complements. The obliqueness hierarchy seems to be involved in a wide range of linguistic phenomena. For example, passivization can be thought of as a process that suppresses the least oblique constituent (i.e., the subject) while promoting the others. Keenan and Comrie (1977) propose a similar “Accessibility Hierarchy” to express some implicational language universals about relativization and verbal agreement.

(4.19) SUBJECT  $\gg$  DIR OBJ  $\gg$  IND OBJ  $\gg$  OBLIQUES  $\gg$  GENITIVES  $\gg$  OBJECTS  
OF COMPARISON

### 4.2.3 O-commanding and O-binding

While the GB approach to binding relies on the configurational notions of c-command and government, HPSG Binding Theory uses the relation obliqueness command, or o-command.

(4.20) LOCAL O-COMMAND

If  $Y$  and  $Z$  are *synsem* objects and  $Y$  is referential, then  $Y$  locally o-commands  $Z$  iff  $Y$  is less oblique than  $Z$ . that is,  $Y$  precedes  $Z$  on some ARG-ST list.

The following two examples are given to illustrate how local o-command works. Note that a referential NP has an INDEX value of type *referential* (*ref*), which is only subject to local o-command. This then immediately excludes expletives like *it* or *there*.

(4.21) John adores Mary.

$$\left[ \begin{array}{l} \textit{adores} \\ \text{ARG-ST} \langle \boxed{1}\text{NP}_{ref}, \boxed{2}\text{NP}_{ref} \rangle \end{array} \right]$$

The *synsem* object labelled ① locally o-commands the one labelled ②, but not vice versa.

- (4.22) There are two students absent.
- $$\left[ \begin{array}{l} \textit{are} \\ \text{ARG-ST} \langle \text{①NP}_{\textit{there}}, \text{②NP}_{\textit{ref}}, \text{③AP} \rangle \end{array} \right]$$

The expletive NP labelled ① is the least oblique item, but because it is non-referential, it does not locally o-command ② or ③. The referential NP ②, on the other hand, does locally o-command ③. Local o-command operates on the co-arguments of the head and is a special case of o-command of a more general relation called simply **o-command**. The definition of o-command is as follows:

- (4.23) If Y and Z are *synsem* and Y is referential, then Y o-commands Z iff Y locally o-commands some X dominating Z.

Here, ‘domination’ is a tree-configurational notion. Suppose X’ is a sign having X as its SYNSEM value and Z’ is a sign with Z as its SYNSEM value. Then X dominates Z iff X’ contains Z’ somewhere in its constituent structure (DTR value). The following example illustrates the notion of o-command:

- (4.24) John thinks Kim adores Mary.
- $$\left[ \begin{array}{l} \textit{‘thinks K adores M’} \\ \text{HD-DTR} \left[ \begin{array}{l} \textit{head-comp-struct} \\ \text{ARG-ST} \langle \text{①NP}_{\textit{ref}} \textit{‘John’}, \text{②S} \rangle \end{array} \right] \\ \text{COMP-DTRS} \left\langle \begin{array}{l} \textit{‘K adores M’} \\ \text{SYNSEM} \quad \text{②} \\ \text{DTRS} \left[ \begin{array}{l} \textit{head-subj-struct} \\ \text{SUBJ-DTR} \langle \text{③NP} \textit{‘kim’} \rangle \end{array} \right] \end{array} \right\rangle \end{array} \right]$$

The subject NP ① *John* o-commands ③ '*Kim*' because (a) it locally o-commands the embedded s ② '*K adores M*' and (b) the sentence ② dominates ③. Corresponding to A-binding in GB, we have **o-binding**, which is coindexation under o-command:

- (4.25) Y (locally) o-binds Z iff Y and Z structure share INDEX values and Y (locally) o-commands Z. Any *synsem* object Z that is not (locally) o-bound is (locally) o-free.

Also note that the index of a nominal-object has internal structure (indicated as a subscripted AVM box), namely, the features PERSON, NUMBER, GENDER as in the following:

- (4.26) ANAPHOR: *himself*<sub>①</sub>  

$$\left[ \begin{array}{l} \text{CONT} \mid \text{INDEX} \quad \text{①} \left[ \begin{array}{ll} \text{PER} & 3rd \\ \text{NUM} & sg \\ \text{GEN} & mas \end{array} \right] \end{array} \right]_{ref}$$

It is token-identity (structure-sharing) of indices that corresponds to the notion of coindexing for NPs. Therefore, coindexed NPs (the binder and the bindee) necessarily bear identical specifications for person, number and gender.

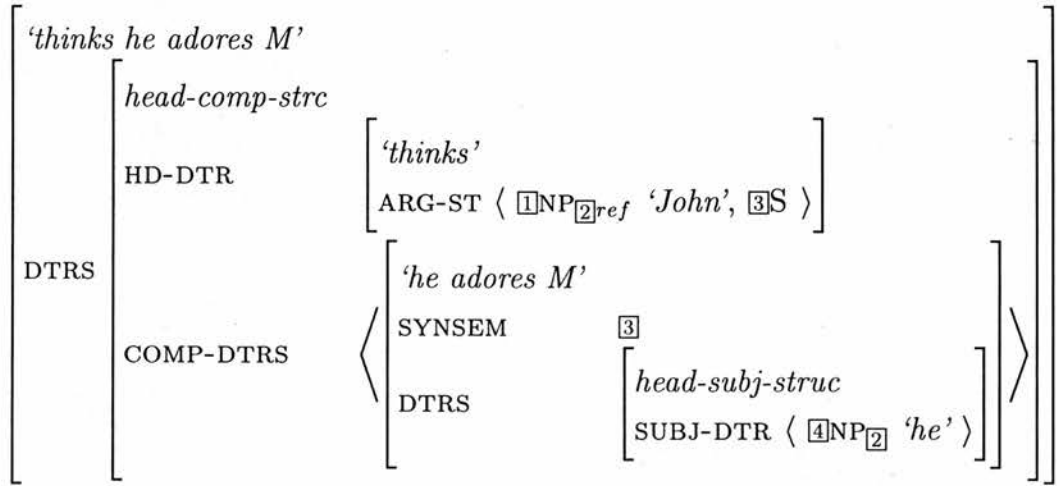
Now let us have a look at the following two examples to illustrate (local) o-binding:

- (4.27) *John*<sub>i</sub> adores *himself*<sub>i</sub>  

$$\left[ \begin{array}{l} adores \\ \text{ARG-ST} \langle \text{①NP}_{2ref}, \text{③NP}_{2ref} \rangle \end{array} \right]$$

① and ③ are coindexed, and ① locally o-commands ③. Therefore, ① locally o-binds ③. On the other hand, ① is locally o-free, since nothing o-commands it.

(4.28) John<sub>i</sub> thinks he<sub>i</sub> adores Mary.



The main clause subject  $\boxed{1}$  ‘John’ o-binds the embedded subject  $\boxed{4}$  ‘he’, but it does not locally o-bind it.  $\boxed{4}$  is o-bound, but it is locally o-free. Also note that since INDEX values in HPSG encode person, number, and gender, coindexing immediately accounts for pronoun-antecedent agreement.

#### 4.2.4 HPSG Binding Principle

There are three binding principles, analogous to the GB Binding theory:

- (4.29)
- **Principle A:** A locally o-commanded anaphor must be locally o-bound
  - **Principle B:** A personal pronoun must be locally o-free
  - **Principle C:** A non-pronoun must be o-free.

Let us have a look at the following three examples with the relevant AVMs which illustrate the binding principles:

- (4.30) \*We know herself
- $$* \left[ \begin{array}{l} \textit{know} \\ \text{ARG-ST} \langle \boxed{1}\text{NP}_{\boxed{2}\textit{ppro}}, \boxed{3}\text{NP}_{\boxed{4}\textit{refl}} \rangle \end{array} \right]$$



The ARG-ST list of *know* contains a reflexive NP ③ that is locally o-commanded (by ①) and therefore subject to Principle A. But ③ is not locally o-bound, because it is not coindexed with a local o-commander, and this violates Principle A.

$$(4.31) \quad *Mary_i \text{ likes } her_i$$

$$* \left[ \begin{array}{l} \textit{likes} \\ \text{ARG-ST} \langle \text{①NP}_{\text{②:npro}}, \text{③NP}_{\text{②:ppro}} \rangle \end{array} \right]$$

In this ARG-ST list we have the personal pronoun *her* (③) coindexed with a less oblique *synsem* element ①. Therefore, the pronoun ③ is locally o-bound, in violation of Principle B.

$$(4.32) \quad *He_i \text{ loves } John's_i \text{ parents}$$

$$* \left[ \begin{array}{l} \textit{'loves } J's \textit{ parents'} \\ \text{DTRS} \left[ \begin{array}{l} \textit{head-comp-struct} \\ \text{HD-DTR} \mid \text{ARG-ST} \langle \text{①NP}_{\text{②:ppro}}, \text{③NP} \rangle \\ \text{COMP-DTRS} \left\langle \begin{array}{l} \textit{'J's parents'} \\ \text{SYNSEM} \quad \text{③} \\ \text{DTRS} \left[ \begin{array}{l} \textit{head-spr-struct} \\ \text{SPR-DTRS} \langle \text{④NP}_{\text{②:npro}} \rangle \end{array} \right] \end{array} \right\rangle \end{array} \right] \end{array} \right]$$

Here we have a non-pronominal NP ④ '*John*', which is dominated by the NP ③ '*John's parents*'. ③ is locally o-commanded by the subject NP① '*he*'. So ① o-commands ④, and since ① and ④ are coindexed, ① o-binds ④. But by Principle C, non-pronominals like ④ must be o-free, so this sentence is ungrammatical.

#### 4.2.5 Exempt anaphors

Principle A of the HPSG Binding Theory is implicational. The following statement is equivalent to the one in (4.29) above:

$$(4.33) \quad \text{If an anaphor is locally o-commanded, then it must be locally o-bound}$$

What about anaphors (reflexives and reciprocals) that are not locally o-commanded? These are exempt from Principle A, and the binding theory (as it stands) puts no constraints on their distribution. For example, anaphors in subject position are not locally o-commanded, because they appear first on the head's AGR-ST list:

- (4.34) \*Himself drinks Whisky.  

$$\left[ \begin{array}{l} \textit{drinks} \\ \text{ARG-ST } \langle \text{NP: } \textit{ana}, \text{NP} \rangle \end{array} \right]$$

Since the NP *himself* is not locally o-commanded, it is not required to be locally o-bound. Pollard & Sag (1994) account for the ungrammaticality of this sentence by assuming that reflexives are always accusative. The relevant lexical entry for *himself* is:

- (4.35) 
$$\left[ \begin{array}{l} \text{PHON} \quad \langle \textit{himself} \rangle \\ \text{HEAD} \quad \left[ \begin{array}{l} \textit{noun} \\ \text{CASE} \quad \textit{acc} \end{array} \right] \\ \text{CONT} \quad \left[ \begin{array}{l} \textit{refl} \\ \text{INDEX} \quad \left[ \begin{array}{l} \text{PER} \quad \textit{3rd} \\ \text{NUM} \quad \textit{sg} \\ \text{GEN} \quad \textit{masc} \end{array} \right] \end{array} \right] \end{array} \right]$$

Since *drinks* requires a nominative subject, the reflexive pronoun will not be able to appear grammatically in sentence (4.34). Here are some other examples of exempt anaphors (from P&S 94, p. 263):

- (4.36) a. The children<sub>i</sub> thought that [each other's]<sub>i</sub> pictures were on sale.  
 b. [The children]<sub>i</sub> thought that [pictures of [each other]<sub>i</sub>/themselves<sub>i</sub> were on sale.  
 c. John suggested that tiny gilt-framed portraits of [each other]<sub>i</sub> would amuse [the twins]<sub>i</sub>.

In (4.36a), the reciprocal NP is the specifier of the noun *pictures*, and in (4.36b), it is the complement. In both cases, the reciprocal is the first element of the ARG-ST list of *pictures*. Since it is not locally o-commanded, it is exempt from Principle A. The example in (4.36c) seems to confirm this analysis, because the anaphor is not o-bound by a less oblique NP. If *each other* were not exempt, the sentence would be ruled ungrammatical as a Principle A violation.

Exempt anaphors are also found in expletive constructions (Pollard & Sag (1994), p. 259), as shown below with the relevant AVM of the ARG-ST list of the head ‘*was*’:

(4.37) ?John<sub>i</sub> was devastated by the loss of his entire family. Now there was only himself<sub>i</sub> remaining.

(4.38) ARG-ST ⟨ NP<sub>there</sub>, NP:*ana*, AP[+PRD] ⟩

Since the first NP (*there*) is non-referential, it does not locally o-command the other elements in the list. So the anaphor *himself* is exempt. The example sentence in (4.37) confirms that Principle A does not apply, because the antecedent of the reflexive occurs in a different sentence altogether. Note that P&S do not suggest that exempt anaphors are completely unrestricted. There seem to be complex constraints on their distribution and reference. Pollard & Sag (1994) suggest that exempt anaphors are subject to nonsyntactic factors in their distributions such as processing, intervention (such as expletive intervenors (Kuno 1987)), perspective (Iida 1996) and a hierarchy of thematic relations (Jackendoff 1972, Wilkins 1988, Hellan 1988). Though the analysis of the factors that are responsible for the distribution of exempt anaphors is not fully developed, Pollard & Sag (1994) conclude that the constraints on the such factors fall outside Binding Theory.

In this section, we have outlined the HPSG Binding Theory in detail. After the extensive speculation of cross-linguistic observation, the Binding Theory in Pollard & Sag (1994) proved to be adequate for English data, and the theory has undergone more developments in the current binding theory. In the next section, we will

evaluate the binding theory in HPSG in terms of *caki* binding. We will start with the general problems of the original form of principles in accounting for *caki*-binding, and then we will also review and evaluate two other developments: one in Xue *et al.* (1994) and the other in Manning (1996) and Manning & Sag (1998a).

## 4.3 HPSG Binding Theory and *Caki*

### 4.3.1 O-Binding for *caki*-Binding

Binding Principle A in HPSG, as it is, is clearly problematic for long-distance anaphors like *caki*. This is primarily because principle A for reflexives restricts its binding domain to be local. Principle A is repeated as below:

(4.39) **Principle A:**

A *locally* o-commanded anaphor must be *locally* o-bound

In this principle, the antecedent of the reflexive can only be a *local* o-binder which *locally* o-commands the reflexive. However, we have seen that *caki* can be bound by an antecedent outside a clausal boundary, or in the HPSG context, by a o-binder (or o-commander). For the current Principle to be able to cover this extended binding domain for long-distance anaphors, we can, of course, establish a parameterised principle which states that a long-distance anaphor must be bound by an o-commander. Such an attempt has been made by Xue *et al.* (1994) who have proposed a binding principle to control long-distance anaphor, *ziji*, in Chinese. Their principle, PRINCIPLE Z, is formally stated below:

(4.40) **Principle Z:**

A locally o-commanded long-distance anaphor must be o-bound.

This Principle allows a long-distance anaphor to be bound either by a local o-binder or an o-binder. Thus, this is a fundamentally distinct constraint from Principle B which blocks any local o-binding. This seems to capture successfully the property of long-distance anaphora having an extended binding domain. This will explain cases

of long-distance subject binding, and given that Principle Z is proposed to account for Chinese *ziji* which is bound exclusively to subjects, it proves to be adequate for *ziji*-binding. However, it only partly satisfies the case of *caki*-binding, that is, it explains only cases of subject-binding, but not non-subject binding in both the local and the extended binding domains.

A further problem comes from the notion of *obliqueness* which has a key role in defining the binding principles in HPSG. Leaving aside the issue of the extended binding domain or cases of non-subject binding, defining local binding relations based on obliqueness results in some complex problems. Firstly, consider the following examples in English:

- (4.41) a. Mary talked to John about himself  
 b. \*Mary talked about John to himself  
 c. John talked to Mary about himself

The three examples above involve PPs headed by nonpredicative ('case-marking') prepositions. For such PPs to make the properties of their internal NPs visible in ARG-ST, HPSG assumes that the head preposition makes no contribution to the CONTENT of the PP, the CONTENT value of such prepositions is structure-shared with that of their NP. With this analysis, the ARG-ST list of *talked* in (4.41a) would thus look something like:

- (4.42) ARG-ST  $\langle$  NP:<sub>Mary</sub>, PP[*to*]:<sub>John</sub>, PP[*about*]:<sub>ana</sub>  $\rangle$

The PP[*about*] is locally o-commanded by both the NP<sub>Mary</sub> and the PP<sub>John</sub> and only the latter has a compatible INDEX value. Thus, the sentence is unambiguously grammatical. (4.41b) is ruled out on the basis of the relative obliqueness of the PP complements. Pollard & Sag (1994) argue that a PP[*to*] is always less oblique than a PP[*about*], and thus, regardless of the absence of linear precedence constraints on their ordering, their (obliqueness) ordering on the ARG-ST remains constant.

As for (4.41c), the PP[*about*] inherits reflexive semantics from its NP object to produce a ARG-ST list as below:

(4.43) ARG-ST  $\langle$  NP:*John*, PP[*to*]:*Mary*, PP[*about*]:*ana*  $\rangle$

This is properly locally o-commanded by the NP:*John* and is therefore correctly predicted to be grammatical.

Now, let us see whether the same logic applies to the equivalent examples in Korean, which we have seen already in chapter (2):

- (4.44) a. \*John<sub>i</sub>-i Bill<sub>j</sub>-ekey *caki<sub>j</sub>*-etayhayse malhayssta  
 John-NOM Bill-DAT **self**-about talked  
 ‘John<sub>i</sub> talked to Bill<sub>j</sub> about **self**<sub>\*j</sub>’
- b. John<sub>i</sub>-i Bill<sub>j</sub>-etayhayse *caki<sub>i/j</sub>*-ekey malhayssta  
 John-NOM Bill-about **self**-DAT talked  
 ‘John<sub>i</sub> talked about Bill<sub>j</sub> to self<sub>i/j</sub>’
- c. John<sub>i</sub>-i Bill<sub>j</sub>-ekey *caki<sub>i</sub>*-etayhayse malhayssta  
 John-NOM Bill-DAT **self**-about talked  
 ‘John<sub>i</sub> talked to Bill<sub>j</sub> about **self**<sub>i</sub>’

Compared with the English examples in (4.41) where the binding patterns were seen to comply with the obliqueness hierarchy, (4.44c) has the same binding relations. However, the judgements on (4.44a) and (4.44b) contrasts with that of (4.41a) and (4.41b). Why should this be? If we observe the binding patterns in (4.44) carefully, the PP[*to*] does not seem to be able to bind anything and it behaves as if it is more oblique than the PP[*about*], which contradicts Pollard & Sag’s (1994) assumption on the obliqueness hierarchy. At this stage, we may speculate that the obliqueness hierarchy can be parametrised. However, while it is not very desirable to assume that each language may have a different obliqueness hierarchy, which lacks generalisation in predicting binding patterns of a language, there is evidence even in English to indicate that the notion of obliqueness may not be the only factor that we should rely on in predicting the correct binding relations. Consider the following examples:

(4.45) \*Mary talked about himself to John

(4.45) has an identical ARG-ST with that given in (4.42) but the indicated binding relation is not allowed. This type of example remains problematic for HPSG where binding possibilities reflect an obliqueness hierarchy of surface grammatical relations; Manning & Sag (1998*a*) admit that ‘consideration of typologically diverse languages show that this consensus is mistaken’ (p6). They then promote the notion of ARG-ST further to bring in the logical concept of grammatical functions which is clearly distinct from surface grammatical relations. I will sketch and evaluate this new development of the HPSG binding theory in the next section.

### 4.3.2 New Development: A-Binding

Manning (1994) argues that syntax should make a clean distinction between two levels: a level of surface grammatical relations and a level of syntactic argument structure, both of which have separate prominence rankings. In more recent work in Manning (1996), he also argues that the more semantic properties of binding are sensitive to prominence at a level of ‘argument structure’. He observed that in many languages, reflexives cannot be bound by just any less oblique NP, but rather their antecedent is restricted to what might be loosely called ‘subjects’. Given that binding theory is defined on argument structure, the natural constraint to suggest is that in those languages, reflexives must be bound by the first element on some argument structure list. This idea is formalised as the definition and principle proposed in Manning (1996) and Manning & Sag (1998*a*):

- (4.46) a. An **a-subject** is an entity that is least oblique at some level of a-structure
- b. A-subject principle: Anaphors must be a-subject-bound

The class of a-subjects overlaps, but differs from, other notions in common use: external arguments are a-subjects, but since the a-subject is simply the most prominent argument of the predicate, an a-subject can be an internal argument, as in the case of unaccusative verbs. The concept of a-subject is thus similar to the “logical

subject” (Jespersen 1924), and again all logical subjects are a-subjects, but Manning and Sag argue that the compound argument structures that occur as a result of passives and causatives yield lexical forms with nested argument structures and thus multiple a-subjects, where for Jespersen only the agent argument of a passive is the logical subject.

The notion of a-subject in A-binding, to a certain extent, resolves the problems that persist with O-binding. For example, it explains binding facts in causative constructions in Korean, illustrated below:<sup>10</sup>:

- (4.47) Mary<sub>i</sub>-ka Bill<sub>j</sub>-ekey caki<sub>i/j</sub>-lul chiryoha-keyhayssta  
 Mary-NOM Bill-DAT self-ACC treat-made  
 ‘Mary<sub>i</sub> made Bill<sub>j</sub> treat self<sub>i/j</sub>’

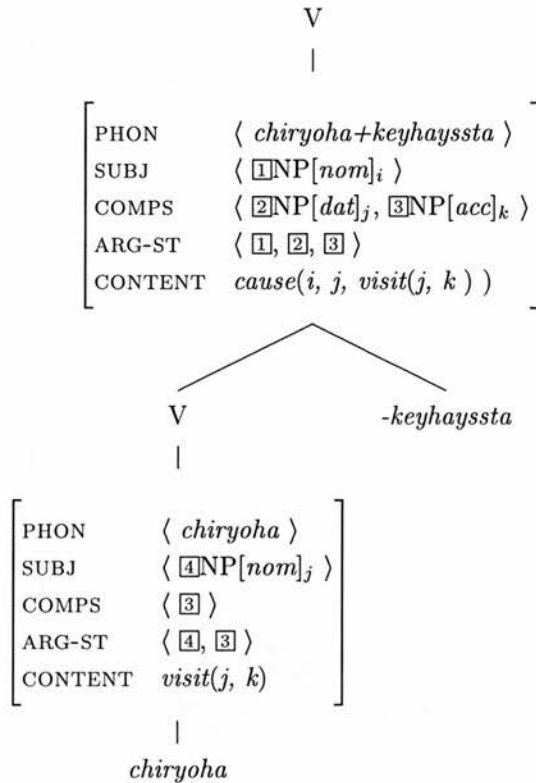
Here, *caki* is bound by either the causer *Mary* or the causee *Bill*. The complex word structure for the causative of the simple transitive verb *chiryoha-* (‘treat’) is as follows:

---

<sup>10</sup>The presentation of the following illustrations is drawn from Manning *et al.* (1999).



(4.48)



Here, the lower object NP (③) is present on both the lower and higher COMPS lists. The subject of the lower list is however distinct from, but coindexed with, the dative NP on the higher list. This coindexing has the effect of identifying this element on the upstairs ARG-ST list semantically with the subject argument of the lower verb stem. In (4.48), we have multiple a-subjects (first on each ARG-ST): One is *Mary* as the first element in the higher ARG-ST and the other is *Bill* as the first element in the lower ARG-ST. As *caki* is bound by either *Mary* and *Bill*, the notion of a-subject predicts the correct binding relation. Note also that although the causee (②) has the properties of an a-subject, it never has the properties of a subject in terms of grammatical relations or valence list positions. The binding predictions in this examples are derived from the nested ARG-ST which should also look like the following (Manning & Sag 1998a):

(4.49) ARG-ST  $\langle \textcircled{1}\text{NP}_i, \textcircled{2}\text{NP}_j, \textcircled{4}\langle \text{PRO}_j, \textcircled{3}\text{NP}_{i/j} \rangle \rangle$

This nested ARG-ST is licensed by the constraints on the underlyingly biclausal causative constructions stated in the type-hierarchy. This allows the binding relations in two ARG-STs in one representation. PRO acts as a placeholder element in ARG-ST lists and is used to mark positions coindexed with an element higher in the ARG-ST which is needed for binding. Note also that PRO here should not be confused with the empty category in GB terms. More discussion of the nested ARG-ST will be discussed in section 5.1.3.

The development of A-binding as a whole may be an improvement on o-binding at least in resolving the restricted effect of the obliqueness of surface grammatical relations. Further, I agree that the binding in constructions like biclausal causatives and passives is neatly explained away by the notion of a-subjects. However, it is quite puzzling that the analysis ignores or does not mention the well known facts of non-subject binding. For example, in the following example, there is no way we can construe the non subject binder *John* as a-subject of any given predicate:

- (4.50) Mary<sub>i</sub>-ka John<sub>j</sub>-uropwute [**caki**<sub>i/j</sub>-ka Bill<sub>k</sub>-lul hyemohanta]-ko tulessta  
 Mary-NOM John-from self-NOM Bill-ACC hate-COMPL heard  
 ‘Mary<sub>i</sub> heard from John<sub>j</sub> self<sub>i/j</sub> hates Bill<sub>k</sub>’

Furthermore, as will become clearer later in the discussion of the issue of ARG-ST in section 5.1.3, unlike (4.47), structures such as (4.50) are difficult to represent in terms of a nested ARG-ST. This means that the nominative *caki* in the lower clause does not have any means to be linked to the higher a-subject *Mary*. At the same time, recall also that exemption is stated on the basis of the anaphor being in initial position (by not being locally c-commanded) and any factors on these exempt anaphors are said to be outside the domain of binding theory. In this situation, *caki* in (4.50) has to be categorised as an exempt anaphor as it comes as the first element in its ARG-ST. Then the binding possibility between *caki* and the matrix subject *Mary* should be explained by unknown factors outside binding theory. However, when there exist legitimate antecedents within the sentential domain, involving mysterious factors to explain the binding relation seems to be far fetched, and thus at least inelegant.

A further problem on the same note can be also found where *caki* is in sentence initial position, which truly demonstrates the type of exempt anaphor. Manning *et al.* (1999) present the Japanese examples (4.51a); the equivalent Korean example is given in (4.51b).

- (4.51) a. Japanese  
 Zibun-zisin<sub>i</sub> ga hihan s-are-ta koto ga Taroo<sub>i</sub> o  
 self NOM criticism do-PASS-PAST COMP NOM Taroo ACC  
 nayamase-teiru  
 bother-PROG-PRES  
 (lit.) ‘The fact that self was criticised bothers Taroo’  
 [p20, ex.42](Manning *et al.* 1999)
- b. Caki<sub>i</sub>-ka pipan-patassta-nunkes-i John<sub>i</sub>-ul koiryophiessta  
 self-NOM criticism-received-COMPL-NOM John-ACC bothered  
 ‘That Self<sub>i</sub> was criticized bothered John<sub>i</sub>’

Manning *et al.* (1999) argue that the basic prediction of HPSG binding theory that anaphors with no local a-commander are exempt from the binding theory is confirmed because examples like (4.51) are still allowed. However, the following kind of example is a serious problem for any such account:

- (4.52) A. Mary<sub>i</sub>-ka ku pati-e kass-ni anim tarun salam-i  
 taysin kass-ni?  
 Mary-NOM the party-to went-Q or other person-NOM  
 instead go-Q  
 ‘Is it Mary<sub>i</sub> who went to the party or somebody else instead?’
- B. \*Ani, caki<sub>i</sub>-**nun** kasse  
 no, self-TOP went

In B of (4.52), *caki* is a topic marked subject, which thus would take up the first position of ARG-ST. Since the exemption is stated on the basis of the anaphor being in initial position, nothing in the binding theory rules out the occurrence of

anaphors in the main subject position. However the simple case like B in (4.52) show that this does not hold generally.

Lastly, a problem for defining the antecedent on the basis of the position in ARG-ST is also pointed out in Bredekamp (1996), where it causes a more serious problem in accounting for structures like ‘a picture of NP’ in English (e.g., *a picture of himself*) where the direct object of a head does not require any subject. However, the NPs in question are subject to a-binding in the same way as arguments of other heads. This case leaves the notion of a-subject rather vacuous which in turn makes a-binding in ARG-ST inappropriate.

### 4.3.3 Conclusion

A-binding is in general certainly more adequate than o-binding. It not only accounts for all the subject-binding patterns, but (as Manning and Sag claim) it also explains certain binding facts in Western Austronesian languages such as Balinese and Toba Batak and ergative languages like Inuit<sup>11</sup>.

Concerning *caki*-binding relations and other similar long-distance anaphora (e.g. *zibun*-binding), A-binding explains not only cases of subject-binding but also some non-subject binding relations such as indirect-object binding in causative constructions as we saw in the last section. However, as pointed out earlier, there are many other cases of non-subject binding which remain unaccounted for by the analysis. Further, defining a nested argument structure list as a constraint on the causative in the type hierarchy might be a necessary step because this helps make all the NPs in the sentence visible in one list to determine the correct binding relation. Although this might be plausible for the constructions like the causative which is morphologically complex as we saw in (4.47), the same intuition will be difficult to extend to the sentences embedding finite clauses, in which environment most of the long-distance binding patterns are observed. This issue will be taken up in more detail in section 5.1.3.

---

<sup>11</sup>For more detailed discussion and evidence for this claim, see Manning (1994), Manning (1996) and Manning & Sag (1998a)

Furthermore, the antecedency for *caki* in terms of a-subject is rather too restricted in accounting for examples like (4.50) above or the following example we saw in section (2.3.1):

- (4.53) John<sub>i</sub>-i Mary<sub>j</sub>-lul caki<sub>i/j</sub> cip-ulo tolye ponayssta  
John-NOM Mary-ACC self house-to return sent  
'John<sub>i</sub> returned Mary<sub>j</sub> to self<sub>i/j</sub>'s house'

The problems for these examples come from the fact that a-subject is defined as the first element in some level of ARG-ST of a predicate. However, *Mary* in (4.53) cannot satisfy this definition of a-subject as it is neither a syntactic subject nor a logical subject (this also applies to *John* in (4.50)). In spite of this, it is a legitimate antecedent. Therefore, the current statement of A-binding based on the definition of a-subject needs extensive modification to be able to predict the non-subject binding instances.

# Chapter 5

## Long-Distance Anaphora: The Core Factors

### 5.1 A New Approach

On the basis of the observations on *caki*-binding relations we made in Chapter 2, the critical review of the approaches in chapter 3 and the discussion of the problems in A-binding in the last Chapter, in this section we will provide a unified account treating all instances of *caki*-binding, that is, a treatment which is able to cover cases of both subject and non-subject binding.

In Chapter 2, we observed a variety of binding patterns of *caki* in terms of various possible antecedents in different constructions. In order to review approaches which have been proposed for long-distance anaphora in Korean and other languages, Chapter 3 has covered 4 groups of different approaches. Two of them include syntactic and non syntactic approaches where the former includes the subjecthood condition and the basic c-command requirement, while the other two groups are divided into two: discourse approaches and approaches based on the thematic hierarchy. As each of these approaches proved not to be adequate for *caki*-binding, we have also seen other accounts where the underlying principles for long-distance

anaphora include elements of both the syntactic and non-syntactic approaches. We have called these approaches the disjunctive theory and the conjunctive theory. Though these accounts considerably reduce the number of problems that remain from the syntactic or non-syntactic approaches alone, we have noted that these approaches fail to be applicable for *caki*-binding crucially because of the discourse or semantic roles adopted. In short, the crucial reason for that is, as we have seen, that *caki* does not allow a discourse antecedent. Also, the semantic part of the account relies on a rather strict thematic hierarchy which cannot adequately explain various possible antecedents of *caki*.

On the other hand, HPSG provides binding principles using the o-command relationship based on grammatical relationships between NPs. The principle on anaphors seems to account for a comprehensive set of data which are inexplicable by other configurationally defined binding principles. We also have reviewed the notion of a-command which is a modified version of o-command, which was noted to be problematic when applied to a greater variety of cross-linguistic data on anaphor-binding. In the previous chapter, we have seen that a-binding based on the notion of a-command still lacks the ability to predict the correct binding patterns of *caki* especially for the cases of non-subject binding. All the more, we have noted, but not yet considered, the problem arising from the locality of argument structure, which makes it hard to explain the binding relation outside the local domain.

This chapter will be devoted to the consideration of the following issues. Firstly, we will approach *caki*-binding from the most problematic non-subject binding cases, which leads us to provide a unified account which also treats cases of subject binding. Secondly, we will show that both syntax and thematic roles are in force in accounting for *caki*-binding and argue that the lexical argument hierarchy proposed by Kiss (1991) plays a key role in defining the necessary syntactic and thematic elements. Thirdly, we will attempt to incorporate the notion of lexical argument hierarchy into HPSG. This will require us to provide an amendment to the current representation of ARG-ST in HPSG, so that it can accommodate long-distance anaphora. Further,

we will argue that for long-distance anaphora, it is the notion of lexical argument hierarchy that should be implemented at the level of ARG-ST for the case of Korean, rather than Manning & Sag's (1998*a*) distinction between core and oblique argument and the inherent thematic hierarchy (Jackendoff 1972). Lastly, having noted the mysterious binding discrepancy between complements and adjuncts, observed across languages, we will also provide a treatment for such binding patterns at the level of ARG-ST. We will use ARG-ST as a level for binding as proposed by Manning (1996) in this section. However, the discussion of whether ARG-ST is a proper level for explaining *caki*-binding will be revisited in Chapter 6.

### 5.1.1 Revisiting Cases of Non-Subject Binding and The Thematic Hierarchy

When we reviewed various approaches to long distance anaphora, we discovered that what seems to be most problematic for each approach to explain is cases of non-subject binding. Cases of subject-binding appear to be quite straightforward from the point of view of each approach. For example, as a syntactic condition, the subjecthood condition directly acts to pick a subject as the antecedent. The notion of c-command guarantees that the subject will be in a c-commanding position, thus, a legitimate antecedent. Among non-syntactic approaches, subjects are also associated with the elements which are the most prominent either in discourse or in thematic ranking. However, we have seen that these approaches cannot equally explain the cases of non-subject binding.

In this section, we will approach the whole issue from the opposite direction. That is, we will first focus on discussing the cases of non-subject binding and then propose an explanation for such cases which is also able to account for subject-binding. In the previous sections, the most problematic cases of non-subject binding involved the following:



- (5.1) a. John<sub>i</sub>-i Mary<sub>j</sub>-lul caki<sub>i/j</sub>-uy chayk-ul ilk-hi-essta  
 John-NOM Mary-ACC self-GEN book-ACC read-make-PAST  
 ‘John<sub>i</sub> made Mary<sub>j</sub> read self<sub>i/j</sub>’s book’
- b. John<sub>i</sub>-i Mary<sub>j</sub>-ekey caki<sub>i/j</sub>-uy pang-lul chiwulako hayssta  
 John-NOM Mary-DAT self-GEN room-ACC to clean did  
 ‘John<sub>i</sub> ordered Mary<sub>j</sub> to clean self<sub>i/j</sub>’s room’

In (5.1) above, *caki* can be bound not only by the subject, but also by the direct object *Mary* in (5.1a), and dative NP *Mary* in (5.1b). Putting the subject-binding aside, the non-subject binding can be problematic, most obviously for the subject-hood condition. On the other hand, the c-commanding condition may explain the antecedency of the direct object *Mary* in (5.1a), but cannot account for the antecedency of the dative NP *Mary* in (5.1b) because it is inside a PP, and thus does not c-command *caki*. The approach using the thematic hierarchy is not properly applicable to these two examples because *caki* is not in the argument structure of the higher verb, thus we cannot predict the binding relations of *caki* in terms of the hierarchy. Within the A-binding analysis, at least (5.1a) can be explained as a case of a causative construction by using a nested argument structure. As for (5.1b) it could possibly be analysed as a causative construction because the main verb *hayssta* which, for convenience, we glossed as *did* but translated as *ordered*, is the verb used for the phrasal causative construction illustrated below:

- (5.2) Bill<sub>i</sub>-i Mary<sub>j</sub>-ekey caki<sub>i/j</sub>-uy cha-lul mol-ke hayssta  
 Bill-NOM Mary-DAT self-GEN car-ACC drive-COMP told  
 ‘Bill<sub>i</sub> made Mary<sub>j</sub> drive caki<sub>i/j</sub>’s car’

Then if this is the case, for (5.1b), a-binding allows *Mary* to be an a-subject of the lower predicate, thus binding *caki*. However, this explanation does not extend to examples such as (5.3a) and (5.3b):

- (5.3) a. John<sub>i</sub>-i Mary<sub>j</sub>-lul caki<sub>i/j</sub> cip-ulo tolye ponayssta  
 John-NOM Mary-ACC self house-to return sent  
 ‘John<sub>i</sub> returned Mary<sub>j</sub> to self<sub>i/j</sub>’s house’

- b. Mary<sub>i</sub>-ka aki<sub>j</sub>-lul caki<sub>i/j</sub> chimday-e nwuphiessta  
 Mary-NOM baby-ACC self bed-in put  
 ‘Mary<sub>i</sub> put (her) baby<sub>j</sub> in self<sub>i/j</sub>’s bed’

The two examples above demonstrate both subject and non-subject binding, but they are not causative constructions. Therefore the non-subject binder in each example cannot act as an a-subject. However, it still binds *caki*. At this point, the binding behaviour seems very confusing and complicated.

The following example exhibits a similar syntactic construction to those in (5.3) except that unlike (5.3) the direct object does not bind *caki*.

- (5.4) a. Mary<sub>i</sub>-ka John<sub>j</sub>-ul caki<sub>i/\*j</sub>-uy kapang-ulo ttalyessta  
 Mary-NOM John-ACC self-GEN bag-with hit  
 ‘Mary<sub>i</sub> hit John<sub>j</sub> with self<sub>i/\*j</sub>’s bag’  
 b. John<sub>i</sub>-i Bill<sub>j</sub>-ul caki<sub>i/\*j</sub>-uy cip-ese mannassta  
 John-NOM Bill-ACC self-GEN house met  
 ‘John<sub>i</sub> met Bill<sub>j</sub> at self<sub>i/\*j</sub>’s house’

At this point, it seems that the most plausible source for the explanation may lie in the thematic relations. However, we have seen that an explanation based on thematic roles is not available in these cases since *caki* is not in an argument position, but it is embedded within an argument. However, there may be a counterargument: For example, Hellan (1988) argues that the restrictions imposed by the thematic hierarchy hold not only of coarguments of a single predicate, but also of more deeply embedded reflexive elements. Consider the following example in Norwegian:

- (5.5) Barnet ble fratatt sine foreldre  
 the child was from-taken its parent  
 ‘The child<sub>i</sub> (MALEFACTIVE/SOURCE) was deprived of self’s<sub>i</sub> parents (THEME)’  
 ‘\*The child<sub>i</sub> (THEME) was taken away from self’s<sub>i</sub> parents (MALEFACTIVE/SOURCE)’

In the example above, the antecedent *barnet* ('the child') is required to be more prominent in the thematic hierarchy than *sine foreldre* ('its parents') in order to bind the anaphor *sine* which is embedded in the noun phrase *sine foreldre*. In explaining the binding relation of the genitive form of anaphor in Norwegian, Hellan argues that examples like (5.5) strongly suggest that the requirement of superiority (in the thematic hierarchy) hold between the antecedent of an anaphor and either (i) the anaphor, if the anaphor is a coargument of the antecedent; or (ii) the coargument of the antecedent in which an anaphor is embedded. With Hellan, we will also assume that in this way the genitive *caki* in (5.3) and (5.4) will be visible to its antecedent in order for the thematic hierarchy to predict the binding relation.

Although this proposal brings reflexives which are not coarguments of their antecedents within the scope of the thematic hierarchy, it still leaves out the cases where the antecedent is not a coargument of the reflexive, but rather embedded within a coargument. This is obvious in examples such as the following:

- (5.6)     John<sub>i</sub>-uy ayin<sub>j</sub>-i     caki<sub>i/j</sub>-lul cohahanata  
           John-GEN lover-NOM self-ACC likes  
           'John<sub>i</sub>'s lover<sub>j</sub> likes self<sub>i/j</sub>'

In (5.6), *John's lover* is an argument of *likes*, and receives its thematic role from that predicate. *John*, however, is not an argument of *likes*. It is an argument of *lover* and receives its thematic role (possessor) from *lover*. Both NPs can antecede *caki*. From the point of view of the thematic hierarchy, this is perhaps just as it should be. Given that the thematic hierarchy organises the arguments of a single predicate, it is expected that it would have nothing to say about arguments of different predicates. However, this conception of the function of the thematic hierarchy predicts that in the following sentence all NPs should be possible antecedents of *caki*, contrary to fact:

- (5.7) John<sub>i</sub>-uy emma<sub>j</sub>-ka Bill<sub>k</sub>-lopwute [caki<sub>\*i/j/k</sub>-ka mossaykiessta]-ko tulessta  
 John-GEN mother-NOM Bill-from self-NOM is ugly-COMP heard  
 ‘John<sub>i</sub>’s mother<sub>j</sub> heard from Bill<sub>k</sub> that self<sub>\*i/j/k</sub> is ugly’

The reason for the unavailability of *John* as the antecedent of *caki* is unclear. It shows however that the above conception of the function of the thematic hierarchy is clearly mistaken. We would like to propose that the converse of Hellan’s (1988) and Dalrymple’s (1993) proposal concerning the anaphor also holds. That is, that the thematic inferiority requirement should hold between the anaphor and the antecedent or the coargument of the anaphor within which the antecedent is contained.

In section 3.2.2, we have reviewed two thematic hierarchies proposed by Jackendoff (1972) and Momoi (1985), repeated respectively below:

- (5.8) a. Jackendoff (1972)  
 Agent > Location, Source, Goal > Theme  
 b. Momoi (1985)  
 Experiencer > Agent > Objective (Patient) > Others (Goal, Source, etc.)

If we employ each thematic hierarchy above to define the command relation in (5.3), Jackendoff’s hierarchy in (5.8a) makes the wrong predictions for (5.3a) and (5.3b) where *caki* is included in the PP (Location). In Jackendoff’s thematic hierarchy, the Location outranks the Theme, but the two examples in (5.3) show that *caki* is bound by the direct object (Theme). On the other hand, Momoi’s thematic hierarchy in (5.8b) explains those binding relations because in his thematic hierarchy, the Theme outranks the Location (which is supposed to be in the same rank as others (Goal, Source, etc.)). However, this fails to account for the example in (5.4b) where *caki* in the PP (Location) is not bound by the direct object (Theme) which means that the Location should outrank the Theme. (5.4a) poses the same problem, given that the Instrumental is in the same rank as the Location.

Note, however, that there are some clear distinctions in binding relations between (5.3) and (5.4). Looking closely at the PPs in these examples, those in the former are complement (selected) PPs, whereas those in the latter are adjunct PPs. The binding relations in this respect imply that *caki* in complement PPs can be bound by the direct object (Theme), while in the adjunct PPs, *caki* is not allowed to be bound by the direct object (Theme). If we suppose that the adjunct, in fact, cannot be involved in the explanation based on thematic hierarchy, then Momoi's (1985) thematic hierarchy seems to predict the right binding relation at least in the examples (5.3) which involve only complement PPs. However, the following examples strongly suggest that this assumption takes the wrong direction. Firstly, consider the following:

- (5.9) John<sub>i</sub>-i Bill<sub>j</sub>-ekey [Mary<sub>k</sub>-ka caki<sub>i/\*j/k</sub>-lul choahanta]-ko malhayssta  
 John-NOM Bill-DAT Mary-NOM self-ACC like-COMPL said  
 'John<sub>i</sub> said to Bill<sub>j</sub> that Mary<sub>k</sub> likes self<sub>i/\*j/k</sub>'

In this example, *caki* is only bound by *John* (Agent) in the higher clause, but not the dative PP *Bill* (Goal). This still suggests that Momoi's thematic hierarchy correctly predict the binding relationship and, in any case, the adjunct *to Bill* is not taken into consideration in determining the binding relationship in the thematic hierarchy. However, the following sentence shows exactly the opposite:

- (5.10) Mary<sub>i</sub>-ka John<sub>j</sub>-uropwute [**caki**<sub>i/j</sub>-ka Bill<sub>k</sub>-lul hyemohanta]-ko tulessta  
 Mary-NOM John-from self-NOM Bill-ACC hate-COMPL heard  
 'Mary<sub>i</sub> heard from John<sub>j</sub> that **self**<sub>i/j</sub> hates Bill<sub>k</sub>'

Firstly, other than the Agent antecedent *Mary*, *John* (Source) which is at the bottom rank in the hierarchy (together with the Goal), antecedes *caki* embedded in the Theme clause. Secondly, this is so in spite of the fact that *John* is in the adjunct PP which is supposed to be opaque to the thematic hierarchy. This enables us to point out two crucial facts from our observations of the application of the thematic role hierarchy to non-subject binding cases.

Firstly, the suggested thematic hierarchies (Jackendoff 1972, Momoi 1985) that treat roles like Goal and Source as equally ranked are problematic for the examples in (5.9) and (5.10). Other hierarchies such as the one suggested by Bresnan & Kanerva (1989) are somewhat different in that the thematic roles involved are more finely categorised as shown in (5.11) below. However, the problem is that this still does not explain the different binding patterns between the Goal and the Source in relation to *caki* (Theme) as in (5.9) and (5.10).

- (5.11) agent > beneficiary > recipient/experiencer > instrument > theme/patient  
> location

Therefore, even though the notion of thematic roles seems to play a key role in explaining cases of non-subject binding, the thematic hierarchy does not seem to be flexible enough to be a source for explaining the distinct binding patterns among obliques such as Source, Goal, Location etc.

The second noteworthy fact comes from the point that we still cannot explain the seemingly different non-subject binding behaviours between arguments and adjuncts as in (5.3) and (5.4), where *caki* is embedded in the PPs either as a complement or an adjunct. As there are clear differences in the binding patterns with respect to the distinction between complement and adjuncts, such a distinction should be considered another dimension for *caki*-binding.

Before we move on to the specifics of our proposal, it may be useful to summarise the empirical and theoretical desiderata that any theory of *caki*-binding should minimally meet. Those are summarised in (5.12):

- (5.12) a. A unified account of both subject and non-subject binding.  
b. An explanation for the different behaviour of *caki* in complements and adjuncts.  
c. An explanation of *caki*-binding must be couched in essentially formal sentence grammar-level terms.

- d. Finally, in order to achieve a maximum explanatory effect, *caki* binding relations should be established and determined within a single representation.

We now turn to a further aspect of our proposal, namely, the thematically motivated lexical argument hierarchy.

### 5.1.2 The Thematically Motivated Lexical Argument Hierarchy

One of the problems in taking advantage of thematic roles in explaining *caki*-binding was that the approaches based on the thematic hierarchy generally are restrained by too rigid a relationship between thematic roles of arguments independent of heads. The thematic hierarchy says that one role is more prominent than another, no matter what kind of head or what kind of construction it is involved in. However, we have witnessed that the thematic hierarchy cannot explain cases like (5.10) and (5.9) where the binding relations demonstrate that with the head, *heard*, the Source is more prominent than the Theme (in 5.10); whereas with the head, *told*, the Theme is more prominent than the Goal (in 5.9). This shows that we need a level of hierarchy which assigns different ranks among thematic roles, depending on what kind of head is involved.

Furthermore, an invariant thematic hierarchy irrespective of the head cannot predict the correct binding relation of *caki* in passive constructions. Consider the following pair of examples:

- (5.13) a. John<sub>i</sub>-i Mary<sub>j</sub>-ekey caki<sub>i/\*j</sub>-uy chayk-ul cwu-ess-ta  
 John-NOM Mary-DAT self book-ACC give-PAST-DC  
 'John<sub>i</sub> gave Mary<sub>j</sub> self<sub>i/\*j</sub>'s book'
- b. Mary<sub>j</sub>-ka John<sub>i</sub>-ekeyse caki<sub>j/i</sub>-uy chayk-ul cwu-eci-ess-ta  
 Mary-NOM John-FROM self-GEN book-ACC give-PASS-PAST-DC  
 'Mary<sub>j</sub> was given self<sub>j/i</sub>'s book by John<sub>i</sub>'



The active sentence in (5.13a) shows that the genitive *caki* in the Theme is only bound by the Agent *John*. This is predicted by the thematic hierarchy as the Agent outranks the Theme. Given the assumption that passivization does not alter the ranking among arguments in the thematic hierarchy, we would expect the same binding relation to be maintained in the passivised counterpart of (5.13a). However, this does not hold in reality. As seen in (5.13b), *caki* is bound not only by the Agent *John*, but also, problematically, by the Goal *Mary*. This leads us to assume that it is not a rigid thematic hierarchy that we can rely on in predicting *caki*-binding, but rather some level where the subject in passivization is promoted to be the most prominent element of the head in question. In other words, the desirable level we are seeking is one where the Agent subject is the most prominent element for the active head (e.g. *gave* in (5.13b)) and also where the passivised subject, whatever thematic role it may be, is the most prominent for the passive head (e.g. *was (given)* in (5.13b)). This strongly supports the argument that an inherent thematic hierarchy which is defined irrespective of the individual lexical head is the wrong place to look for explanations of *caki*-binding.

A similar problem has been pointed out by Kiss (1991) in her account of anaphoric binding relations in Hungarian. One of the components in her proposal involves a lexical argument hierarchy which is thematically motivated, but different from the inherent thematic hierarchy in Jackendoff (1972) and Momoi (1985). She attempts to provide a more flexible thematic hierarchy without affecting the universality of the thematic hierarchy by incorporating the notion of a 'primacy condition'. For example, Kiss (1991) proposes that for flat-structured languages like Hungarian, reflexivization will be best explained in terms of the thematic hierarchy. The hierarchy that she adopts, however, is not based on an absolute, inherent hierarchy of  $\theta$ -roles, independent of lexical items, but is the hierarchy of the  $\theta$ -roles associated with a given head in the lexicon. That is, for example, for the verb *give*, the source is the most prominent argument and the goal is the least prominent one, while in the case of the verb, *get*, it is the other way around.



Furthermore, in determining the order of prominence among thematic roles dependent on a lexical item, syntactic realization also contributes to deciding the order of the arguments. Kiss (1991) proposes that the most prominent argument of a verb, denoting the most active participant of the action or state, is the argument marked for external Case assignment (in English, and for nominative assignment in Hungarian). It is more typically represented by the agent or the experiencer, if the  $\theta$ -grid of the verb has any. The second most prominent argument is the direct argument of the verb: most typically a Theme or accusative marked argument. A locative or temporal argument, denoting a completely passive element of the action or state, typically occupies the lowest position in the argument hierarchy. That is, the first elements and the last elements of the hierarchy are fixed. The first will be the most distinguished, as the external argument<sup>1</sup>, followed by the internal argument which is directly case assigned. Locatives and temporals will be ordered last in the hierarchy. In between the first and the last elements, the Source, the Goal and the Theme will be ordered according to the prominence in thematic roles determined by the lexical head.

We adopt the basic notion of Kiss's (1991) lexical argument hierarchy except for one point about directly case assigned arguments, which Kiss assumes are mostly Theme and which should be the second most prominent in the hierarchy. We will see the empirical reason for this from Korean later in this section. However, there is a more general reason why Theme should not always be considered a second most prominent element. Whereas the Agent is typically realised as a DP external argument (except for the passive constructions), the Theme is realised in various ways, for example, DP, PP, or even as a complement clause. Further, it is not only the Theme which

---

<sup>1</sup>It should be noted that the term external argument here should not be confused with that in Grimshaw (1990), Williams (1981), di Sciullo & Williams (1987) and Marantz (1984). For example, Grimshaw (1990) proposes the structured internal organisation of argument structure, where it represents an external/internal argument distinction. In such work, the external argument is more of a lexical semantic notion in that it is defined as an argument that is maximally prominent in both aspectual and thematic prominence. This structured argument structure makes it possible to explicate some of the core properties of external arguments in order to understand how theta-marking works in cases of multiple theta-marking domains, and ultimately to explain the grammatical behaviour of various verb classes. Compared with this, on the other hand, the external argument we have in mind in terms of the lexical argument hierarchy is associated with the syntactic subject.

is typically realised as the direct argument, but the Patient also often occurs in the direct argument position, which Kiss does not say anything about. If the lexical item takes the Patient, though it may be realised as the directly case-marked internal argument, it should be considered to be the least active participant of the action or state. Therefore, being directly case-assigned does not necessarily relate to the second more prominent in the hierarchy. From this stand, we will assume that it is only the external argument which is invariably taken as the most active participant, and thus, should be at the top of the hierarchy. This is a general principle to operate on our lexical argument hierarchy. On the other hand, the direct argument of a head should be treated like the rest of the dependents, which means that it will be governed by a lexically specific (idiosyncratic) constraint.

Having noted the difference between Kiss's (1991) and other thematic hierarchies, adopting Kiss's (1991) approach helps in explaining *caki*-binding. Firstly, in (5.13), we have seen that thematic hierarchies cannot account for *caki* being bound by the passive subject. However, such binding relations can be easily predicted by Kiss's (1991) theory. The lexical argument hierarchy of the active sentence in (5.13a) can be represented in the following way:

(5.14) John<sub>i</sub> (External Arg./Agent) > *caki*<sub>i/\*j</sub>'s book (Theme) > Mary<sub>j</sub> (Goal)

The lexical argument hierarchy in the above shows that *John*, the Agent/external argument will be ordered first in the hierarchy. Given the head *gave*, the Goal *Mary* is the least prominent element, thus placed at the bottom of the hierarchy. Then the theme *caki's book* naturally goes in the middle of the hierarchy. Note that the dative in Korean is realised by the postposition '*ekey*' (to), thus, it is treated as a PP rather than a direct argument of the head as in English examples like (5.15)<sup>2</sup>:

(5.15) Mary gave *him* the book

---

<sup>2</sup>This is of course not to say that the dative can only be realised as the direct argument of the head in English. It can also be a PP as in *Mary gave a book to John*.

The hierarchy predicts the binding relation of *caki* correctly, with *John* being the only antecedent. On the other hand, the lexical argument hierarchy for the passive counterpart shown below is not the same as (5.14), in contrast to what is assumed in the standard thematic hierarchy. Consider the following:

(5.16) Mary<sub>j</sub> (External Arg.) > John<sub>i</sub> (Agent) > caki<sub>j/i</sub>'s book (Theme)

The hierarchy in the above begins with *Mary* which is the external argument (nominative marked). This is followed by *John* which is the Agent (which also can be seen as the Source), the most prominent role, but as the first position is occupied by the external argument, it is ordered as the next most prominent argument. The last element comes with *caki's book*. This order in the hierarchy also predicts the correct binding relation with *caki* being able to be bound by not only the Agent role argument, but also by the passive subject. Therefore, the lexical argument hierarchy circumvents the problem raised by the general thematic hierarchy in passivization.

The advantages of the lexical argument hierarchy can also be seen in explaining a pair of examples where the Goal and the Source exhibit different binding relations. We will repeat the examples in (5.10) and (5.9) as below, each followed by its own hierarchy:

- (5.17) a. Mary<sub>i</sub>-ka John<sub>j</sub>-uropwute [**caki**<sub>i/j</sub>-ka Bill<sub>k</sub>-lul hyemohanta]-ko tulessta  
 Mary-NOM John-from self-NOM Bill-ACC hate-COMPL heard  
 'Mary<sub>i</sub> heard from John<sub>j</sub> that self<sub>i/j</sub> hates Bill<sub>k</sub>'  
 Mary<sub>i</sub> (External Arg./Agent) > John<sub>j</sub> (Source) >  
 S:(caki<sub>i/j</sub>, Bill)(Theme)
- b. John<sub>i</sub>-i Bill<sub>j</sub>-ekey [caki<sub>i/\*j</sub>-ka Mary-lul choahanta]-ko malhayssta  
 John-NOM Bill-DAT Mary-NOM self-ACC like-COMPL said  
 'John<sub>i</sub> said to Bill<sub>j</sub> that self<sub>i/\*j</sub> likes Mary'  
 John (External Arg./Agent) > S:(caki<sub>i/\*j</sub>, Mary)(Theme) > Bill<sub>j</sub>  
 (Goal)

In (5.17a), *Mary*, which is the Agent as well as the external argument, comes as the first element in terms of the general principle on the external argument. On the other hand, in terms of the idiosyncratic information about a head, for *heard*, the second element is *John*, the Source, which is the most prominent element after the external argument, and the Theme clause is the last. This order of rank correctly predicts that *caki* embedded in the Theme clause is bound either by *Mary* or by *John*. The binding relation in (5.17b) is different, although the constructions are similar. *John*, like *Mary* in (5.17a), takes the first rank in the hierarchy. Further, for *told*, *Bill*, the Goal, is the least prominent, thus comes last in the hierarchy. This again correctly predicts that *caki* embedded in the Theme clause, the second element in the hierarchy, is only bound by *John*. Such ordering applies to sentences which do not involve a theme clause but rather involve a Theme noun phrase as shown below:

- (5.18) a. John<sub>i</sub>-i Bill<sub>j</sub>-ekey caki<sub>i/\*j</sub>-uy sacin-ul poyecwuessta  
 John-NOM Bill-DAT self-GEN picture-ACC showed  
 ‘John<sub>i</sub> showed a picture of self<sub>i/\*j</sub> to Bill<sub>j</sub>’
- b. John<sub>i</sub>-i Bill<sub>j</sub>-ekey caki<sub>i/\*j</sub>-uy erinsicel-ul iyakihayssta  
 John-NOM Bill-DAT self-GEN childhood-ACC told  
 ‘John<sub>i</sub> told about self<sub>i/\*j</sub>’s childhood to Bill<sub>j</sub>’

As the data we have seen so far does not actually involve any directly case-assigned argument, let us now check whether adopting the lexical hierarchy predicts the correct binding relations in the constructions where a direct internal argument is involved. The following sentences are repeated from the last section:

- (5.19) a. John<sub>i</sub>-i Mary<sub>j</sub>-lul caki<sub>i/j</sub> cip-ulo tolye-ponayssta  
 John-NOM Mary-ACC self house-to return-sent  
 ‘John<sub>i</sub> returned Mary<sub>j</sub> to self<sub>i/j</sub>’s house’
- b. Mary<sub>i</sub>-ka aki<sub>j</sub>-lul caki<sub>i/j</sub> chimday-e nwuphiessta  
 Mary-NOM baby-ACC self bed-in put  
 ‘Mary<sub>i</sub> put (her) baby<sub>j</sub> in self<sub>i/j</sub>’s bed’

Example (5.19a) above has the direct internal argument *Mary* with the Location *to caki's house*, and the lexical argument hierarchy is shown as below:

- (5.20) John<sub>i</sub> (External Arg./Agent) > Mary<sub>j</sub> (Theme) > caki<sub>i/j</sub>'s house (Location)

In the above, as the Location is suggested as the least prominent role no matter what the lexical item is (Kiss 1991), the Location is placed at the bottom of the hierarchy and the rest of the arguments go as expected: the external argument first and the Theme in the middle. The resulting hierarchy correctly predicts the binding relations of *caki*; *caki* in the Location at the bottom of the hierarchy is bound either by the external argument *John* or by the Theme *Mary*. This is the same hierarchy that applies in (5.19b) with the correct prediction for *caki*-binding.

So far, the lexical argument hierarchy seems to provide an adequate basis to explain not only subject binding but also non-subject binding patterns of *caki*, which have been most problematic for the approaches we have reviewed. However, there is a group of data we have considered but not yet explained: the binding relations in the constructions which involve adjunct phrases. We repeat those we have seen in the last section as (5.21) below:

- (5.21) a. Mary<sub>i</sub>-ka John<sub>j</sub>-ul [caki<sub>i/\*j</sub>-uy kapang-ulo] ttalyessta  
 Mary-NOM John-ACC self-ACC bag-with hit  
 'Mary<sub>i</sub> hit John<sub>j</sub> with self<sub>i/\*j</sub>'s bag'
- b. John<sub>i</sub>-i Bill<sub>j</sub>-ul caki<sub>i/\*j</sub> cip-ese mannassta  
 John-NOM Bill-ACC self house met  
 'John<sub>i</sub> met Bill<sub>j</sub> at self<sub>i/\*j</sub>'s house'

Both sentences in the above involve direct arguments (Theme) and adjuncts (Instrumental and Location, respectively). The lexical argument hierarchy of each sentence is illustrated below:

- (5.22) a. Mary<sub>i</sub>(External Argument/Agent) > John<sub>j</sub>(Theme) > caki<sub>i/\*j</sub>'s bag  
 (Instrumental)

- b. John<sub>i</sub>(External Argument/Agent) > Bill<sub>j</sub>(Theme) > caki<sub>i/\*j</sub>'s house  
(Location)

Thematic roles like Instrumental and Location are typically the most passive roles in the event or state, thus they take the bottom position in the hierarchy as shown in the above. The Themes in (5.22) should outrank both the Instrumental and Location thematic roles without question. Then, according to the ordering in the hierarchy where *caki* is embedded in the bottom element, it should be able to be bound either by the Agent or the Theme. However, this is not the case: the Theme arguments, though they outrank the obliques, can not antecede *caki* embedded inside the obliques.

Overall, even though the lexical argument hierarchy plays a key role in providing a unified account both for subject and non-subject binding of *caki*, it does not resolve the problems coming from the mysterious discrepancy of *caki*-binding patterns between arguments and adjuncts. For now, we will leave this section with this problem unresolved, and return to it below. In the next section, we will provide a formal account of *caki*-binding within HPSG, adopting the notion of lexical argument hierarchy and we will show that our formal analysis within HPSG will provide the answer to the mysterious binding patterns with respect to adjuncts.

### 5.1.3 Formal Treatment of *Caki*-Binding Within HPSG

In this section, we will attempt to provide a unified account of *caki*-binding including both subject and non-subject binding, which can be reduced to one formal level of representation of argument structure incorporated into HPSG. The main subsections of this section will explore in more detail the arrangement of dependents of a head in argument structure in HPSG proposed by Manning & Sag (1998a) and show that long-distance anaphora are best explained when we adopt the notion of lexical argument hierarchy in arranging the arguments within argument structure in HPSG. However, before we directly move on to this discussion, we would like to address the problems of the current representation of the argument structure within



HPSG in explaining long-distance binding. Recall also that we will use ARG-ST as a level for binding in this section, but we will return to the issue whether ARG-ST can be a proper representation for our analysis of *caki*-binding in Chapter 6.

### *Parametrising the locality of ARG-ST*

When we reviewed the account of anaphor-binding in HPSG (Manning 1996), we pointed out in section 4.3.3 that determining anaphoric binding relations at the level of argument structure is problematic for long-distance anaphora. The problem in question focuses on the representation of argument structure in HPSG, which cannot properly account for the binding relation of *caki* and its antecedent outside the local domain, such as in cases where the antecedent binds *caki* outside a finite complement clause. The same problem can also be detected in the binding of genitive *caki* where the property of *caki* as an anaphor does not have any means to surface in the argument structure. We will attempt to amend this problem in this section so that long-distance anaphora like *caki*-binding can be properly accommodated within the level of argument structure which we believe to be the right place to determine the binding patterns of *caki*. Then we will provide a formal account of *caki*-binding based on the amended representation of argument structure.

The relevant data which are problematic for the representation of argument structure in HPSG, as we just noted above, include cases like the following:

- (5.23) John<sub>i</sub>-i [Mary<sub>j</sub>-ka [Bill<sub>k</sub>-i caki<sub>i/j/k</sub>-lul choahanta]-ko sayngkakhanta]-ko  
 John-NOM Mary-NOM Bill-NOM self-ACC like-COMP thinks-COMP  
 malhayssta  
 said  
 ‘John<sub>i</sub> said that Mary<sub>j</sub> thinks that Bill<sub>k</sub> likes self<sub>i/j/k</sub>’

The example above illustrates that the long-distance anaphor *caki* is bound not only by the local subject *Bill*, but also by the higher subjects *Mary* and *John* outside the local domain. The standard representation of the argument structure of (5.23) would be as follows:

- (5.24) a. *said* :ARG-ST<John<sub>i</sub>, S>  
 b. *thinks* :ARG-ST< Mary<sub>j</sub>, S>  
 c. *likes* :ARG-ST< Bill<sub>k</sub>, caki<sub>i/j/k</sub> >

(5.24) illustrates the ARG-ST of the three lexical heads in the sentence (5.23): *said*, *thinks*, *likes*. The argument structure should only include arguments selected by the head, and thus it is realised as a lexical sign in HPSG (as opposed to a phrasal sign). Hence, in (5.24a) and (5.24b), the head selects one subject noun and a complement clause, but the arguments inside the complement clause are not visible. If they were, that would mean that the verbal head selects the argument inside the complement clause. This is precisely based on the notion of locality of argument structure. However, this is a direct problem for LDA because HPSG assumes that the binding relation is determined at the level of argument structure. For example, in the argument structure of *said* in (5.24a), *John* is supposed to be a suitable antecedent, but the bindee *caki* is not visible, or more precisely, it is an irrelevant element in that level of argument structure. The same applies to the antecedency of *Mary* (5.24b). Further, in (5.24c), *caki* is marked as being possibly bound by both *John* and *Mary*, which are not within its argument structure. Given Manning's (1996) principle that an anaphor has to be bound by a-subject, the only a-subject available in the argument structure is *Bill* and nothing else. Therefore, it seems either that the assumption of determining the anaphoric binding relation in the argument structure is wrong, or that the representation of argument structure needs reformulating.

The explanation of genitive *caki*-binding suffers from the same problem. We have seen many examples of genitive *caki* binding, one of which is given below:

- (5.25) John<sub>i</sub>-i    caki<sub>i</sub>-uy ayin-eytayhay malhayssta  
 John-NOM self-GEN lover-about    talked  
 'John<sub>i</sub> talked about self<sub>i</sub>'s lover'

Firstly, we have seen in section 4.3.1 that HPSG assumes that, given that the head preposition makes no contribution to the CONTENT of the PP, the latter is structure-shared with the CONTENT of the prepositional object, so that anaphoric



properties can be visible at PPs at the level of argument structure, which makes the PPs capable of participating in the binding relations. However, in (5.25) above, the complement NP of *about* is *ayin* (lover) but not *caki*, which is the complement of the head noun *ayin*. The argument structure of (5.25) is given below:

- (5.26) a. *talked* :ARG-ST< John<sub>i</sub>, PP[*about*] >  
 b. *about* :ARG-ST< lover >  
 c. *lover* :ARG-ST< caki<sub>i</sub> >

The PP in (5.26a) cannot be anaphoric because *caki* cannot be visible in the ARG-ST, thus *John* cannot antecede anything. On the other hand, *caki* in (5.26b) is stranded with no antecedent available in its argument structure. Therefore, we can see that genitive *caki* does not have any means of showing up in the same level of representation as its antecedent.

One way to tackle this problem would be to assume that for languages which exhibit long-distance anaphora, argument structure can be a HEAD feature that is inherited via the HFP from the lexical head of the phrase. If we make the argument structure visible at the phrasal level, the argument structures in (5.23) and (5.25) would now look respectively like the following:

- (5.27) a. *said* : ARG-ST< John<sub>i</sub>, S: < Mary<sub>j</sub>, S:< Bill<sub>k</sub>, caki<sub>i/j/k</sub> > > >  
 b. *talked* :ARG-ST< John<sub>i</sub>, PP:< NP:< caki<sub>i</sub> > > >

The effect of making the argument structure visible at the phrasal level as in (5.27) above will make the higher arguments relevant for *caki*-binding, that is, *caki* is visible to its possible antecedents. However, given the standard assumption of locality of argument structure, representing the argument structure as in (5.27) risks introducing an excessive degree of non-locality into complement selection. This means that (5.27) can be interpreted such that the head *said* selects all the arguments (*Mary*, *Bill* and *caki*) in the (potentially infinite) sequence of lower argument structure of other heads (*thinks* and *likes* in this case).

However, representations of this type are not unfamiliar. We have already seen in section 4.3.2 that Manning & Sag (1998a) and Manning *et al.* (1999) propose a similar representation, the so-called nested argument-structure. They argue that constructions such as causatives and passives in Japanese, Western Austronesian languages, Turkish and Inuit, where those constructions are realised as morphologically complex predicates (illustrated below), strongly suggest a nested argument structure configuration.

(5.28) a. Causative construction in Japanese

Yumiko<sub>i</sub> ga Ziroo<sub>j</sub>-ni sono hon<sub>k</sub>-o yom-ase-ta

Yumiko NOM Ziroo-DAT that book-acc read-CAUS-PAST

‘Yumiko<sub>i</sub> made Ziroo<sub>j</sub> read that<sub>k</sub> book’

(Manning *et al.* 1999, p.2, ex 2 Modified)

b. Passive construction in Inuit

Amiit Jaaku-mit qimmi-nut niri-tsaali-niqar-put

skin.PL.ABS Jaaku-ABL dog-PL.TERM eat-prevent-PASS-IND.3PL

‘The skins<sub>i</sub> were prevented by Jaaku<sub>j</sub> from the dogs<sub>k</sub> eating *t<sub>i</sub>*’

(Manning & Sag 1998a, p.29, ex 55 Modified)

In (5.28a), the causative verb cluster *yom-ase-ta* appears where the Japanese causative affix *(-)ase* is attached to a verb stem *yom* (to read). In the same way, Inuit verbs for passives are realised by attaching the passive affix *niqar* to the verb stem *niri* (to eat) and *tsaali* (to prevent). The corresponding argument structure for each sentence is proposed by Manning & Sag (1998a) and Manning *et al.* (1999) as follows:

(5.29) a. ARG-ST < Yumiko<sub>i</sub>, Ziroo<sub>j</sub>, < PRO<sub>j</sub>, book<sub>k</sub> > >

b. ARG-ST < skins<sub>i</sub>, < Jaaku<sub>j</sub>, PRO<sub>i</sub>, < dog<sub>k</sub>, PRO<sub>i</sub> > > >

The mechanism behind these nested argument structures is an independent type hierarchy which defines a single derivation type *caus(ative)-drv*, or *passive-drv*. The constraints on such types are posited to be the following:

(5.30)

$$\begin{array}{c}
 \text{caus-drv:} \\
 \left[ \begin{array}{l}
 \text{RESULT} \\
 \text{SOURCE}
 \end{array} \right.
 \left[ \begin{array}{l}
 \text{caus-}lxm-1 \\
 \text{PHON} \quad F_{(s)ase}(\boxed{1}) \\
 \text{ARG-ST} \quad \langle NP_i, NP_j, \boxed{4} \langle PRO_j, \dots \rangle \rangle \\
 \\
 \text{CONT} \\
 \left[ \begin{array}{l}
 \text{cause-rel} \\
 \text{ACTOR} \quad i \\
 \text{UNDERGOER} \quad j \\
 \text{EFFECT} \quad \boxed{3}
 \end{array} \right] \\
 \\
 \text{v-}lxm \\
 \text{PHON} \quad \boxed{1} \\
 \text{ARG-ST} \quad \boxed{4} \\
 \text{CONT} \quad \boxed{3}
 \end{array} \right.
 \end{array}
 \right]$$

(5.31)

$$\begin{array}{c}
 \text{passive-drv:} \\
 \left[ \begin{array}{l}
 \text{RESULT} \\
 \text{SOURCE}
 \end{array} \right.
 \left[ \begin{array}{l}
 \text{pass-v-}lxm \\
 \text{ARG-ST} \quad \langle \boxed{2}_j, \langle \boxed{1}_i, PRO_j \rangle \oplus \boxed{3} \rangle \\
 \text{CONT} \quad \boxed{4} \\
 \\
 \text{trans-v-stem} \\
 \text{ARG-ST} \quad \langle \boxed{1}, \boxed{2} \rangle \oplus \boxed{3} \\
 \text{CONT} \quad \boxed{4}
 \end{array} \right.
 \end{array}
 \right]$$

These derivational types are to be read as saying that the basic and other derived lexemes of the SOURCE type license additional lexemes of the RESULTS type. In the case of passive in (5.31) for example, if the source is the regular active verb, this type will license a passive lexeme with a nested argument structure such as in (5.29b). On the other hand, the causative derivational type in (5.30) licenses the construction of causative lexemes, based on other lexemes made available by the lexicon. This type will thus license a causative verb with a nested argument structure list as in (5.29a). In Manning & Sag (1998a) and Manning *et al.* (1999), it is argued that the anaphoric binding relations in many languages (e.g., Japanese, West Austronesian languages, etc) support this kind of representation of passives and causatives where

logical and surface subject bind the anaphors, which prompts the proposal of the anaphor principle based on the notion of a-subjects (see section 4.3.2).

We have seen in section 4.3.2 that Korean causatives and passives also involve a complex predicate and that the representation of a nested argument structure seems to explain the binding relations of *caki*. We want to argue that the nested argument structure also carries the same risk of introducing an excessive degree of non-locality into complement selection, as discussed in assuming the nested argument structure for the binding relation of *caki* embedded in a complement clause. That is, in (5.29a), it could be assumed that the higher head, the causative lexeme, can select the complement *hon* (book) of the lower head, the lexical stem *yom* (to read). This point may be defended by the fact that morphological causatives are monoclausal. However, Manning & Sag (1998a) themselves also argue that such constructions are *underlyingly* biclausal, which also supports the argument that the head selects the complement of the lower head. In fact, they show that in causatives in Inuit, where they also assume the nested argument structure configuration, the causative affix agrees with the lower object as below:

- (5.32) Aani-p miiqqa-t Juuna-mut paari-tip-pai  
Aani-EG child-PL.ABS Juuna-TERM look.after-CAUS-3SG.3PL  
'Aani had Juuna look after the children'

(Manning & Sag 1998a, p.21, ex.38a)

Here the causative affix *tip* agrees with the lower object *miiqqat*(children). And it seems indeed that there is a selectional relation between the higher head and the complement of the lower head. This however cannot hold at least for Japanese or Korean causatives: these languages lack agreement. Moreover, if the morphologically complex predicate is the justifying evidence for the nested argument structure, another point to be acknowledged is that causatives in Korean and Japanese can be not only morphological but also syntactic. Consider the following example of Korean causatives:

- (5.33) Mary<sub>i</sub>-ka John<sub>j</sub>-ul caki<sub>i/j</sub>-uy chayk-ul ilk-ke hayessta  
 Mary-NOM John-ACC book-ACC read-KE did  
 ‘Mary<sub>i</sub> made John<sub>j</sub> read a self<sub>i/j</sub>’s book’

The above example shows another way of constructing causatives in Korean where the causative verb *hayessta* (did) is connected to the lexical verb with the complementizer *ke*. If nested argument structure can only be used for analysing morphological causatives, then it is dubious that the binding relation between *caki* and the causer *Mary* is licensed in (5.33). This clearly shows that nested argument structure cannot be restricted to morphologically formed constructions. From this position, therefore, we suggest that nested argument structure should be available in languages which exhibit long-distance anaphora<sup>3</sup>.

Based on this assumption, we can now illustrate what the representation of the argument structures of (5.23) and (5.25) should be (examples repeated):

- (5.34) a. John<sub>i</sub>-i [Mary<sub>j</sub>-ka [Bill<sub>k</sub>-i caki<sub>i/j/k</sub>-lul choahanta]-ko sayngkakhanta]-ko  
 John-NOM Mary-NOM Bill-NOM self-ACC like-COMP thinks-COMP  
 malhayssta  
 said  
 ‘John<sub>i</sub> said that Mary<sub>j</sub> thinks that Bill<sub>k</sub> likes self<sub>i/j/k</sub>’  
 b. ARG-ST < John<sub>i</sub>, S:< Mary<sub>j</sub>, S:< Bill<sub>k</sub>, caki<sub>i/j/k</sub> > > >
- (5.35) a. John<sub>i</sub>-i caki<sub>i</sub>-uy ayin-eytayhay malhayssta  
 John-NOM self-GEN lover-about talked  
 ‘John<sub>i</sub> talked about self<sub>i</sub>’s lover’  
 b. ARG-ST < John<sub>i</sub>, PP:< NP:< caki<sub>i</sub> > > >

With this kind of nested argument structure at hand, we can now make *caki* visible through to all possible antecedents and the binding patterns of *caki* can be explained based on the standard a-binding principle.

<sup>3</sup>One will have to be careful in stating the binding conditions so that arguments in higher clauses are not accessible to anaphors in embedded clauses in languages which do not have long-distance binding.

We have now discussed the basic structure of ARG-ST which we have to assume in order to determine the binding relations of *caki*. However, we have not yet shown or discussed the exact and detailed structure inside the local domain of argument structure. That is, we have not considered how many and what kind of arguments are allowed in the argument structure and how they should be ordered. We also have another important issue left to consider; how to make the lexical argument hierarchy relevant to ARG-ST. These issues will be dealt with in the following sections and for now we will conclude the current section with the suggestion that languages with long-distance anaphora should allow an argument structure to embed another argument structure of its complement head, so that long-distance anaphors can have access to their possible antecedents in higher clauses, and such configuration of the argument structure should not be limited to morphologically complex causatives and passives.

### *The Lexical Argument Hierarchy in ARG-ST*

In this section, we will show that what is assumed about the arrangement of dependents of a head in argument structure in HPSG cannot be extended to explain long-distance anaphora. We will argue that it is, instead, the lexical argument hierarchy that should determine the order of the elements in the argument structure in order to predict the correct bindings in subject and non-subject binding patterns of long-distance anaphora in a more unified way.

**Problems with the standard representation of ARG-ST** We briefly discussed in section 4.2.2 how the level of argument structure came to be incorporated into HPSG (ARG-ST), and how the elements in ARG-ST are ordered in terms of obliqueness, which is a crucial notion for defining the a-binding principle (previously, o-binding). Further, it is proposed in Manning (1996) that the argument structure adopted in HPSG is a syntactic level of representation. Earlier work on argument structure has argued for the existence of a level of argument structure separate from

grammatical relations or S-structure configurations (Bresnan & Kanerva 1989, Bresnan & Zaenen 1990, Grimshaw 1990, Alsina 1996)<sup>4</sup>. Though the level of ARG-ST in HPSG is inspired by those conceptions, it is developed as a particular conception of a level of *syntacticized* argument structure within HPSG.

In order to show that the argument structure does not involve any aspects of the semantics of the sentence, Manning & Sag (1998*a*) draw the following distinctions between argument structure and the value of the CONTENT feature:

- (5.36)
- a. ARG-ST is a syntactic level, where things like expletive core arguments are represented, while CONTENT only contains semantic arguments.
  - b. CONTENT represents semantic phenomena like quantification, whereas ARG-ST does not.
  - c. CONTENT provides a fine-grained semantic classification, whereas ARG-ST does not.
  - d. Passivization rearranges core arguments, and so affects ARG-ST, but it does not change the meaning and so the CONTENT remains unchanged.

(Manning & Sag 1998*a*, p.8, fn.7)

Furthermore, the arrangement of the dependents of a head primarily relies on the notion of obliqueness, but it is also handled by the syntactic distinction of core arguments (or terms) and oblique arguments. Following Hellan (1988) and Dalrymple (1993), Manning & Sag (1998*a*) divide the elements in ARG-ST into two groups: one for core arguments (terms) and the other for obliques<sup>5</sup>, where core arguments a-command obliques and, within each of the groupings of core arguments and obliques, a-command reflects argument prominence. The supporting evidence

---

<sup>4</sup>We will discuss ARG-ST from this view in Chapter 6

<sup>5</sup>Terms are considered to be direct arguments of the verb and the subject, while obliques are indirect arguments.



is taken by Manning & Sag (1998*a*) from Arka (1998) who provides the following example in Balinese:

- (5.37) a. Nyoman nakomang awakue teken ia  
 Myoman AV.ask self to 3SG  
 ‘Nyoman<sub>i</sub> asked (to) him/her<sub>j</sub> (about) self<sub>i/\*j</sub>’
- b. Iang ngedengang I Wayan sig awakue  
 1SG AV.show I Wayan to self  
 ‘I showed I Wayan<sub>i</sub> to himself<sub>i</sub>’

(Manning & Sag 1998*a*, p.20, ex.31)

In the above, the marked binding relations show that in (5.37a) an oblique goal *ia* cannot bind a term Theme *awakne*, while in (5.37b), a term Theme *I Wayan* can bind an oblique goal *awakne*. This proves the o-commanding relation between term arguments and oblique arguments.

While Manning & Sag (1998*a*) argue for a syntactic level ARG-ST with a distinction of core vs. oblique arguments, they oppose a thematic approach to binding based on the commonly assumed thematic hierarchy, for example, that of Jackendoff’s (1972) which we saw (in section 3.2.2), which is used to explain data such as:

- (5.38) a. I talked to John about himself  
 b. ??I talked about John to himself

While the thematic hierarchy accounts for the examples above, Manning & Sag (1998*a*) argue that it crucially fails to explain the examples shown below:

- (5.39) a. Mary explained John<sub>i</sub> to himself<sub>i</sub>  
 b. Mary introduced Bob<sub>i</sub> to himself<sub>i</sub>

Given Jackendoff’s (1972) thematic hierarchy where the Goal outranks the Theme, the binding relations where the Theme *John* in (5.39a) and *Bob* in (5.39b) bind the Goal reflexives cannot be explained. However, for Manning & Sag (1998*a*), the distinction between core and oblique arguments becomes very useful for examples like

(5.39) because *John* and *Bob* as term arguments are allowed to antecede the oblique Goal reflexives. With this distinction between core/oblique arguments, they also adopt the notion of thematic role hierarchy for the arrangement of elements within each grouping of core and oblique arguments. For example, in (5.38), the arguments apart from the subject are both obliques, thus they are competing amongst themselves as obliques. Within each of the core or oblique groupings, it is suggested that arguments will be ordered according to role prominence, roughly in accord with traditional thematic hierarchies (they assume that of Jackendoff (1972)). Thus, “the ordering within each grouping will differ *little* across languages” (Manning & Sag (1998*a*), emphasis ours). This proposal also corresponds to proposals by Hellan (1988) and Dalrymple (1993). This way, keeping the notion of the thematic hierarchy among each of the core and oblique groups in ARG-ST proves to be advantageous in explaining the binding relations in (5.38) where the Goal can bind the Theme but not the other way around. This is suggested by the evidence in English as in (5.38) and also in Balinese as follows:

(5.40) Tiang matakon teken anake ento unduk awakne  
 1SG AV.ask to person that about self  
 ‘I asked (to) the person<sub>i</sub> about him/herself<sub>i</sub>’

However, we argue that this distinction between core and oblique arguments incorporating the notion of thematic role hierarchy does not provide a complete answer to long-distance anaphora. Firstly, even though the suggestion about incorporation of the thematic hierarchy predicts the correct binding relations as suggested by the examples in (5.38) and (5.40), at the same time, there can be found examples in long-distance anaphora such as *caki*-binding and *zibun*-binding in Japanese which contradict that suggestion. Observe for example the following sentences:

(5.41) a. Korean  
 John<sub>i</sub>-i Bill<sub>j</sub>-ekey *caki*<sub>i/\*j</sub>-etayhayse malhayssta  
 John-NOM Bill-DAT self-about talked  
 ‘John<sub>i</sub> talked to Bill<sub>j</sub> about self<sub>i/\*j</sub>’

b. Japanese

John<sub>i</sub> ga Bill<sub>j</sub> ni zibun<sub>i/\*j</sub> no koto o hanasita  
 John NOM Bill DAT **self** GEN matter ACC talked  
 'John<sub>i</sub> talked to Bill<sub>j</sub> about self<sub>i/\*j</sub>'

(Kuno 1972)

(5.41a) and (5.41b) correspond to the English example in (5.38), but the binding relations differ in that the anaphor in the Oblique Theme is not bound by the Goal oblique. This casts doubt on employing the thematic role hierarchy to arrange the elements within each grouping for predicting the correct binding relations. Furthermore, it seems that even local anaphora in some languages cannot be explained by the representation suggested by Manning & Sag (1998a) as shown below:

(5.42) a. Italian

Gianni ha parlato a Maria<sub>i</sub> di se stessa<sub>\*i</sub>  
 Gianni talked to Maria about herself  
 'Gianni talked to Maria<sub>i</sub> about herself<sub>\*i</sub>'

(Giorgi 1991, p.203, fn.2, ex. iib)

b. Greek

\*O Kostas milise sti Maria gia ton eafto tis  
 the Kostas talked to the Maria for the self her  
 '\*Kostas talked to Mary about herself'

In Italian, *se stessa* is the clause-bound anaphor (Giorgi 1991, p.186), thus we would expect it to behave in a similar way to the locally-bound anaphor in other languages such as English (e.g., in (5.38)). However, the given binding relation indicates that the Goal oblique *Maria* cannot bind the Theme oblique anaphor (*se stessa*), which contradicts the ARG-ST which assumes Jackendoff's (1972) thematic role hierarchy. The same applies to Greek as illustrated in (5.42b). On a slightly different note, moreover, even if it did predict the correct binding relations, employing the thematic role hierarchy does not seem to reflect Manning and Sag's argument for a *syntacticized* ARG-ST. It also shows that their distinction between core and oblique

arguments does not generalise over locally-bound anaphors cross-linguistically, and this is related to the following second problem.

The more serious problems come from Manning and Sag's distinction between core and oblique arguments. We have seen that they suggest that the term arguments o-command the oblique arguments. This, however, contradicts the binding of morphologically complex, and locally bound, reflexives in Norwegian as below:

- (5.43) a. Jon<sub>i</sub> fortalte oss om seg selv<sub>i</sub>  
 Jon told us about himself  
 b. Vi fortalte Jon<sub>i</sub> om seg selv<sub>\*i</sub>  
 we told John<sub>i</sub> about himself<sub>\*i</sub>  
 [p.33, ex.10a and 11](Hellan 1991)

The binding relation shows that the term subject binds the Theme Oblique but the term object cannot antecede it, and this is where the core/oblique distinction breaks down. This is also suggested by evidence from Polish and Russian:

- (5.44) a. Polish  
 Jan<sub>i</sub> opowiada Piotrowi<sub>j</sub> o sobie<sub>i/\*j</sub>  
 Jan-NOM talks Piotr-DAT about self  
 'Jan<sub>i</sub> told Piotr<sub>j</sub> about self<sub>i/\*j</sub>'  
 (Reinders-Machowska 1991, p.140, 0ex.1)  
 b. Russian  
 Milicjoner<sub>i</sub> rassprašival arestovannogo<sub>j</sub> o sebe<sub>i/\*j</sub>  
 policeman-NOM questioned suspect-ACC about self-LOC  
 'The policeman<sub>i</sub> questioned the suspect<sub>j</sub> about himself<sub>i/\*j</sub>'  
 ((Iida 1996, p133, ex.106) taken from (Rappaport 1986))

In Polish as in (5.44), the binding relation of the local anaphor *sobie* shows that even though *Piotr* is a term argument, it cannot bind *sobie* in the Theme oblique. The same applies to Russian as in (5.44b).

We have seen so far two problems in the representation of ARG-ST proposed by Manning & Sag (1998*a*). The first one lies in employing the thematic role hierarchy to arrange the order of the elements in each of core/oblique groups. The second problem is, as suggested in anaphor-binding in Korean and Norwegian, that the core/oblique distinction may contribute to the explanation of subject binding, but does not hold for cases of non-subject binding. We agree that both thematic and syntactic prominence is in operation in explaining long-distance anaphor binding. However, we argue that the definitions of both thematic and syntactic prominence must be altered. The first should be modified along the lines of the lexical argument hierarchy, and the second in terms of the fundamental difference between external argument and internal arguments, in order to define the representation of ARG-ST to predict the correct binding behaviour in long-distance anaphora.

**Implementing the Lexical Argument Hierarchy in ARG-ST** In the previous section 5.1.2, we proposed that both subject and non-subject binding of *caki* can be consistently explained in terms of the lexical argument hierarchy. It was shown that the lexical argument hierarchy is defined in terms of both thematic roles and syntactic terms. It is partially motivated by thematic roles in that the thematic hierarchy employed in the lexical argument hierarchy is not an absolute and inherent one, but is dependent on each particular lexical head. Further, it is also syntactic because the priority for the most prominent element goes to the external argument. The prominence order of the rest of the arguments then relies on the lexically determined thematic hierarchy. We argue that it is the ordering by such syntactic priority and thematic hierarchy that should be reflected in ARG-ST, replacing the distinction between core and oblique arguments and the notion of absolute thematic role hierarchy suggested by Manning & Sag (1998*a*).

Firstly, with the lexical argument hierarchy in ARG-ST, we are able to explain the long-distance binding relations in Korean and Japanese as in (5.41). The following illustrates what the ARG-ST of each sentence should look like:

(5.45) a. ARG-ST < John<sub>i</sub>, PP[*about*]:<caki<sub>i/\*j</sub>>, PP[*to*]:<Bill<sub>j</sub>> >

- b. ARG-ST < Ziroo<sub>i</sub>, PP[*about*]:<zibun<sub>i/\*j</sub>>, PP[*to*]:<Hanako<sub>j</sub>> >

In each sentence, the external argument *John* in (5.45a) and *Ziroo* in (5.45b) comes as the first element in the ARG-ST. This is followed by the Theme argument *caki* and *zibun*, respectively. As the Goal is the least prominent argument specified by the lexical head *talked*, it takes the last position in the ARG-ST, indicated by *Bill* and *Hanako* in (5.45). In the given ARG-STs, the binding relations of *caki* and *zibun* are subsequently predicted.

Secondly, our representation of ARG-ST also surmounts a problem from the following data which contradict Manning & Sag's (1998a) ARG-ST representation:

- (5.46) John<sub>i</sub>-i Bill<sub>j</sub>-lopwute caki<sub>i/j</sub>-uy chayk-ul patassta  
 John-NOM Bill-from self-GEN book-acc received  
 John<sub>i</sub> received self<sub>i/j</sub>'s book from Bill<sub>j</sub>'

In Manning & Sag's (1998a) ARG-ST, *caki's book* as a term argument o-commands the Goal oblique, thus *caki* should not be bound by the Goal oblique. However, the prediction is not borne out. On the other hand, our ARG-ST will put the elements in the right order for the Goal oblique to bind the term argument as shown below:

- (5.47) ARG-ST < John<sub>i</sub>, PP[*from*]:< Bill<sub>j</sub> >, NP:< caki<sub>i/j</sub> > >

The first element is the external argument *John*. As for the lexical head *received*, the most prominent thematic role is the Source, that is, *Bill*, which is followed by the remaining argument, the Theme, *caki's book*. Thus, our representation of ARG-ST correctly predicts that *caki* as the last element in the ARG-ST can take either *John* or *Bill* as the antecedent.

So far, we have examined in detail the representation of ARG-ST in HPSG proposed by Manning & Sag (1998a) and argued that even though it explains anaphoric binding relations in languages like English and Balinese, it cannot generalise over locally bound anaphors in Italian, Greek, Polish and Norwegian, let alone over long-distance anaphors like *caki*. We have suggested that at least for long-distance

anaphora, it is not the term/oblique distinction and the absolute thematic role hierarchy that should be incorporated into the representation of ARG-ST, but the lexical argument hierarchy that should be reflected in representing the order of the arguments in ARG-ST in order to provide the correct analysis for long-distance anaphora. However, as one might have noticed, encoding the external argument as the most prominent element in ARG-ST causes a problem as the traditional notion of argument structure is not generally assumed to be a level of syntactic representation (for example, Grimshaw (1990)). This is also problematic for Manning & Sag's (1998a) distinction of core and oblique arguments in ARG-ST. This will be again pointed out in the next section and the main discussion of this problem will be dealt with in Chapter 6.

Having provided a representation of ARG-ST, what is still left unexplained is the different binding relations between complements and adjuncts, a question that we have raised in section 5.1.2 but put off. The next section will explore this issue, and thus will provide a more complete picture of the representation of ARG-ST.

### *Caki-Binding in Structured ARG-ST*

In section 5.1.2, we noted that there is a discrepancy in *caki*-binding in constructions which involve complements and adjuncts, which we repeat below as (5.48) and (5.49) with the corresponding ARG-ST:

- (5.48) a. John<sub>i</sub>-i Mary<sub>j</sub>-lul caki<sub>i/j</sub> cip-ulo tolye ponayssta  
 John-NOM Mary-ACC self house-to return sent  
 'John<sub>i</sub> returned Mary<sub>j</sub> to self<sub>i/j</sub>'s house'  
 ARG-ST < John<sub>i</sub>, Mary<sub>j</sub>, PP[*to*]:< NP: < self<sub>i/j</sub> > > >
- b. Mary<sub>i</sub>-ka aki<sub>j</sub>-lul caki<sub>i/j</sub> chimday-e nwuphiessta  
 Mary-NOM baby-ACC self bed-in put  
 'Mary<sub>i</sub> put (her) baby<sub>j</sub> in self<sub>i/j</sub>'s bed'  
 ARG-ST < Mary<sub>i</sub>, baby<sub>j</sub>, PP[*in*]: < NP:< self<sub>i/j</sub> > > >



- (5.49) a. Mary<sub>i</sub>-ka John<sub>j</sub>-ul [**caki<sub>i/\*j</sub>-uy** kapang-ulo] ttalyessta  
 Mary-NOM John-ACC **self-ACC** bag-with hit  
 ‘Mary<sub>i</sub> hit John<sub>j</sub> **with self<sub>i/\*j</sub>’s bag**’  
 ARG-ST < Mary<sub>i</sub>, John<sub>j</sub>, PP[with]: < NP:< self<sub>i/\*j</sub> > > >
- b. John<sub>i</sub>-i Bill<sub>j</sub>-ul caki<sub>i/\*j</sub> cip-ese mannassta  
 John-NOM Bill-ACC self house met  
 ‘John<sub>i</sub> met Bill<sub>j</sub> at self<sub>i/\*j</sub>’s house’  
 ARG-ST < John<sub>i</sub>, Bill<sub>j</sub>, PP[at]: < NP:< self<sub>i/\*j</sub> > > >

In ARG-ST in (5.48a) and (5.48b), we have the external argument as the first element, followed by the Theme, which is in turn followed by the least active role, the Location. In this order, the binding relation is correctly predicted so that *caki* in the Location role is bound by the higher elements in the ARG-ST. Compared with this representation, the ARG-STs in (5.49) show the same order as in (5.48) in terms of the order of the external argument, the Theme and the Location. However the given binding relation is not quite the same as in (5.48). We have noted that this different binding relation is due to the distinction between complement PPs as in (5.48) and adjunct PPs in (5.49). Japanese also exhibits this peculiar binding relation in adjuncts where *zibun* in the Location adjunct can be bound by the subject but not by the direct object as shown below:

- (5.50) Taroo<sub>i</sub> wa kodomo<sub>j</sub> o zibun<sub>i/\*j</sub> no kutu de butta  
 Taroo TOP child ACC SELF ’s shoe with hit  
 ‘Taroo<sub>i</sub> hit the child<sub>j</sub> with self<sub>i/\*j</sub>’s shoe’

(Hellan 1991, p.40, ex.24)

Such binding relations between complements and adjuncts can also be found in binding patterns of local anaphora in other languages, Czech being a case in point:

- (5.51) a. Karel<sub>i</sub> narovnal desticky<sub>j</sub> na sebe<sub>i/j</sub>  
 Karl stacked plates on himself/themselves

- b. Kouzelník<sub>i</sub> zkrřížil tyče<sub>j</sub> přes sebe<sub>i/j</sub>  
 ‘magician crossed bars across himself/themselves’

(Toman 1991, p.155, ex.14a,b)

The binding domain for reflexives in Czech never extends beyond the domain of a finite clause, thus exhibiting local anaphora. The two examples in (5.51) have complement PPs and the reflexives in the PPs are bound by either the subject or the direct object, just as in the Korean examples in (5.48). On the other hand, the following Czech examples with adjuncts display different binding patterns, again just as in Korean and Japanese in (5.49) and (5.50):

- (5.52) a. Jana<sub>i</sub> pomáhala Karlovi<sub>j</sub> in svém<sub>i/\*j</sub> bytě  
 Jana helped Karl<sub>dat</sub> in her-REFL/his-REFL apartment
- b. Jana<sub>i</sub> čekala na karla<sub>j</sub> ve svém<sub>i/\*j</sub> bytě  
 Jana waited for Karl in her-REFL/his-REFL apartment

(Toman 1991, p.158, ex.20)

- (5.53) Jana<sub>i</sub> zavraždila Karla<sub>j</sub> ve svém<sub>i/\*j</sub> bytě  
 Jana murdered Karla in her-REFL apartment

(Toman 1991, p.157, ex.18)

The above data from Czech show that reflexives in adjunct PPs can be bound by the subject but not by the direct object, which was also observed in Korean and Japanese ((5.49) and (5.50)). Further, this can be also demonstrated in Russian:

- (5.54) Maša<sub>i</sub> videla Natašu<sub>j</sub> v svoej<sub>i/\*j</sub> kvartire  
 Masah-NOM saw Natasha-ACC in self apartment-LOC  
 ‘Masha<sub>i</sub> saw Natasha<sub>j</sub> in her<sub>i/\*j</sub> apartment.’

(Iida 1996, p133, ex.107) taken from (Rappaport 1986)

It seems therefore that the difference in binding relations with respect to the distinction between complement and adjunct PPs is not only limited to long-distance anaphora but is also applicable to cases of local anaphor-binding patterns. Does this mean that adjuncts should be treated differently from complements?

In HPSG, there has been a proposal by Bouma *et al.* (1998) that adjuncts should not be included in ARG-ST, but should appear in an additional list called DEPS-list (dependents list)<sup>6</sup>. This proposal was motivated by the need for a unified account of extraction, including subject, complement and adjunct extraction. If we borrow this representation for the moment, and contemplate its effects on our binding phenomena, it would certainly explain why the reflexives are not bound by the direct objects (because reflexives in adjunct PPs are not in ARG-ST), but at the same time, we will immediately have a problem in accounting for the cases of subject-binding (because the reflexives in adjunct PPs are not in ARG-ST, thus are not able to access the subject antecedent). Thus, making adjuncts unavailable in ARG-ST only complicates the problem rather than solving it. On the other hand, if we keep adjuncts in ARG-ST, we will have the same problematic representation of ARG-ST we saw in (5.48) and (5.49). Therefore, with either of those options, we find ourselves back to square one.

In order to overcome the problem, we propose first that adjuncts should be included in ARG-ST, otherwise, the anaphor-binding principle in terms of a-binding will never suffice to explain any legitimate case of binding relations in adjuncts. Second, we propose that adjuncts should be made distinct from complements within ARG-ST by distinguishing complements from subject within ARG-ST in a similar way that subjects are split from the complements in the VALENCE list (Borsley 1987, Pollard & Sag 1994). In consequence, the resulting representation of ARG-ST should look roughly like the following:

---

<sup>6</sup>There has been a concern that postulating DEPS-LIST alongside ARG-ST causes redundancy in HPSG (for example, Alexopoulou (1999)). We will return to this issue in the next Chapter 6.

(5.55) ARG-ST < subject, < comps, ...>, adjuncts >

Given this configuration of ARG-ST, we formulate the principle for *caki*-binding as follows:

(5.56) Principle for *Caki*-Binding (Preliminary)

*Caki* must be bound by a preceding element in the same, or higher, level of ARG-ST.

This principle differs from that of Manning's (1996), in that it does not use the notion of *a-subject* in predicting a possible antecedent. We have repeatedly pointed out that the notion of *a-subject* is too restricted to cover *caki*-binding, for example, in cases like (5.48) where the direct objects can be an antecedent of *caki* even though they can not be understood as either a syntactic subject or a logical subject (see also section 4.3.2 and 4.3.3).

Based on our representation of ARG-ST and the proposed principle in (5.56), the examples in (5.48) and 5.49) should contain the following ARG-ST lists:

(5.57) a. (= (5.48a))

*returned*: < John<sub>i</sub>, < Mary<sub>j</sub>, PP[*to*]:<NP:<self<sub>i/j</sub>>> > >

b. (= (5.48b))

*put*: < Mary<sub>i</sub>, < baby<sub>j</sub>, PP[*in*]:<NP:<self<sub>i/j</sub>>> > >

(5.58) a. (= (5.49a))

*hit*: < Mary<sub>i</sub>, < John<sub>j</sub> >, PP[*with*]:<NP:<self<sub>i/\*j</sub>>> >

b. (= (5.49b))

*met*: < John<sub>i</sub>, < Bill<sub>j</sub> >, PP[*at*]: <NP:<self<sub>i/\*j</sub>>> >

In (5.57a) and (5.57b), the subjects are separated from the complements (the direct object and the complement PPs) and *caki* in the PPs is correctly bound by the *a*-subjects (the first element in the ARG-ST). On the other hand, in (5.58a) and (5.58b), the complements are distinguished from the subject as well as from the adjuncts. *Caki* is firstly bound by the subject which is the first element in the

ARG-ST which contains the anaphor. However, the direct object cannot antecede *caki* since it is not in the same level of ARG-ST.

Assuming that PPs as in *heard from* and *talked to.. about* are selected by the head, they should be treated as complements, and the representation above should satisfy every instance of *caki*-binding.

#### 5.1.4 Summary and Conclusion

In this section, we have provided a unified account for *caki*-binding including cases of both subject and non-subject binding.

Firstly, we looked at the most problematic cases of *caki*-binding, concentrating on cases of non-subject binding which most existing approaches have failed to consistently explain at the same time as subject-binding.

Secondly, we have argued that the right way to approach such problems is to base the explanation on the lexical argument hierarchy proposed by Kiss (1991), which not only promotes the syntactic notion of external argument, but also incorporates the notion of a lexically determined thematic hierarchy where the thematic prominence is dependent on the lexical item, rather than on an absolute and inherent hierarchy such as that of Jackendoff (1972).

Thirdly, we have shown that it is the lexical argument hierarchy which should be incorporated into ARG-ST in HPSG in order to account for long-distance anaphora, instead of the distinction between core and oblique arguments and Jackendoff's (1972) thematic hierarchy which is proposed by Hellan (1988) and (Dalrymple 1993), and adopted by Manning & Sag (1998*a*) in HPSG terms.

Fourthly, we proposed that the distinct binding relations between complements and adjuncts, which can be widely observed across languages, can be explained by separating the complements from the subject within the level of ARG-ST under our principle that anaphors must be bound by a higher element in the same or higher level of ARG-ST. This principle does without the notion of a-subject proposed by

Manning (1996) as this notion is not adequate to accommodate all the possible antecedents of *caki*.

Having formulated a complete frame which underlies *caki*-binding, we are left with one question about the problem raised in section 5.1.3. In implementing Kiss's (1991) lexical argument hierarchy, we have shown that the external argument should take the first position as the most prominent element despite the fact that argument structure is traditionally viewed as an autonomous level of syntactic structure (Alsina 1996). This problem relates not only to our distinction between subject and complements, and in turn, between complements and adjuncts in ARG-ST as proposed in the last section, but also to the Manning & Sag's (1998a) distinction between core and oblique arguments. The following Chapter will highlight this issue, together with other theoretical consequences of our new principle of *caki*-binding.

Even though we have represented our analysis in ARG-ST throughout this section, whether ARG-ST provides a satisfying base for it is another issue. In the following chapter, we will address a problem where our binding principle using the lexical argument hierarchy is well suited in terms of both ARG-ST and the traditional argument structure. As for what is proposed in Manning (1996), it is not clear whether ARG-ST is a level independent of the traditional argument structure which happens to bear the same name, or whether it is a convenient subset of the traditional argument structure. Therefore, the following chapter involves two parts to be considered together: i) problematic parts in the explanation of *caki*-binding for both ARG-ST and the traditional argument structure, ii) problematic aspects of a-binding in ARG-ST from the point of view of the traditional argument structure. In the course of the discussion, we will show what the proper level for *caki*-binding should be like and will discuss its theoretical consequences. Even though we will eventually abandon ARG-ST for the level of *caki*-binding, this does not necessarily mean that the result is a more complicated theory. On the contrary, our proposal will get rid of the redundancy that the theory currently suffers from, and thus will make it more elegant.

## Chapter 6

### Theoretical Consequences:

### Argument Structure, ARG-ST, DEPS and *Caki*-binding

In the course of providing a binding principle for *caki* in the level of the ARG-ST representation, we have proposed three major modifications to the representation of ARG-ST in Manning (1996), Manning & Sag (1998*a*) and Bouma *et al.* (1998). First, we have made ARG-ST a head feature, so that it can appear at the phrasal level. The motivation for this was to allow the antecedent in a higher clause to be visible to *caki* in the lower clause. Second, unlike the distinction between term and oblique arguments in ARG-ST in Manning & Sag (1998*a*), in adopting the notion of the lexical argument hierarchy (Kiss 1991) in ARG-ST, we have proposed that the external argument should be assigned to the first position as the most prominent element and that the rest of the arguments are ordered in terms of a lexically defined thematic hierarchy. Third, in order to explain the discrepancy in *caki*-binding between complements and adjuncts, we have not only suggested that adjuncts should be included in ARG-ST but also that the representation of ARG-ST must be structured in such a way as to allow subjects (external arguments) to be separated from complements which, in turn, are distinguished from adjuncts.



Even though these modifications enable us to adequately account for *caki*-binding, the question whether these modifications can be represented in the level of ARG-ST remains open. For example, the external argument appears as the first element in ARG-ST even though such grammatical specification is not directly encoded in the standard argument structure (Grimshaw 1990, Alsina 1996)<sup>1</sup>. In fact, what is standardly accepted as a representation of ARG-ST as in Manning & Sag (1998a) also leads to the same problem (Alsina 1993) since it encodes syntactic specifications such as expletives and the notion of term and oblique arguments.

In this chapter, we will focus on providing a proper level at which *caki*-binding should be determined, and argue that it is DEPS where *caki*-binding should be determined, satisfying the principle we proposed in the last section. To do so, we will first highlight problems coming from explaining *caki*-binding both in ARG-ST and in the traditional argument structure on the grounds of the three points summarised above (section 6.1). Secondly, we will show that ARG-ST is not an adequate level for stating the binding principle for *caki*. Instead, we will provide a preliminary level, called BINDING-ST as an example in order to show what this level should be like (section 6.2). We will however suggest that such a level overlaps with DEPS, the level of extraction phenomenon, and that construing two independent levels for different syntactic phenomena is theoretically unattractive (section 6.3). After examining the extraction mechanism in HPSG, we will show that assuming that DEPS should be viewed as an interface between syntax and semantics, just like ARG-ST, we can reduce two independent levels, BINDING-ST and DEPS to one (DEPS), without disturbing the extraction operation in HPSG. This way, DEPS will be made a level for both binding and extraction without incurring unnecessary theoretical costs.

---

<sup>1</sup>Throughout this section, we will refer to the representation of argument structure adopted in HPSG as ARG-ST and argument structure as in Grimshaw (1990), for example, as argument structure.

## 6.1 Problems in ARG-ST in HPSG

### 6.1.1 Syntactic Functions in ARG-ST

Since argument structure was introduced into generative linguistics in works such as Bresnan (1978) and Williams (1980), one of the important things about it is that it is not generally assumed to be a level of syntactic representation that encodes information about syntactic category. Argument structure is a representation which represents relative prominence relations among arguments. The prominence relations are (in part) determined by the thematic properties of the predicate (via the thematic hierarchy)<sup>2</sup>. Further, an important property of the prominence theory of argument structure is that it does not make use of theta role labels in the argument structure representation. Argument structure represents the argument-licensing capacity of a predicate without specifying any semantic information about its arguments, except for their relative prominence. Hence, two verbs with different theta roles but the same prominence relations will be indistinguishable as far as argument structure is concerned (Grimshaw 1990).

While argument structure is sensitive to semantics, the syntactic information of a predicate (e.g., what kind of grammatical functions it takes) is not directly encoded in argument structure. Argument structure only provides the minimal information

---

<sup>2</sup>Another prominence relation involved in argument structure is the aspectual properties of the predicate. The most directly relevant data come from the *fear* and *frighten* classes of verbs. Consider the following:

- (6.1)     a.    The baby feared the barking dog  
          b.    The barking dog frightened the child  
          c.    \*The child frightened the barking dog  
              (given that it is the child who is frightened)

Even though the thematic hierarchy always assigns the Experiencer *the child* more prominence than the Theme *barking dog* for both *fear* and *frighten*, it does not explain why the Experiencer is realised as the subject in the *fear* class and not in the *frighten*. Grimshaw (1990) argues that in order to explain such examples, it is the aspectual prominence based on event structure of the predicates that should jointly determine the prominence relations among arguments together with the thematic hierarchy. For details, see Grimshaw (1990, section 2.3). For similar views, see Pustejovsky (1988) and Belletti & Rizzi (1988).

necessary for deriving alternative syntactic frames, or subcategorisation of predicates, and for deriving the alternative syntactic frame when an alternation exists such as in the active and passive pair given below:

- (6.2)    a.    John built the cottage  
          b.    The cottage was built by John

In (6.2a) and (6.2b), the predicate takes two arguments: the builder (Agent) and one corresponding to the entity being built (Theme). However, whereas the builder is the subject (6.2a), it is not in (6.2b) where, instead, the entity built is the subject. In spite of the important differences in surface syntax between an active sentence and its corresponding passive, as argument structure only represents the relative prominence among arguments, the argument structures in both cases in (6.2) are essentially the same. Hence syntactic functions such as subject and object do not play any role in determining the arrangement of the arguments of the predicate. Further, syntactic functions are directly specified in argument structure.

Given this representation of argument structure, accommodating the lexical argument hierarchy which our binding principle is based on is problematic. By our arguing that the external argument should be put in the first position in ARG-ST, we depend on the grammatical function for the arrangement of the arguments of the predicate. The same problem also results from our distinction between subject and complements within ARG-ST because this distinction is associated with grammatical functions rather than with semantics. Therefore, at this point, it is not clear whether ARG-ST in HPSG is parallel to what is traditionally viewed as argument structure.

It is, however, clear that ARG-ST in HPSG as proposed in Manning & Sag (1998*a*) is a very much syntacticized notion. First, as we have seen in section 5.1.3, it was made clear that ARG-ST is a syntactic level, and thus passivization, for example, rearranges core arguments, and so affects ARG-ST (see (5.36)). Secondly, they propose a distinction between term and oblique arguments and these are defined in terms of syntactic functions. This also indicates that syntactic function is involved

in arranging the arguments of the predicate. Thirdly, the a-binding principle (also in Manning (1996)) predicts that both syntactic and logical subjects (a-subject) must bind anaphors. And anaphors should be bound by the first element in some level of ARG-ST which is identified with either the syntactic or the logical subject. This makes the syntactic subject necessarily appear in the first position in ARG-ST, which again makes the arrangement of arguments in ARG-ST dependent on syntactic functions.

Therefore, ARG-ST behaves very differently from what we would have expected from the traditional argument structure. However, there are certain points where ARG-ST certainly takes advantage of argument structure. For example, elements in each group of term and oblique arguments are ordered in terms of thematic prominence which also determines the order of arguments in argument structure. Further, representing the logical subject as the most prominent element at some level in ARG-ST is very much like in argument structure. For example, in passive, the NP realised by the *by* phrase should be represented as the logical subject in the ARG-ST of the lexical verb, thus the first position in the nested ARG-ST. This certainly is the same in argument structure where the passive has the same representation as its active counterpart such that the Agent expressed by the *by* phrase will be the most prominent element. At this point, then, it is not clear whether ARG-ST is a distinct level of representation from argument structure, or simply some level of representation where semantics and syntax are jointly involved in arranging arguments of the predicate. We will carry on this issue in re-examining our proposal that adjuncts should be included in ARG-ST in the following section.

### 6.1.2 Adjuncts in ARG-ST

In section 5.1.3, we argued that adjuncts should be taken into consideration in determining the binding relations of *caki*, thus we have proposed that adjuncts should be included in ARG-ST. However, this is not what is generally assumed in both traditional argument structure and ARG-ST in HPSG.

A fairly clear distinction in terms of argument structure has been made between arguments and adjuncts in linguistic theories (for example, Grimshaw (1990) and McConnell-Ginet (1982) a.o.). For instance, arguments can be selected and subcategorised in the sense that their presence and the form they take are under the control of individual predicates. Further, arguments must be licensed; they can occur only if they are theta-marked by a predicate as a function of the predicate's argument structure. On the other hand, adjuncts are not theta-marked and do not need to be licensed by a relationship to argument structure. They are not subcategorised. Hence, their form is free, and they are never required in argument structure. Thus although arguments and adjuncts are semantic participants, hence have a representation at semantic structure, they are complementary in representing and defining argument structure; arguments are semantic participants also represented at argument structure, while adjuncts are not represented at argument structure.

In HPSG, on the other hand, in order to provide an account of adjunct extraction on a par with complement extraction, van Noord & Bouma (1994) propose the following Adjunct Lexical Rule:

$$(6.3) \quad \left[ \begin{array}{l} verb \\ \text{SUBJ} \quad \boxed{1} \\ \text{COMPS} \quad \boxed{2} \end{array} \right] \mapsto \left[ \begin{array}{l} \text{SUBJ} \quad \boxed{1} \\ \text{COMPS} \quad \boxed{2} \oplus \langle 'adjuncts' \rangle \end{array} \right]$$

The Adjunct Lexical Rule above adds an adverbial *synsem* to COMPS. Once an adverbial is added to the COMPS list, adjunct extraction is subsumed under the

Complement Extraction Rule of Sag (1997)<sup>3</sup>:

$$(6.5) \quad \left[ \begin{array}{l} \textit{word} \\ \text{SUBJ} \quad \boxed{1} \\ \text{COMPS} \quad \boxed{2} \end{array} \right] \mapsto \left[ \begin{array}{l} \text{SUBJ} \quad \boxed{1} \\ \text{COMPS} \quad \boxed{2} \ominus \langle \textit{gap-ss} \rangle \end{array} \right]$$

Furthermore, in recent developments in HPSG, *synsem* includes two subsorts: *canonical-synsem* (*canon-ss*) and *gap-synsem* (*gap-ss*) as illustrated below:

$$(6.6) \quad \begin{array}{l} \text{a. } \textit{canon-ss} \longrightarrow \left[ \text{NONLOC} \mid \text{SLASH} \quad \{\} \right] \\ \text{b. } \textit{gap-ss} \longrightarrow \left[ \begin{array}{l} \text{LOC} \quad \boxed{1} \\ \text{NONLOC} \mid \text{SLASH} \quad \{\boxed{1}\} \end{array} \right] \end{array}$$

*Canon-ss* is associated with all overt expressions, that is, all signs are required to have a SYNSEM value of type *canon-ss*. On the other hand, as for *gap-ss*, the crucial difference is that there is no lexical entry with a SYNSEM value of type *gap-ss*, which is never realized syntactically in a head-valence structure. Returning to (6.3), the Complement Extraction Lexical Rule in (6.5) removes an element from COMPS and instantiates it as a ‘gap’ (i.e., an element of type *gap-ss*). As a result of the Adjunct Lexical Rule in (6.3), an adjunct as an element in the COMPS list can go through the Complement Extraction Rule in order to be extracted. However, adding adjuncts to the COMPS list via the Adjunct Lexical Rule blurs the distinction between adjuncts and truly selected dependents. To preserve this distinction, Bouma *et al.* (1998)

<sup>3</sup>A  $\ominus$  B denotes the list A minus the elements in B. As for the details of the motivation and the operation of the Complement Extraction Lexical Rule, HPSG aims to eliminate the use of null elements, arguing that the use of empty categories in linguistic analysis is unnecessary (Pickering & Berry 1991, Sag & Fodor 1994). The approach to Unbounded Dependency Constructions given in Pollard & Sag (1994) (Ch.9) does not rely on the empty lexical item trace. Instead, heads are assumed to be ‘SLASHED’ via a lexical rule.

$$(6.4) \quad \begin{array}{c} \text{COMPLEMENT EXTRACTION LEXICAL RULE (CELR)} \\ \left[ \begin{array}{l} \text{COMPS} \quad \langle \dots [\text{LOC } \boxed{1}] \dots \rangle \\ \text{INHER} \mid \text{SLASH} \quad \boxed{2} \textit{set} \end{array} \right] \mapsto \left[ \begin{array}{l} \text{COMPS} \quad \langle \dots \rangle \\ \text{INHER} \mid \text{SLASH} \quad \boxed{2} \cup \{\boxed{1}\} \end{array} \right] \end{array}$$

The function of this lexical rule is that a *synsem* object is removed from the head’s COMPS list, and its LOCAL value is moved into the head’s INHER|SLASH set.



assume the level of ARG-ST, which contains all and only the selected arguments of a lexical head, and thus adjuncts do not have a position to satisfy in ARG-ST. Therefore, a lexical entry of a transitive verb which has a singleton COMPS list and one adverbial element should look like the following:

$$(6.7) \quad \left[ \begin{array}{l} \textit{verb} \\ \text{SUBJ} \quad \langle \text{[1]NP} \rangle \\ \text{COMPS} \quad \langle \text{[2]NP, [3]'} \textit{adjuncts}' \rangle \\ \text{ARG-ST} \quad \langle \text{[1], [2]} \rangle \\ \text{DEP-ST} \quad \langle \text{[1], [2], [3]} \rangle \end{array} \right]$$

In the AVM above, an adverbial can be realised in both COMPS and DEPS-ST, but not in ARG-ST<sup>4</sup>. However, in some recent works, it has been proposed that adjuncts should be present in ARG-ST. For example, Przepiórkowski (1997) assumes that adjuncts should be put in ARG-ST (via a lexical rule) from the evidence of case assignment in Korean, Finnish, Russian and Polish, given the assumption that cases are assigned in ARG-ST. Further, under the analysis by Manning *et al.* (1999) for Japanese causatives, the composed adjuncts are actually on the ARG-ST lists of both the verbal stem and the morphologically derived causative word.

Even though these proposals are in line with our intuition concerning *caki*-binding, the addition of adjuncts to the level of the traditional argument structure remains problematic. Hence, it again casts doubt on whether argument structure is an appropriate level at which *caki*-binding should be determined.

### 6.1.3 ARG-ST as a phrasal sign

In section 5.1.3, we have deliberately made ARG-ST a HEAD feature so that it can appear at a phrasal level in order to explain long-distance binding of *caki* to an

---

<sup>4</sup>As briefly mentioned in section 5.1.3, Bouma *et al.* (1998) assume dependency structure (DEPS-ST) as an extended argument structure which consists of the selected arguments plus an underspecified list of adverbial *synsem*. We will discuss DEPS-ST in more details in the upcoming sections 6.3 and 6.4.



antecedent outside the local domain. However, as a representation of selected arguments of a lexical item, both ARG-ST and the traditional argument structure strictly remain at the lexical level. Even though we agree that argument structure is information associated with a lexical item, whether ARG-ST should be so in HPSG is not uncontroversial.

Some linguists have expressed their concern that if ARG-ST remains on lexical signs, there may be problems in accounting for various syntactic phenomena, some of which are resolved by making ARG-ST a HEAD feature and thus appear at the phrasal level. Such phenomena include purpose infinitives in English (David Bexter, June 1998), French causatives and clitic climbing (Mike Calcagno and Carl Pollard, June 1998), Auxiliary-Initial Sentences in Welsh and Polish (Bob Borsely, June 1998)<sup>5</sup>.

Following those lines of argumentation, our proposal of making ARG-ST a HEAD feature does not seem too far fetched. Further, as pointed out in the previous sections, ARG-ST seems to be different from traditional argument structure in that it specifies grammatical functions (6.1.1), includes adjuncts in some cases (6.1.2) and appears at the phrasal level (as above). From this stand, we suggest that even though ARG-ST is essentially motivated by binding theory (Manning 1996), the problems above actually come from the fact that ARG-ST is too restricted a level to explain many syntactic phenomena because it has been (partly) derived from the traditional level of argument structure. Concerning specifically *caki*-binding, an account within the level of ARG-ST seems to necessitate too much modification of the status of ARG-ST itself which results in a theoretically inelegant, not to say incoherent, treatment. Therefore, the preceding discussion points to a double question. First, to what extent is the ARG in ARG-ST justified, and second, given that the main purpose of ARG-ST is to provide a locus for the determination of binding relations, would it not be more appropriate (perhaps more honest) to redefine and reformulate this level of representation as BINDING STRUCTURE? If this turns out to be a viable course

---

<sup>5</sup>The issue of making ARG-ST a HEAD-feature was discussed on line in the HPSG mailing list. The reference in the text refers to the contributions of these scholars to their discussion. The discussion can be retrieved on-line at <http://hpsg.stanford.edu/hpsg-l/1998>.

of action, then the BINDING STRUCTURE in question should be a level of representation independent of the notion of argument structure which can also function as the interface between syntax and semantics. It is this question that we will take up in the following section.

## 6.2 Is ARG-ST a Binding Structure?

Since ARG-ST came to replace the SUBCAT list, there was a concern that it overlaps with VALENCE, as ARG-ST can be derived from valence by means of a constraint like the following:

$$(6.8) \quad word \implies \begin{bmatrix} \text{SUBJ} & \boxed{1} \\ \text{COMPS} & \boxed{2} \\ \text{ARG-ST} & \boxed{1} \oplus \boxed{2} \end{bmatrix}$$

(6.8) above shows that ARG-ST value is simply the APPEND ( $\oplus$ ) of the SUBJ and COMPS list. While this representation may be thought of as merely summarising the valence of a lexical sign, Manning & Sag (1998*b*) present evidence that motivates the introduction of ARG-ST as a level of representation which cannot be reduced to valence. One piece of evidence given by Manning & Sag (1998*b*) is that ARG-ST, unlike valence, can represent *pro* which is syntactically unexpressed yet plays a grammatically significant role.

(6.9)    Naoki-ga    mi-ta  
           Naoki-NOM see-PAST  
           ‘Naoki saw (it)’

(Manning & Sag 1998*b*, ex.3)

The Japanese sentence in (6.9) has an unexpressed object (*pro*) and the lexical entry of the verb is as follows:

$$(6.10) \quad \left[ \begin{array}{l} \text{SUBJ} \quad \langle \boxed{1}\text{NP}[\text{NOM}] \rangle \\ \text{COMPS} \quad \langle \rangle \\ \text{ARG-ST} \quad \langle \boxed{1}_i, \boxed{2}\text{NP}_j \rangle \\ \text{CONT} \quad \left[ \begin{array}{l} \textit{seeing} \\ \text{SEER} \quad i \\ \text{SEEN} \quad j \end{array} \right] \end{array} \right]$$

(Modified from Manning & Sag (1998*b*, ex.4))

The pro-dropped object NP does not appear on the COMPS list as it is not realised on the surface. However, it must still appear on ARG-ST so that it can explain properties such as binding; for instance, here we need to explain that (6.9) cannot mean ‘Naoki saw himself’. Therefore ARG-ST is used to explain properties such as binding in terms of ‘deep’ subcategorisation (Manning & Sag 1998*b*) unlike valence.

Furthermore, even more important evidence for the need of an ARG-ST list separate from valence comes from binding facts in passives, causatives and ergativity in languages such as West Austronesian languages, Inuit and Russian (Manning & Sag 1998*a*, Manning 1996) (see also section 5.1.3). The advantage of having ARG-ST as an independent level in such languages is that ARG-ST can not only represent the syntactic subject as the first element, but can also promote the logical subject to first position. Since anaphors tend to be bound by either syntactic or logical subject (a-subjects) in those languages, ARG-ST provides a relevant level of representation for the explanation of binding phenomena.

In the light of *caki*-binding, however, we have already suggested that the distinction between the syntactic and logical subjects is not relevant (see section 5.1.3) because an NP which is not an a-subject can be a possible antecedent of *caki*. This is also the case in the Czech examples in (5.51) where the direct object can bind the anaphor in the complement PP, and also in Balinese as follows:

- (6.11) a. Iang ngedengin I Wayan awakne  
 1SG AV.show I Wayan self  
 'I showed I Wayan<sub>i</sub> himself<sub>i</sub>'
- b. Tiang matakon teken anake ento unduk awakne  
 1SG AV.ask to person that about self  
 'I asked (to) the person<sub>i</sub> about him/herself<sub>i</sub>'

(Manning & Sag 1998a, ex.30)

The examples above are presented by Manning & Sag (1998a) to show the distinction between core and oblique arguments. In (6.11a), a term goal *I Wayan* can bind a term theme *awakne*, while in (6.11b), an oblique goal *anake* can bind an oblique theme *awakne*. Even though Manning & Sag (1998a) use this distinction in arranging arguments to which a-binding applies within ARG-ST, it seems quite unclear how the binding relations in (6.11) can be explained in terms of a-subject. The antecedents in both examples above can be viewed neither as the syntactic subject nor as the logical subject. Therefore, a-binding based on the notion of a-subject fails to explain binding relations in some languages as indicated above.

At least for *caki*, its binding relations can be explained without depending on the notion of a-subjects which are representable in ARG-ST. This, together with the facts about adjuncts in ARG-ST leads to the conclusion that it would be much easier if we claimed that it is not the level of ARG-ST, but it is a level, which we call Binding Structure (BINDING-ST) at which we can determine the binding relations of *caki*. BINDING-ST then should be the level which can not only represent a grammatical specification so that it can pick out the external argument structure, but also incorporate semantics so that arguments can be ordered in terms of the lexical argument hierarchy. Further, as a level of the dependents of the predicate, it should be able to contain adjuncts. In other words, binding structure incorporates the advantages of argument structure in that it can arrange arguments in terms of their semantics. Further, it can be freely a syntactic level not restricted by the

notion of argument structure<sup>6</sup>. The internal structure of binding structure then should look like the following:

- (6.12) The internal structure of Binding Structure  
BINDING-ST < subject, < comps, ...>, adjuncts >

This is essentially the same as the argument structure we have proposed in (5.55). Recall that we have angled brackets around complements to explain the different binding of *caki* between complements and adjuncts. This makes a clear distinction between adjuncts and truly selected dependents, which was one of the important functions of ARG-ST pointed out earlier.

Having established an independent level of representation, BINDING-ST, we note that it has an overlapping representation with DEPS as proposed by Bouma *et al.* (1998), which automatically creates redundancy. We will discuss this problem in the next section.

### 6.3 Redundancy in BINDING LIST and DEPS-LIST

In accounts of extraction before Bouma *et al.* (1998) (Sag 1997, van Noord & Bouma 1994), different types of extraction such as subject, complement and adjunct extraction are treated via different lexical rules. However, Bouma *et al.* (1998) present two objections to the lexical rule based account of extraction. Firstly, they argue that it is inelegant to have separate mechanisms for complement, subject and adjunct extraction. Furthermore, it is descriptively unsound: cross-linguistic evidence suggests that extraction is a more unified phenomenon. In languages that “register” the presence of gaps (for example, in Chamorro where verbal morphology reflects the presence of a gap), subject and adjunct gaps pattern with complement gaps. Secondly, lexical rules are problematic in general, and seem particularly inappropriate for analysing extraction. If extraction is a lexical process, we might expect to see lexical exceptions (e.g., particular lexical items that cannot undergo

---

<sup>6</sup>We will leave the discussion of BINDING-ST being a HEAD feature to section 6.4.

the Complement Extraction Lexical Rule (CELR)). The apparently universal absence of such cases argues against a UDC analysis driven by lexical rules.

In order to overcome such problems originating from using lexical rules, Bouma *et al.* (1998) provide a constraint-based account for extraction. One of the important feature of this account is that they introduce a new DEPENDENCY STRUCTURE (DEPS) which specifies the list of dependents of a lexical head. It takes a value of type *list (synsem)* and keeps track of all the syntactic dependents of a word. For this new feature DEPS, two relevant constraints are introduced:

$$(6.13) \quad \text{ARGUMENT REALIZATION}$$

$$word \Rightarrow \left[ \begin{array}{ll} \text{ARG-ST} & \boxed{1} \\ \text{DEPS} & \boxed{1} \oplus \text{list}(\text{adjuncts}) \end{array} \right]$$

$$(6.14) \quad \text{DEPENDENT REALIZATION}$$

$$word \Rightarrow \left[ \begin{array}{ll} \text{SUBJ} & \langle \boxed{1} \text{ synsem} \rangle \\ \text{COMPS} & \boxed{2} \ominus \text{list}(\text{gap-ss}) \\ \text{DEPS} & \langle \boxed{1} \rangle \oplus \boxed{2} \end{array} \right]$$

Argument Realization in (6.13) says that the SYNSEM values of the core syntactic dependents of a word (i.e., subject, specifier, complements) appear in the DEPS list. In addition to these, the SYNSEM values of zero or more adjuncts can appear at the end of the DEPS list. On the other hand, Dependent Realization in (6.14) states that the first member of the DEPS list is the subject, while the COMPS list contains the rest of the members of DEPS minus any that happened to be of type *gap-ss*. Note that in combination with Argument Realization, this means that adjuncts will appear on COMPS, as long as they are not of type *gap-ss*. (The detailed operation of extraction (e.g., the percolation of NONLOC information) will be illustrated in the next section).

Considering Argument Realization together with Dependent Realization, ARG-ST also needs to contain *gap-ss* just like the DEPS list. Given this, the minimal difference between ARG-ST and DEPS can then be reduced to the fact that ARG-ST does not have the position for adjuncts, while DEPS contains them. Now recall that for *caki*-binding, we have replaced ARG-ST by BINDING-ST which also contains adjuncts. Our BINDING-ST will thus have exactly the same representation as the DEPS list. This is apparently redundant. If there is any reason that the DEPS list cannot serve a purpose which BINDING-ST (or ARG-ST) can, then we may be able to justify the existence of BINDING-ST and DEPS list in parallel. Otherwise, one list should suffice. One last thing to be noted is the matter whether ARG-ST should be retained or not in parallel with DEPS. Even though, for binding, a representation of ARG-ST seems to be largely in parallel with that of *deps*, retaining ARG-ST can be advantageous in capturing syntactic domains other than binding. This will be pointed out when it becomes more relevant in section 8.2. In the next section, we will examine whether DEPS can be an appropriate level in determining *caki*-binding.

## 6.4 Binding in the DEPS List

In abandoning BINDING-ST in favour of DEPS, there are especially two points to be carefully considered: One is whether the arrangement of arguments in terms of the lexical argument hierarchy would be appropriately done in DEPS list, and the other is whether it is possible to make the DEPS list appear on a phrasal level, unlike a lexical sign as proposed by Bouma *et al.* (1998).

First, in Dependent Realization as in (6.14), it has been shown that the elements in DEPS are the APPEND of the SUBJ and COMPS lists (minus *gap-ss*), in that order. Further, Argument Realization as in (6.13) shows that zero or more adjuncts can appear at the end of the DEPS list. Overall, therefore, elements in the DEPS list will be in the order of subject, complements and adjuncts. In fact, this is what we predict for the order of the elements in BINDING-ST as shown in (6.12). The only difference is that among complements the order can be altered in terms of the lexical argument



hierarchy in BINDING-ST. However, this different order among complements will not cause any problem in the DEPS list in accounting for extraction. If any complement is of type *gap-ss*, regardless of its position, it can be identified with the filler by structure-sharing its LOCAL value and properly discharged. Therefore, our internal structure of BINDING-ST can be readily realised in the DEPS list.

Second, in Bouma *et al.* (1998), the DEPS feature (together with ARG-ST) is relevant for words only, and thus does not appear on a phrasal level. However, we would like to examine whether DEPS can be made a HEAD feature as we did for BINDING-ST (or ARG-ST), so that we can account for long-distance binding of *caki* at the level of DEPS. Since the DEPS list is a level of representation for handling extraction, making the DEPS list a HEAD feature should not disturb the extraction mechanism. To see this, it is necessary first to understand the operation of extraction in HPSG.

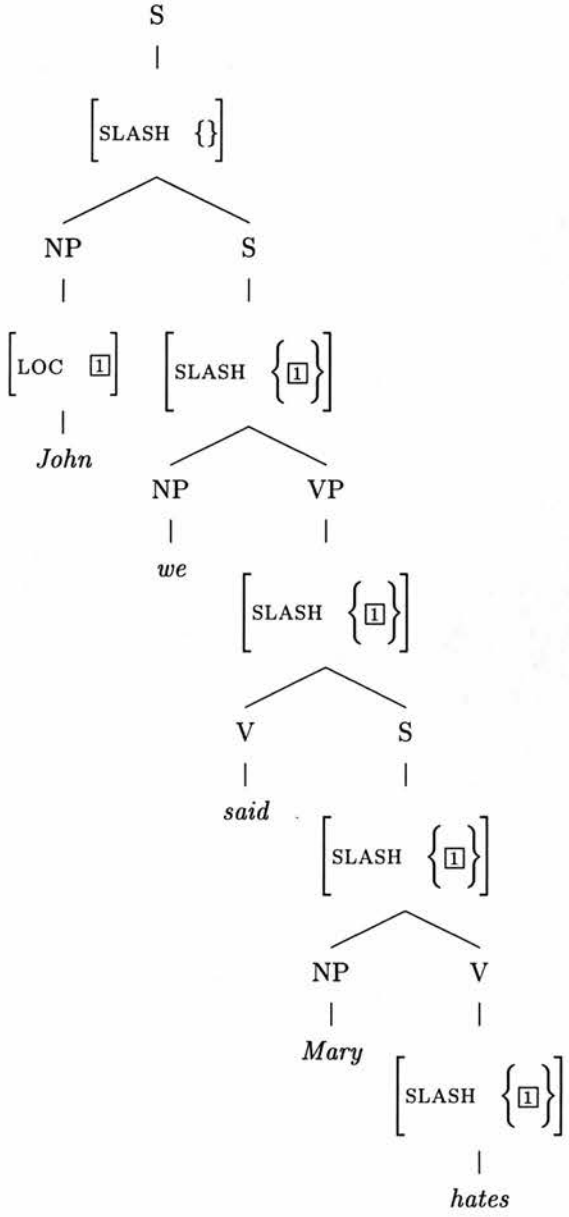
#### 6.4.1 DEPS and Constraints on SLASH

In HPSG, unbounded dependency constructions such as *wh*-questions, topicalisation in English and relative clauses are handled by the Filler-Head Schema. The Filler-Head Schema deals with a phrase where an extracted object (Filler daughter) combines with a clause containing a missing element (*gap*) which is identified with the filler. Furthermore, the information that a phrase contains a *gap* is encoded in its NONLOCAL|SLASH value. To illustrate this point, consider the following example:

(6.15) John, we said Mary hates *e*

In (6.15), *John*, extracted from the object position, is a filler and ‘*e*’ identifies the position of *gap*. (6.15) has the following simplified feature structure:

(6.16)



In (6.16), the missing element is identified at the bottom of the tree and it appears on the SLASH value of the verb (indicated by  $\boxed{1}$ ). The 'slashed' verb combines with the remaining argument and the SLASH value of a head daughter is successively passed up to its mother via the SLASH Inheritance Principle (shown below) up to the point where it is discharged when it is combined with the compatible dislocated element (the filler). For this, there are two constraints in operation as given below:

(6.17) SLASH AMALGAMATION

$$word \rightarrow \left[ \begin{array}{l} LOC \quad \left[ \begin{array}{l} CAT \quad \left[ \begin{array}{l} DEPS \left\langle \left[ SLASH \boxed{1} \right], \dots, \left[ SLASH \boxed{n} \right] \right\rangle \\ BIND \boxed{0} \end{array} \right] \\ SLASH \quad \left( \boxed{1} \cup \dots \cup \boxed{n} \right) - \boxed{0} \end{array} \right] \end{array} \right]$$

With the constraint above, Manning & Sag (1998a) propose that the SLASH value of a lexical item is defined in terms of the SLASH values of its dependents in DEPS<sup>7</sup>. SLASH Amalgamation ensures that if a dependent is slashed then the head which selects it will also be slashed. Also note that DEPS takes a SYNSEM list as its value as follows:

(6.18)  $\left[ \begin{array}{l} DEPS \quad list(synsem) \end{array} \right]$

Given this, the lexical entry for *hates* in (6.16) should look like the following:

(6.19)  $\left[ \begin{array}{l} \text{'John, we said Mary hates } e' \\ \text{'hates'} \\ SUBJ \quad \langle \boxed{1} \rangle \\ COMPS \quad \langle \rangle \\ DEPS \quad \left\langle \left[ \begin{array}{l} \boxed{1} NP_{Mary}, \boxed{2} \left[ \begin{array}{l} gap-ss \\ LOC \quad NP \\ SLASH \quad \boxed{3} \end{array} \right] \end{array} \right\rangle \\ BIND \quad \{ \} \\ SLASH \quad \boxed{3} \end{array} \right]$

By SLASH Amalgamation, the verb *hates* is slashed because its complement is slashed in DEPS. Also, by Dependent Realisation in (6.14), the COMPS list has

<sup>7</sup>The feature BIND is used by the words that introduce *tough*-constructions. These words subcategorise for SLASHed complements but they lexically discharge the dependency so it never gives rise to a head-filler phrase. See section 3.3. in Manning & Sag (1998a) for more detailed illustration.

an empty value as its value is *gap-ss* which is realised only in DEPS. Then the configurational passing of the SLASH feature is associated with the following constraint:

$$(6.20) \quad \text{SLASH INHERITANCE PRINCIPLE (SIP)}$$

$$\text{Head-val-phrase} \longrightarrow \left[ \begin{array}{l} \text{SLASH} \quad \boxed{1} \\ \text{HD-DTR} \quad \left[ \text{SLASH} \quad \boxed{1} \right] \end{array} \right]$$

The SLASH Inheritance is defined as a constraint on *head-val-phrases*, which is the type of phrase involving only head, complement, or subject daughters. Thus, it does not apply to *head-filler* phrases, where a SLASH should be discharged by the phrase combining with the filler. This stops SLASH being passed up (indicated by the empty SLASH value on the head-filler phrase in (6.16)). The *head-filler* phrase has the following constraint<sup>8</sup>:

$$(6.21)$$

$$\text{hd-filler-ph} \longrightarrow \left[ \begin{array}{l} \text{SUBJ} \quad \langle \rangle \\ \text{SLASH} \quad \boxed{2} \uplus \boxed{3} \\ \text{HD-DTR} \quad \left[ \text{SLASH} \quad \boxed{2} \uplus \left\{ \boxed{1} \right\} \right] \\ \text{NON-HD-DTRS} \quad \left\langle \left[ \text{LOC} \quad \boxed{1} \right] \right. \\ \quad \left. \left[ \text{SLASH} \quad \boxed{3} \right] \right\rangle \end{array} \right]$$

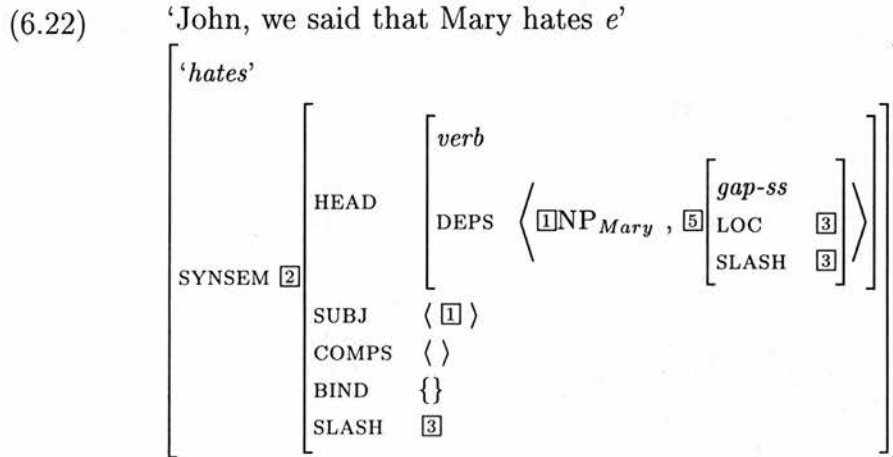
The constraint above simply states that the SLASH value of a head-filler phrase is the SLASH value of the head daughter minus the filler plus the SLASH value of the filler.

With the relevant constraints for the extraction operation in place, let us now move on to examine the effects of making DEPS a HEAD feature.

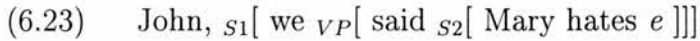
<sup>8</sup>' $\uplus$ ' designates the operation of disjoint set union, which is just like familiar set union except that the disjoint union of two sets with a nonempty intersection is undefined.

### 6.4.2 DEPS as a HEAD Feature?

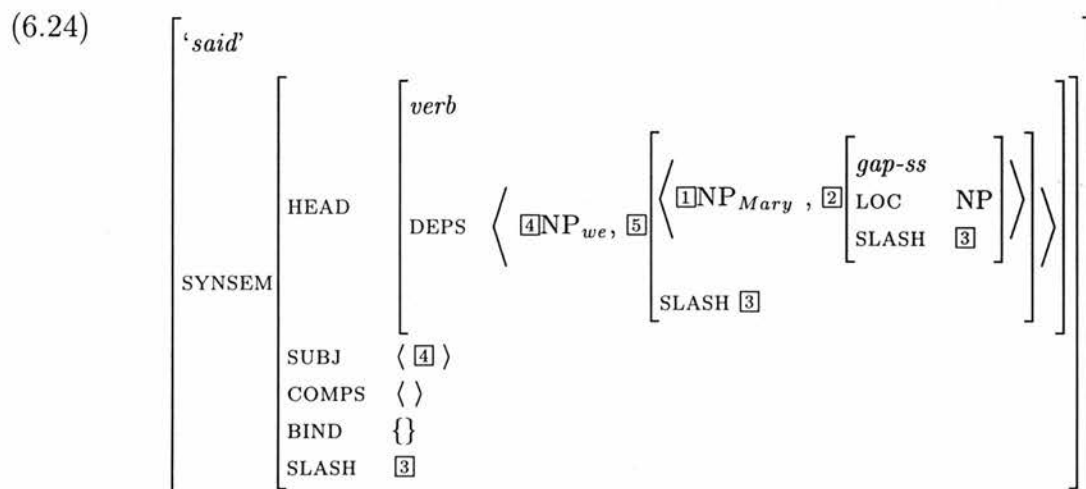
First, including DEPS as one of HEAD features will alter the configuration of the feature structure of (6.19) in the following way:



DEPS as a HEAD feature is passed up and will be available on the phrase level by the Head Feature Principle (HFP). For the sake of illustration, example (6.19) is repeated below:



Since DEPS is a HEAD feature, its value will be represented at the S2 level together with other relevant SYNSEM values of the phrase. When the phrase S2 combines with the higher verb *said*, the SYNSEM of S2 (indicated by  $\boxed{5}$  below) will appear in the DEPS of *said* as the second element as illustrated below:



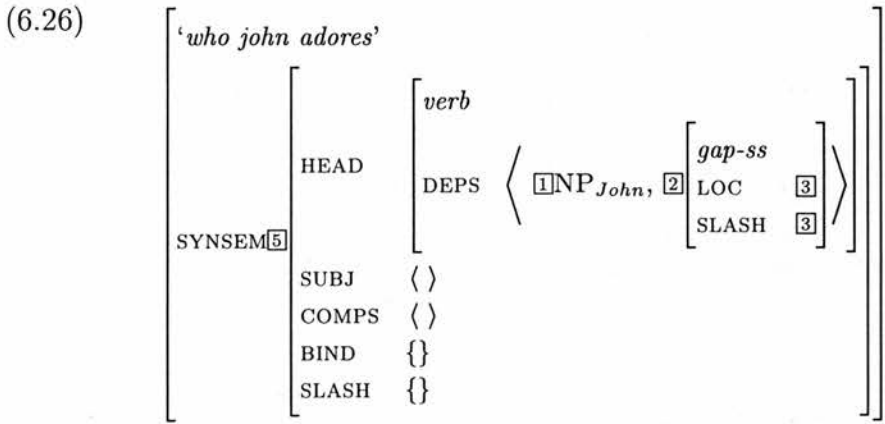
In DEPS above, the matrix NP *We* (SYNSEM $\boxed{4}$ ) comes as the first element, followed by the sentential complement SYNSEM $\boxed{5}$ . This internal presentation of the SYNSEM $\boxed{5}$  is essentially the same as in (6.22), but is simplified to help the illustration by omitting features like SUBJ, COMPS and BIND. Inside the SYNSEM $\boxed{5}$ , the two embedded arguments are represented in the embedded DEPS list: the lower subject ( $\boxed{1}$ ) and the *gap-ss*. The SLASH value ( $\boxed{3}$ ) at the bottom of the feature structure is inherited via the SLASH Inheritance Principle. And this is the representation of DEPS we intended to create in order for long-distance binding of *caki* to be licensed. The representation will not have any problem for extraction as it does not disturb SLASH, satisfying both SLASH Amalgamation and the SLASH Inheritance Principle; a lexical head *hates* will amalgamate SLASH values from DEPS. Once this is done, SLASH will be percolated by the SLASH Inheritance Principle.

Now, let us look at a slightly different construction of extraction such as the following where the filler is available in the middle of the sentence (i.e., not at the top as in (6.16)).

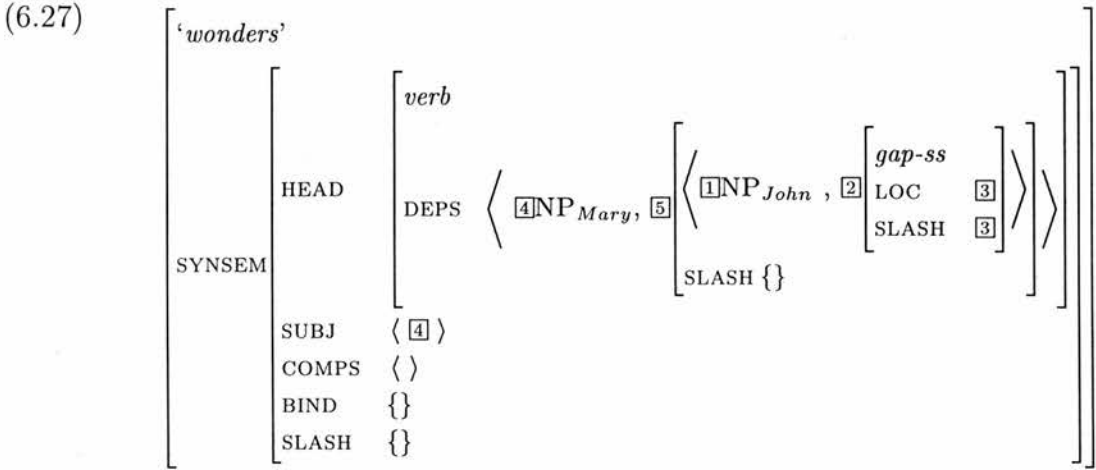
(6.25) Mary  $_{VP}$ [ wonders  $_S$ [ who John adores  $e$  ]]

In the lexical entry of *adores*, the complement is realised as *gap-ss* in DEPS, whose SLASH value will be added in the NONLOC|SLASH value by SLASH Amalgamation.

And the SLASH value will be passed up by the SLASH Inheritance Principle in the same way as in (6.23). The difference is that in (6.25), the SLASH is discharged before being passed up to the higher clause by being combined with the filler (*who*). The feature structure on the filler-head phrase *who John adores* is shown below:



Note that in (6.26), the SLASH value is empty following the constraint on the filler-head-phrase as in (6.21). When the phrase in (6.26) is combined with the higher verb *wonder*, its SYNSEM value  $\boxed{5}$  will be represented in the DEPS of *wonder*. The feature structure of *wonder* is shown below:



In (6.27), we again have the DEPS representation where the arguments in the higher clause are visible to the arguments of the complement clause. As for the extraction mechanism, on the other hand, the presentation in (6.27) will not violate SLASH



Amalgamation because none of the elements in DEPS is slashed ( $\boxed{4}$  is not slashed and the SLASH of the SYNSEM $\boxed{5}$  is empty).

Therefore, making DEPS a HEAD feature would not violate or require any modifications of any principal constraints on extraction. This means that we can explain the binding relations of *caki* at the level of DEPS. Since we have now just one level of representation for both binding and extraction, we can circumvent some redundancy in the theory.

## 6.5 Conclusion

So far, we have established the argument that DEPS is the level at which *caki*-binding should be determined. We started by examining the uncomfortable notion of traditional argument structure. We concluded that the notion of argument structure is too limited to be the relevant level for explaining *caki*-binding, for which we have provided, in essence, three arguments, based on syntactic specifications, the involvement of adjuncts, and the representation of arguments outside the local domain.

In the beginning of this chapter, in providing a unified account of *caki*-binding, we concluded that ARG-ST is not an appropriate level either. This conclusion is partly based on the three arguments above, but it is also due to the fact that a-subjects are not the best concept to explain *caki*-binding. We have also shown that the notion of a-subject is not necessarily a determining factor in predicting the correct antecedent in other languages such as Balinese, Japanese and Czech.

As both argument structure and ARG-ST fail to be the relevant level for *caki*-binding, we have provisionally provided an independent level, namely, BINDING-ST as a HEAD feature. However, this level has been shown to be almost identical to DEPS, a level for extraction, except for the fact that DEPS is a lexical representation. Nonetheless, having two identical levels in parallel for different syntactic phenomenon is clearly seems unnecessary. In order to surmount this redundancy in the theory, we have

reduced two levels to one in DEPS. The only question in doing so was the possibility of making DEPS a HEAD feature just like BINDING-ST. We have witnessed that it does not affect any extraction mechanism. Thus we proposed that binding and extraction can be now explained in one level of presentation, DEPS.

Therefore, DEPS can now appear not only at the lexical level, but also in a phrasal level, which makes both local and long-distance binding of *caki* explicable in one level of representation. Given this, we can now state our final version of the principle for *caki*-binding with DEPS as the required locus as follows:

(6.28) Principle for *Caki* (Final version)

*Caki* must be bound by a preceding element in the same, or higher, level of DEPS.

This, we argue, is the core principle that has to be satisfied by all instances of *caki*-binding. In the following chapter, we will explore other cases of *caki*-binding, namely, topic-binding. Due to the discourse function of topics and the cases where *caki*-binding is licensed without an overt antecedent, topic-binding has attracted various discourse approaches (e.g., Sells (1987), Iida (1996) and Huang (1984)). Our treatment in the next chapter will not involve any extra principle. Rather, we will show that if proper attention is paid to the way that topic constructions in Korean are derived, topic-binding can be explained consistently by our principle without any unacceptable extra theoretical costs.

# Chapter 7

## Other Components of LDA

### 7.1 Introduction

In section 3.2.1 of chapter 3, we have seen that the relevance of contextual factors in accounting for long-distance anaphora has been repeatedly stressed in the literature in various forms (Perspective (Iida 1996), Logophoricity (Sells 1987) or Point of view (Kuroda 1973, Banfield 1982, Zribi-Hertz 1989)). We also reviewed syntactic conditions such as the subjecthood condition which states that the Japanese Long-Distance Anaphor *zibun* allows a subject antecedent, but not a nonsubject one (Kuroda 1965, Kuno 1973). It is not, however, always clear whether these conditions are proposed in order to replace the contextual ones or to be taken in conjunction with them, as we have witnessed from various binding data where either conjunction or disjunction of two groups of theories (syntactic and contextual) cannot be extended to account for Korean long-distance anaphora.

In this chapter, we will focus on the binding data that have been generally considered to be the domain of contextual rather than syntactic conditions. Typically, such data show *caki*-binding without any overt antecedent within the sentence domain. The relevant data in question also include the cases of non-subject binding such as topic-binding. By reconsidering data in Korean, usually cited in favour

of the contextual/discourse approach to long-distance anaphora, I will propose an account of long-distance anaphora in more structural and formal terms rather than depending on the contextual factors. The structural account we propose in this section does not involve an extra constraint on binding principles. Rather we will account for the binding relations in terms of interaction with already existing relevant constructions.

The first section will deal with cases where *caki* occurs without an overt antecedent within the sentence. We will argue however that this is only an epiphenomenon. We will show that the underlying structure is such that an antecedent is always available. The key property of Korean that we will use is what is known as the multiple subject, or the double nominative construction.

The second section will look at data involving different types of morpho-syntactic topicalisation in relation to different topic-binding behaviours. We will argue that the observed different topic-binding relations do not need any extra binding constraints or contextual explanations, but that they follow naturally from the nature of different topic constructions. Thus, such binding data are explained in terms of simple structural principles. At the end of the section, we will address some theory internal problems within HPSG in providing a complete explanation for topic-binding in terms of DEPS and will examine various possible explanations.

## 7.2 Caki Without an Overt Antecedent

### 7.2.1 The Major Factors

It has been generally observed (e.g., Li & Thompson (1976)) that East-Asian languages such as Korean, Japanese and Chinese present a certain cluster of common features such as the following:

1. Topic-orientedness
2. Double nominative constructions

### 3. Long-distance Anaphora

Firstly, one of the long established characteristics of languages such as Chinese, Japanese, and Korean is their context dependence; or to put it in Li & Thompson's (1976) terminology, their "topic-orientedness". Unlike other pro-drop languages (e.g., those of the Romance family) empty pronouns in East Asian languages are licensed not by strong agreement such as inflected verbs but by their ability to be identified via strong contextual or discourse features<sup>1</sup>. Their second common feature, double nominative constructions, represents their ability to generate two subject positions. These languages also consistently exhibit long-distance anaphoric patterns (*ziji* in Chinese, *zibun* in Japanese, and *caki* in Korean). Even though each of these common features has been a widely discussed issue (for examples, Li & Thompson (1976), Yoon (1987), Doron & Heycock (1999), Chang (1995) and Moon (1994)) their interaction has not been thoroughly investigated. In this section, we will highlight the effects of the second feature in the account of the third one. In this way, it will become clear that cases of *caki*-binding without overt antecedents do not need to rely on an independent contextual account.

#### 7.2.2 The Basic Facts and Previous Accounts

In this subsection, we will observe the basic long-distance anaphoric phenomena that are central to the accounts which are highly dependent on contextual factors. The first set of data which has been given as evidence for the discourse based account of long-distance anaphora involves cases such as the following:

---

<sup>1</sup>This is relevant to topic-binding which we will deal with in the next section.

- (7.1) A. *Mary<sub>i</sub>-ka ku pati-e kass-ni anim tarun salam-i*  
*taysin kass-ni?*  
 Mary-NOM the party-to went-Q or other person-NOM  
 instead go-Q  
 'Is it *Mary<sub>i</sub>* who went to the party or somebody else instead?'
- B. Ani, ***caki<sub>i</sub>-ka kasse***  
 No, ***self<sub>i</sub>-NOM*** went  
 'No, ***self<sub>i</sub>*** went'

(7.1) illustrates an exchange between two speakers A and B. In A's utterance, *Mary* is mentioned and remains a prominent topic throughout the exchange. In B's utterance, *caki* occurs without any overtly expressed antecedent in its own sentence. As the indexing indicates, *caki* is anteceded by *Mary*. This sort of example has been cited in most of the literature as a case of discourse-binding (Huang 1984, Ueda 1984). In fact, *caki* in (7.1) looks as if it was bound by a discourse antecedent. In other words, with no possible antecedent available in its own sentence, it looks for its antecedent in the previous discourse. In this case, *caki* is bound by the prominent topic *Mary* in the discourse which is introduced by the subject in the initial utterance (A).

The reason that discourse binding applies to these examples is that *topic* is viewed as a discourse function interpreted as what is being talked about or what is presupposed or understood by *the speaker*. This definition of topic is well suited for the notion of *Perspective* or *Point of view* used in the discourse based accounts of long-distance anaphora. However, this type of account describes rather than explains the data. One would like to have a formal account of how the prominent topic or the level of prominence of any given topic is formally represented in order to disambiguate and decide amongst several possible topic antecedents. This is particularly obvious in the following situation:

- (7.2) A. Mary-ka pati-ey ka-ss-ni? anim John-i ka-ss-ni?  
 Mary-NOM party-to go-PAST-Q or John-NOM go-PAST-Q  
 'Did Mary go to the party or is it John who went to the party?'
- B. \*Caki-ka ka-ss-e  
 self-NOM go-PAST-DC  
 'Self went'

(7.2) has the same structure as (7.1) except that in (7.2), *John* replaces *tarun salam* (*other person*). When there are two equally prominent entities in the discourse such as *Mary* and *John* in (7.2), *caki* in B's reply is not licensed. If *caki* is indeed bound by a prominent entity in discourse according to discourse approaches, then *caki*-binding should be licensed having either *Mary* or *John* as the antecedent. Unless such accounts can provide a good explanation for the ungrammaticality of B in (7.2), it is hard to see how the binding relation is completely dependent on the discourse. Furthermore, even if discourse approaches can deal with instances like (7.2), a superior account would be one which can be concretely formalised so that a legitimate antecedent is clearly visible in relation to *caki* at some formal level, rather than leaving the prediction of the correct antecedent to the discourse context. Given this, in attempting to move the explanation of the phenomena into the realm of a structural account, we will propose that a sentence like (7.1B) is structurally more complex than it seems and that there is an element akin to a 'topic' structurally present although not necessarily syntactically identified as such. More specifically we will argue that such examples are instantiations of the double nominative construction.

### 7.2.3 The Double Nominative Construction

It is well known that Korean and Japanese allow two nominative marked NPs to occur with a one place predicate. Consider the following:

- (7.3) a. John-i/un ton-i issta  
 John-NOM/TOP money-NOM exist  
 'John has money'



- b. Mary-ka/nun    meri-ka    norahta  
 Mary-NOM/TOP hair-NOM yellow is  
 'Mary's hair is yellow'
- c. LA-ka/nun    hankukin-i    manhta  
 LA-NOM/TOP Korean-NOM many  
 'LA has many Koreans'

The precise nature and function of each of the nominative NPs has long been a matter of debate. For our purposes, it is sufficient to observe that the outer nominative has subject-like properties<sup>2</sup>. The outer nominative can be considered an argument of the complex predicate created by the combination of the inner nominative and the core predicate, very much like the topic-comment relation holding between a sentential topic and the rest of the sentence (Heycock & Lee 1990). Another interesting feature of this construction is that it displays strictly rigid word order which contrasts sharply with the freedom of word order which is observed in Korean.

- (7.4) a. John-i    pal-i    kilta  
 John-NOM arms-NOM are long  
 'John's arms are long'
- b. \*Pal-i    John-i    kilta  
 arms-NOM John-NOM are long

Furthermore, it is even more interesting to notice that it is not only word order permutations such as scrambling, as above, but also any other operation that would disturb that order, such as topicalisation, that are disallowed:

- (7.5) a. Mary-ka    aiyn-i    sihem-ey hapkyekhayssta  
 Mary-NOM lover-NOM exam-in passed  
 'Mary's lover passed the exam'
- b. \*Aiyn-nun Mary-ka    sihem-ey hapkyekhayssta  
 lover-TOP Mary-NOM exam-in passed

---

<sup>2</sup>For this point, see for example the extensive discussion in Doron & Heycock (1999).

It is however possible to *-nun'* mark (topicalise) the outer nominative nominal as shown in (7.3).

On the interpretive side, the semantic relations that must hold between the first and second nominatives are more restricted than the general 'aboutness' relation that needs to hold between the sentential topic and the comment part of the sentence. The relations in question are alienable possession (7.3a), inalienable possession (7.3b), part-whole (7.3c) and identity. With this description in mind, let us now turn to the behaviour of seemingly antecedentless *caki*.

#### 7.2.4 *Caki*-Binding in Double Nominative Constructions

The first observation concerning the behaviour of *caki* in double nominative constructions is that when it occurs, it can only occupy the second nominative position:

- (7.6) a. John-i caki-ka tampay-lul kkulhessta  
John-NOM caki-NOM cigarette-ACC cut  
'John himself stopped smoking'
- b. \*Caki-ka John-i tampay-lul kkulhessta  
caki-NOM John-NOM cigarette-ACC cut

Observe next that the understood antecedent of *caki* in cases where it is not explicitly present has to bear a special type of semantic relation to the subject of the sentence where *caki* occurs. Interestingly enough, this type of relation is the same as the relation between the outer nominative and the subject of the rest of the sentence in double nominative constructions as shown below.

- (7.7) A. Nay-ka ne-ekey cwu-n chayk-ul poassni? Bill-i cachkoi-sstentey?  
 I-NOM you-DAT give-REL book-ACC saw? Bill-NOM was looking for  
 ‘Did you see the book I gave you yesterday? Bill was looking for  
 it’.
- B. (Bill<sub>i</sub>-i) Apeci-ka ku chayk-ul caki<sub>i</sub>-ekey onul cwuessnunte  
 Bill-NOM father-NOM the book-ACC self-DAT today gave  
 (Bill<sub>i</sub>’s) father gave the book to caki<sub>i</sub> today’

Finally, observe that in the cases that interest us here, it is impossible to topicalise (‘-*nun*’ mark) *caki* itself.

- (7.8) B’. \*Ani, caki<sub>i</sub>-**nun** kasse  
 no, self-TOP went

In a discourse-based account, this example will not differ from B in (7.1) and there is no way of explaining the ungrammaticality of (7.8). This cluster of facts can be explained if we assume that antecedentless *caki* occurs in what is in reality a double nominative construction with an unrealised first nominative. This simple account would explain first why *caki* takes in these cases the kind of antecedents that it does (simply because they have to fulfill the conditions of appearance as the first nominative). And second, why *caki* cannot be topic marked which has long been a mysterious fact (Kim 1993). The reason for this is to be found in the word order rigidity observed in the double nominative construction. Crucially, observe that there is no general prohibition against topicalisation of *caki* itself as the following example shows:

- (7.9) Caki<sub>i</sub>-nun Mary<sub>i</sub>-ka ceil silhehanta  
 self-TOP Mary-NOM the most hates  
 ‘As for self<sub>i</sub>, Mary<sub>i</sub> hates the most’

We will return to this type of example in section 7.3.

Despite the fact that this word order permutation is string-vacuous, it would still forbid the establishment of the proper semantic relation. Therefore, the representation of B's reply in (7.1) is as follows:

- (7.10) B. Ani, [ $e_i$ ]<sub>nom</sub> **caki**<sub>i</sub>-ka kasse  
 No, **self**<sub>i</sub>-NOM went  
 'No, [ $e_i$ ]<sub>nom</sub> **self**<sub>i</sub> went'

Therefore, it appears that under this simple framework of assumptions, there is no need to appeal to contextual factors to account for the antecedentless *caki*. Moreover, it seems that the purely syntactic account is more desirable as it also explains the behaviour of *caki* in this particular context to a fuller extent.

### 7.2.5 Topic-Orientedness and the Expression of Topics

The last question that remains is what allows the outer nominative not to be expressed as in (7.10). The answer is related to the general notion of topic-orientedness of Korean. In highly discourse oriented languages like Korean, topic phrases or phrases that function like topics are very often suppressed in sentences subsequent to the first occurrence of the topic in the discourse. Consider the following examples:

- (7.11) A. Yong-i nuwku-hako ssawa-ss-ta-ko?  
 Yong-NOM who-with fight-PST-DC-Q  
 'Who did you say Yong fought with?'  
 B. Heyn-hako-yo  
 Heyn-with-PO  
 'With Hyen'

(Chang 1995, p.200, ex.57(c-d))

Given the topic *Yong* in the initial dialogue A in (7.11), B's reply is elliptical: the topic elements are not repeated. And once the topic is introduced in an unstressed form, it is suppressed in the subsequent utterance or realised in a pronominal form. Otherwise, the same topic is repeated as shown below:

(7.12) Mia-nun ko-sam-i tway-yo.

Mia-TOP high-three become-SE

Nay-nyen-ey (kay-nun/Mia-nun) tayhak-ey ka-yo.

next-year-at she-TOP/Mia-TOP college-to go-SE

Kulayse (kay-nun/Mia-nun) Yelsimhi Kongpwuha-ko iss-e-yo

So she-TOP/Mia-TOP diligently study-ing iss-PO-DC

'Mia becomes a high school 3rd grader. She goes to college next year. So she is studying hard'

(Chang 1995, p.200, ex.58)

In the above discourse setting, *Mia* is the topic in the first sentence. It continues to be the topic and is realised by zero or the pronominal *kay-nun* (*child/she*). More importantly, the only case where the topic is obligatorily overt is when there exists some ambiguity (when there is more than one prominent topic in the given context), when the topic has been just changed from the previous one or when there is a need to reintroduce the topic for clarification. Let us imagine the same context as in (7.2), but with two individuals *Bill* and *John* who are equally prominent in the context. In this case, the reply of B should express the topic explicitly to clarify which one he/she is talking about. In this situation, the use of pronouns in the reply is simply disallowed, shown as below. This is because the pronoun itself would not be able to disambiguate between *Bill* and *John* as a prominent topic.

(7.13)

(B) \***Ku**-ka ka-ss-e

**he**-NOM go-PAST-DC

'**He** went'

Thus, this suggests that the failure of the discourse pronominal binding is attributed to discourse. The same reasoning applies to *caki*-binding. We have already seen in (7.2) that *caki* is not allowed when there is ambiguity in the prominent topic and the topic or the outer nominative in (7.13) is not explicitly expressed. However, the failure to reintroduce the prominent topic explicitly affects *caki*-binding differently

from cases of pronominal-binding. When the topic is suppressed when it needs to be expressed explicitly, thus does not satisfy the requirement that *caki* should be bound sentence internally. That is, without such an explicit topic, *caki*-binding is not properly licensed, thus, ungrammatical.

Furthermore, analysing the topic as a pronominal element of a particular kind disallows split antecedents<sup>3</sup>. In this way, the chain of an overt or covert topic is systematic. Understanding this chain of topics, which is structurally accommodated by double nominative constructions, provides an elegant way of predicting the observed binding patterns. Before we move on to the HPSG implementation of the approach outlined earlier, one terminological cautionary note is in order. We recognise that some confusion may arise from our use of topic. In what precedes, topic is used in order to refer to a set of discourse functions and behaviour regardless of its precise syntactic realization. More specifically, it is not necessarily the case that only '-*nun*' marked elements can function as a topic. As shown earlier, the outer nominative marked elements in double nominative constructions function as topics although they are not formally identified as such. We will return to cases where there is a '-*nun*' marked element and a different subject in the sentence in section 7.3.

### 7.2.6 An HPSG Approach

In the last section, we saw that the outer nominal can take the topic marker *-nun* instead of the nominative marker. It may be argued that the topic marked nominal should not be considered a selected argument. For example, Yoon (1987) argues that the DNC should be analysed as a gapless topic/focus construction such that the outer nominal may be licensed by the same principles that license *as for* phrases and other parentheticals in English. In fact, the analysis of the outer nominal as a pure topic is not uncontroversial. There is, however, plenty of evidence that the outer nominal has legitimate argument status. For example, Doron & Heycock

---

<sup>3</sup>Perhaps the pronominal in question is akin to PRO, which also needs to be bound and cannot take split antecedents. However, we will leave that issue open for the time being.

(1999) argue that the outer nominative nominal has in all cases the properties normally associated with subjects in languages like Japanese and Semitic. One piece of evidence comes from the fact that an anaphor such as *zibun* can be bound by the outer nominative nominal as in the following:

- (7.14) Sono hito<sub>i</sub>-ga kidomo-ga zibun<sub>i</sub>-yori atama-ga ii (koto)  
 that person-NOM child-NOM self-than head-NOM good (fact)  
 ‘That person<sub>i</sub> [is such that his/her] child is more intelligent than him/her’

(Doron & Heycock 1999, ex. 16)

This indicates that the outer nominative nominal is in an argument position, given the general assumption that binding of anaphors is only possible from argument positions. Also, the fact that *zibun* is a subject-oriented anaphor shows that the outer nominative nominal behaves like a subject. As the DNC in relation to *caki*-binding is similar to Japanese, our view is in line with Doron & Heycock (1999) and the outer nominal is treated as a subject member of DEPS. The fact that it occurs in the leftmost position identifies it as the most prominent element, as is common in other languages. This explains simply why the nominative case of the outer nominative nominal can freely alternate with the topic marker as shown in (7.3). For example, the Korean sentence corresponding to (7.14) is shown below:

- (7.15) Ku saram-**i/un** ai-ka caki-pota meri-ka cohta  
 the person-NOM/TOP child-NOM self-than head-NOM good

When the outer nominal is topic-marked as above, because it is a topic phrase, it is normally treated as an adjunct phrase or a dislocated phrase (Yoon 1987, Moon 1994). One may object to the assumption that it should be contained in the argument structure, given the hypothesis of Manning (1996) and Manning & Sag (1998a) that binding principles are stated at a level of syntacticized argument structure. As we explain *caki*-binding at the level of DEPS list which also contains adjuncts, including the topic may not be a problem. Given this, though, there is some evidence that the topic phrase in question should not be treated either



as a pure adjunct phrase or as a dislocated phrase. Firstly, our argument that the topic phrase is an argument comes from the fact that the marker *nun* is not a postposition, but rather behaves like a case marker. This assumption is based on the phenomenon of case stacking in Korean. In Korean, one NP can bear two post nominal markers, but the combination is not entirely free. (Gerdts 1985, Yoon & Yoon 1990). There are two types of postnominal elements in Korean: Case markers and Postpositions. The case markers include *-i/ka* for nominative, *(l)ul* for accusative and *-uy* for genitive. Postpositions are like *(uro)pwute* (from), *-ey* (in/at), etc.. The generalisation is that a postposition can be inside another element, but a case marker cannot. That is, case-markers are not allowed to combine with each other or with a following postposition, whereas postpositions are able to combine with either a case marker or a postposition. If the topic marker *-nun* is truly a postposition, it should be able to combine with other postpositions as well as a case marker. However, while it combines freely with postpositions, it resists combining with a case marker as follows<sup>4</sup>:

- (7.16) a. Hakkoy-ka-\*nun ku cip-uroputeo-nun melta  
 school-NOM-TOP the house-from-FOC far  
 'The school is far from the house'
- b. Untong-ul-\*nun John-i kyeul-ey-nun haci anhnunta  
 exercise-ACC-TOP John-NOM winter-in-FOC do not  
 'John does not do exercise in winter'
- c. Seoul-lopute-nun amwu sosik-i-\*nun epsta  
 Seoul-from-TOP any news-NOM-TOP not  
 'There isn't any news from Seoul'

From the above, the topic marker cannot be combined with the nominative case (7.16a), or with accusative case (7.16b), but it can still co-occur with the postposition *from* in (7.16c). This shows that the topic marker behaves just like a case marker but not like a postposition.

---

<sup>4</sup>The *-nun* marks either topic or focus. One of the differences between topic marker and focus marker is position: a topic phrase occurs in the sentence initial position whereas a focus remains in its base position

A second piece of evidence for the argument status of topic comes precisely from its behaviour as a binder. A comparison of the behaviour of anaphor and pronominal in DNCs shows that the topic phrase in such environments seems to behave like a nominal locally subcategorised by a head. Consider the following examples:

- (7.17) a. John<sub>i</sub>-un caki<sub>i</sub>-ka kassta  
 John-TOP self-NOM went  
 'As for John<sub>i</sub>, (he) himself<sub>i</sub> went'
- b. \*John<sub>i</sub>-un ku<sub>i</sub>-ka kassta  
 John-TOP he-NOM went  
 '\*As for John<sub>i</sub>, he<sub>i</sub> went'

In (7.17a), *caki* is bound by the topic phrase. As the anaphor should be bound by a nominal in an argument position either locally or at a distance (for long-distance anaphora), (7.17a) shows that the topic phrase is actually in an argument position. Furthermore, replacing *caki* in the sentences above by *cakicasin*, a local anaphor, yields perfectly grammatical results in (7.18), which shows that (7.17a) is an instance of local, not long-distance anaphora.

- (7.18) John<sub>i</sub>-un cakicasin<sub>i</sub>-i kassta  
 John-TOP selfself-NOM went  
 'As for John<sub>i</sub>, (he) himself<sub>i</sub> went'

Given this, (7.17b) demonstrates that the topic phrase is actually within the local domain; the pronoun is not bound by the topic phrase and this is because it violates principle B. These two points together (the topic phrase being an argument and locally subcategorised) support the argument that the topic phrase in DNCs should be considered an argument.

Returning to the normal double nominative construction, I propose the constraint on this construction as follows<sup>5</sup>:

---

<sup>5</sup>The feature topic we adopt here is the one referred to in Information Packaging (Vallduví 1992). The feature structure of the CONTEXT value is schematized as follows:

(7.19)

## (7.20) DOUBLE NOMINATIVE CONSTRUCTION LEXICAL RULE

$$\left[ \begin{array}{l} \text{HEAD} \left[ \begin{array}{l} \textit{verb} \\ \text{DEPS} \langle \textcircled{3}\text{NP}_{\textcircled{4}} \rangle \oplus (\langle \textcircled{5} \rangle) \end{array} \right] \\ \text{VAL} \left[ \text{SUBJ} \langle \textcircled{3}\text{NP}_{\textcircled{4}} \rangle \right] \end{array} \right] \Rightarrow \left[ \begin{array}{l} \text{HEAD} \left[ \begin{array}{l} \textit{verb} \\ \text{DEPS} \langle \textcircled{1}\text{NP}_{\textcircled{2}}, \textcircled{3}\text{NP}_{\textcircled{4}} \oplus \textcircled{5} \rangle \end{array} \right] \\ \text{VAL} \left[ \text{SUBJ} \langle \textcircled{1}\text{NP}_{\textcircled{2}}, \textcircled{3}\text{NP}_{\textcircled{4}} \rangle \right] \\ \text{CONT} \left[ \begin{array}{l} \dots \\ \text{R} \langle \text{NP}_{\textcircled{2}}, \text{NP}_{\textcircled{4}} \rangle \\ \dots \end{array} \right] \\ \text{CONX} \left[ \text{TOPIC} \textcircled{6}\text{NP}_{\textcircled{2}} \right] \end{array} \right]$$

As for the basic mechanism of this lexical rule, the input of the rule is a lexical entry with a single subject (indicated by  $\textcircled{3}$ ), and the output contains two subjects ( $\textcircled{1}$  and  $\textcircled{3}$ ) resulting in double nominative constructions. Also, the DEPS list of the input shows that it takes one subject as an argument and this is followed by a list of other arguments ( $\textcircled{5}$ ) if any, whereas the DEPS of the output has two subjects. To license the two nominative nominals as subjects, two non syntactic conditions have to be met. Firstly, the two subjects should hold certain semantic relations such as inalienable possession, relation possession, identity etc, which we will refer to as

$$\left[ \text{CONTEXT} \left[ \text{INFO-STR} \left[ \begin{array}{l} \text{FOCUS} \\ \text{GROUND} \left[ \begin{array}{l} \text{LINK} \\ \text{TAIL} \end{array} \right] \end{array} \right] \right] \right]$$

Vallduvì's (1992, 1995) information structure consists of three primitives, i.e., FOCUS, LINK or TOPIC, and TAIL, that encode the information status of each element of the sentence. He first partitions a sentence into two parts, FOCUS and GROUND. Ground is the part that anchors the sentence to the previous discourse or the hearer's 'mental knowledge' whereas focus is the 'informative' part that makes some contribution to the discourse. Then ground is further divided into LINK and TAIL. Link can in more traditional terms be interpreted as topic (Gundel 1987, Reinhart 1982). Tail can be understood as 'other' given information or the rest of the ground information which is less conspicuous in the sentence. What we refer to as topic here is what is meant by LINK in Vallduvì and that is where the topic information should be stored. For the lexical rules we will use throughout this chapter, we will omit the detailed path up to the feature TOPIC in CONX.

R-relations. In the output of the lexical rule, this is indicated in the CONT feature, taking an R relation feature as one of its values. The R feature also should take a list of two members which are coindexed with the two first elements in the DEPS (that is, the subjects). At the same time, the CONX feature indicates the prominent element in the discourse, that is, what the topic is in the context. The CONX has the TOPIC feature whose value is instantiated with a prominent element and explicitly with *-nun* marking. Further, the TOPIC value has to be coindexed with the first element in DEPS. This element should be coindexed with the left element in the R relation and at the same time, it should also be coindexed with the first element in the DEPS. With these two features (CONT and CONX) appropriately satisfied, the two nominals are placed as the first elements in the DEPS list.

More importantly, however, it is only when the topic value is explicitly provided that the outer nominal is allowed to be covert. The unrealised outer nominal can be understood as an empty pronominal which can be recovered through a given context in languages that freely allow pro-drop and lack a rich inflectional system like Korean<sup>6</sup>. Given this, therefore, the ungrammaticality of (7.17) below is due to the fact that the outer nominal is suppressed despite the lack of an explicit topic.

- (7.21) \*Caki-ka kasse  
 Self-NOM went  
 ‘\*Self went’

---

<sup>6</sup>This, however, should not be understood as a missing element or a *gap*. In recent developments in HPSG (Sag 1997), the type of synsem is divided into two subtypes: canonical synsem (*canonical-ss*) and non canonical synsem (*noncanonical-ss*). The former is associated with all overt expressions and the latter with non-overt expressions. The *noncanonical-ss* has again two subtypes: *pro-ss* and *gap-ss*. The *pro-ss* type represents all the unexpressed objects, standardly referred to as a little *pro* in pro-drop languages. The *pro* we use in this section indicates the *pro-ss* type. As the other kind of *noncanonical-ss*, *gap-ss* is associated with extracted elements in unbounded dependency constructions. We will discuss this type in more detail in section 7.3.3.

$$(7.22) \quad \left[ \begin{array}{l} \text{HEAD} \quad \left[ \begin{array}{l} \textit{verb} \\ \text{DEPS} \quad \langle \text{pro}, \text{NP}_{\text{ref}} \rangle \end{array} \right] \\ * \text{VAL} \quad [\text{SUBJ} \langle \text{NP}_{\text{ref}} \rangle ] \\ \text{CONT} \quad [\text{R} \langle [], \text{NP}_{\text{ref}} \rangle ] \\ \text{CONX} \quad [\text{TOPIC} [] ] \end{array} \right]$$

Furthermore, as the outer nominal is covert there is no way of ensuring that the DNC-relevant relations (R-relations) are properly satisfied.

Returning to (7.17), observe that it becomes fully grammatical when a context where a topic can be identified is supplied, as is also the case for (7.10) whose feature structure is illustrated in (7.23):

$$(7.23) \quad \left[ \begin{array}{l} \text{HEAD} \quad \left[ \begin{array}{l} \textit{verb} \\ \text{DEPS} \quad \langle \text{pro}_{\text{ref}}, \text{NP}_{\text{ref}} \oplus \text{NP} \rangle \end{array} \right] \\ \text{VAL} \quad [\text{SUBJ} \langle \text{NP}_{\text{ref}} \rangle ] \\ \text{CONT} \quad [\text{R} \langle \text{NP}_{\text{ref}}, \text{NP}_{\text{ref}} \rangle ] \\ \text{CONX} \quad [\text{TOPIC} \text{NP}_{\text{ref}} ] \end{array} \right]$$

In this case, even though the outer nominal is unrealised, there is a contextual back-up which allows the recovery of the content of the missing element as indicated in [1]. With the outer nominal properly identified, the CONT attribute indicates that the two nominals are in a proper R-relation. Furthermore, on the basis of the DNC lexical rule, the example below is straightforwardly explained.

$$(7.24) \quad *John_i\text{-un} \text{ Bill-i} \quad \textit{ku} \textit{ chayk-ul} \textit{ caki}_i\text{-ekey} \textit{ cwuessta}$$

John-TOP Bill-NOM the book-ACC self-DAT gave  
 ‘\*As for John<sub>i</sub>, Bill gave the book to self<sub>i</sub>’

In the above, the topic maker is licensed only from DNC constructions and the first two nominals *John* and *Bill* are not in any of the R-relations. Therefore, DNC is not licensed, let alone any *caki*-binding.

In conclusion, we brought together a set of seemingly unrelated properties of a particular set of languages under a special mode of interaction of the syntactic, semantic and pragmatic components of linguistic theory and showed that this set of properties, instead of being a mere curious and interesting set of “areal features”, in fact represent a tightly knit network and one of the best (perhaps the optimal) solution to the long-distance anaphora question.

## 7.3 Topicalisation and *Caki*-Binding

In the last section, we have seen that a topic can be a possible binder of *caki* when it appears in double nominative constructions. In this section, we will look at further cases of topic-binding in constructions resulting from other topicalising mechanisms such as scrambling and extraction.

### 7.3.1 Topic and Word-order

As we briefly saw in the last section, in Korean topichood is encoded by morphology with the topic marker *nun*. This marker induces different and restricted readings than case-makers in term of information features [+/-Prominent (Prom)] and [+/-New] (Choi 1999). *Newness* is to do with whether an element in the sentence is already anchored in the discourse, or whether it is old/new(informative) information. On the other hand *prominence* is more relevant to the question of whether an element of a sentence can be singled out and then talked about among several potential alternatives. This is illustrated in the following sentences:

- (7.25) a. Swuni-**ka** Inho-lul mannassta  
 Swuni-NOM Inho-ACC met  
 ‘Swuni met Inho’
- b. Swuni-**nun** Inho-lul mannassta  
 Swuni-TOP Inho-ACC met  
 ‘As for Swuni, she met Inho’

(Choi 1999, p89,ex.47)

(7.25a) and (7.25b) have the same structure except that the former has a nominative marked subject whereas the latter has a topic-marked subject. With the nominative marked subject, (7.25a) gives a neutral reading in terms of the information features, that is, *Swuni* can be either [+New] or [-New] and [+Prom] or [-Prom] information in the sentence. However, with the topic marking in (7.25b), *Swuni* conveys [-New,+Prom] information which means that *Swuni* is presented as a distinct entity from the rest of the sentence: it is what the sentence is about and what the remaining part of the sentence is commenting on. However, when it comes to phrases other than the subject, which is generally in the sentence initial position, suffixing the topic marker *nun* does not seem to be enough to promote the topic reading. Consider the following topic context:

- (7.26) a. Inho-nun? Inho-lul nwuka manna-ss-e?  
 Inho-TOP Inho-ACC who.NOM meet-PST-Q  
 ‘What about Inho? Who met Inho?’
- b. \*Swuni-ka Inho-nun manna-ss-e  
 Swuni-NOM Inho-TOP meet-PST-DSE  
 ‘Swuni met Inho’
- c. Inho-nun Swuni-ka manna-ss-e  
 Inho-TOP Swuni-NOM meet-PST-DSE  
 ‘As for Inho, Swuni met him’

(Choi 1999, p175,ex.16)

This is a context that coerces the object to be interpreted as topic and the object *Inho* is the prominently presented given information (i.e. [-New, +Prom]). With this context, the default-order sentence (7.26b) is not good whereas (7.26c) is good when the object *-nun* phrase is fronted to the sentence initial position. This indicates that the topic reading with *-nun* phrases can only be guaranteed in the sentence initial position. (7.26c) is an example of a sentence with a fronted topic. It is cases of *caki*-binding where those mechanisms are involved that we will deal with in this section.



### 7.3.2 Data: Fronted Topics and Antecedents of *caki*

We have seen that a topic can be a binder for *caki* and this is the case for topic constructions such as the following:

- (7.27) a. \*Caki-ka Mary-lul ceil cohahanta  
 self-NOM Mary-ACC the most like  
 ‘Self likes Mary the most’  
 b. Mary-nun caki-ka ceil cohahanta  
 Mary-TOP self-NOM the most like  
 ‘Mary<sub>i</sub>, self<sub>i</sub> likes the most’

In the canonically ordered sentence (7.27a), *caki* cannot be bound by the lower object *Mary*. However, when the object becomes a topic by being topic marked and appearing in the sentence initial position, the topic marked object can now bind *caki*. The following sentence demonstrates a similar topic-binding, where the fronted topic-marked dative phrase binds *caki* in the complement clause.

- (7.28) a. John<sub>i</sub>-i Bill<sub>j</sub>-ekey [Mary<sub>k</sub>-ka caki<sub>i/\*j/k</sub>-lul silhehanta]-ko malhayssta  
 John-NOM Bill-DAT Mary-NOM self-ACC hate-COMPL said  
 ‘John<sub>i</sub> said to Bill<sub>j</sub> that Mary<sub>k</sub> hates self<sub>i/\*j/k</sub>’  
 b. Bill<sub>j</sub>-ekey-nun John<sub>i</sub>-i [Mary<sub>k</sub>-ka caki<sub>i/j/k</sub>-lul silhehanta]-ko malhayssta  
 Bill-DAT-TOP John-NOM Mary-NOM self-ACC hate-COMPL said  
 ‘To Bill<sub>j</sub>, John<sub>i</sub> said that Mary<sub>k</sub> hates self<sub>i/j/k</sub>’

In (7.28a), the dative NP, *Bill*, does not bind *caki*, whereas if the dative NP is *-nun* marked and fronted as in (7.28b), it qualifies as a legitimate binder of *caki*. However, the obvious generalisation that topics may bind *caki* does not hold across all constructions such as below:

- (7.29) a. Bill<sub>i</sub>-i [caki<sub>i</sub>-ka John<sub>j</sub>-ul silhehanta]-ko malhayssta  
 Bill-NOM self-NOM John-ACC hate-COMPL said  
 ‘Bill<sub>i</sub> said that self<sub>i</sub> hates John<sub>j</sub>’

- b. John<sub>j</sub>-un Bill<sub>i</sub>-i [caki<sub>i/\*j</sub>-ka e silhehanta]-ko malhayssta  
 John-TOP Bill-TOP self-NOM hate-COMPL said  
 ‘John<sub>j</sub>, Bill<sub>i</sub> said that self<sub>i/\*j</sub> hates ’

In (7.29a), the lower object *John* is not a potential binder of *caki*. But if the generalisation on topic binding is correct, then the topic *John* binding *caki* in (7.29b) should yield a grammatical sentence, contrary to fact. Even more strikingly, there are cases where topicalisation makes an apparently legitimate binder of *caki* incapable of binding it:

- (7.30) a. Mary<sub>i</sub>-ka [John<sub>i</sub>-i caki<sub>i/j</sub>-lul silhehanta]-ko malhayssta  
 Mary-NOM John-NOM self-ACC hate-COMPL said  
 ‘Mary<sub>i</sub> said that John<sub>i</sub> hates self<sub>i/j</sub>’  
 b. John<sub>i</sub>-un [Mary<sub>i</sub>-ka [e caki<sub>i/\*j</sub>-lul silhehanta]-ko malhayssta]  
 John<sub>j</sub>-TOP Mary-NOM self-ACC hate-COMPL said  
 ‘As for John<sub>j</sub>, Mary<sub>i</sub> said that e hates self<sub>i/\*j</sub>’

The lower subject *John* in (7.30a) binds *caki* as a case of subject-binding. However, once *John* is topicalised, i.e., fronted to the sentence initial position outside the complement clause as in (7.30b), it no longer binds *caki*. The ungrammatical binding between *John* and *caki* in (7.30b) may be attributed to a processing problem, whereby *Mary* is parsed as the subject of the embedded verb. However, a processing account of the data in (7.30b) would predict that the same pattern arises in all cases of long-distance topicalisation, regardless of the presence or not of an anaphor. As (7.31) demonstrates, this is simply not the case:

- (7.31) a. Mary-ka [John-i Bill-ul silhehanta]-ko malhayssta  
 Mary-NOM John-NOM Bill-ACC hate-COMPL said  
 ‘Mary said that John hates Bill’  
 b. John-un [Mary-ka [e Bill-lul silhehanta]-ko malhayssta]  
 John-TOP Mary-NOM Bill-ACC hate-COMPL said  
 ‘As for John, Mary said that e hates Bill’

type of unbounded dependencies is not allowed. We would like to identify what this property could be and how it can be explained.

### 7.3.3 “Scrambling” vs. Extraction in HPSG

In this section, we will show how the two kinds of topicalisation we saw in the last section can be explained in HPSG. As will become apparent, the two kinds of topicalisation belong to completely different types of construction, which explains the discrepancy in topic-binding patterns. The two following subsections will provide an HPSG account for each topicalisation and the representation of DEPS, at which level binding relations are determined. With the proper representation of DEPS for the two kinds of topicalisation, we will then examine how these different accounts for the two types of topicalisation will predict the topic-binding relations.

#### *Short-Distance Topicalisation*

For short-distance topicalisation as in (7.27b) and (7.28b), we argue that it should be analysed by the ‘Head-Complement Schema’ that is assumed to license clausal structure in free word order languages such as Japanese and Korean (Pollard & Sag 1994, 40)<sup>7</sup>. Contrary to long-distance topicalisation, we propose that short-distance topicalisation does not involve extraction, but is rather to be treated on a par with “scrambling”. This is supported by the distribution of resumptive pronouns in the two constructions.

- (7.33) a. \*Mary-nun, John-i *kunye-lul* cohahanta  
 Mary-TOP John-NOM she-ACC likes  
 ‘As for Mary, John likes *her*’
- b. John-un [Mary-ka [Bill-i {ku-lul/e} cohahanta]-ko sayngkakhanta]  
 John-TOP Mary-NOM Bill-NOM he-ACC likes-COMPL thinks  
 ‘As for John, Mary thinks that Bill likes (*him*)’

---

<sup>7</sup>This type of free ordering is also referred to as scrambling. A term originally due to Ross (1967).

(7.33a) and (7.33b) have resumptive pronouns in the relevant positions. As resumptive pronouns are only allowed in positions from which an object is extracted, the ungrammaticality of (7.33a) indicates that short-distance topicalisation does not involve extraction. On the other hand, long-distance topicalisation allows resumptive pronouns as in (7.33b) and this supports the hypothesis that long-distance topicalisation does involve extraction. Furthermore, as will become clear in the following section, the binding patterns in (7.27a and 7.27b) would be mysterious if extraction is involved.

Having argued that short-distance topicalisation does not involve extraction, let us have a look at the DEPS representation of such constructions. Consider first the following two examples:

- (7.34) a. Mary-ka John-ul cohahanta  
 Mary-NOM John-ACC likes  
 ‘Mary likes John’  
 DEPS ⟨ Mary, John ⟩
- b. John-ul Mary-ka cohahanta  
 John-ACC Mary-NOM likes  
 ‘Mary likes John’  
 DEPS ⟨ Mary, John ⟩

From the default sentence in (7.34a), (7.34b) is a scrambled version where the object *John* is scrambled over the subject *Mary*. As the sentences (7.34a) and (7.34b) only differ in word-order, having semantically the same interpretation, the predicational relation among the arguments of the head remains the same such that the external argument *Mary* comes as the first element and the Theme *John* as the second. Therefore, scrambling itself does not alter the order of elements in DEPS.

### ***Long-distance Topicalisation: Extraction***

We have seen in the last section that long-distance topicalisation involves extraction, thus it is of the type of *gapped topic constructions*. Two main types of approach for

gapped topic constructions in Korean and also for similar structures in Japanese have been proposed in literature. The first one is a Movement account which argues that the gap in the comment part is a trace (of scrambling) (e.g., Kuroda (1965); Hasegawa (1984); Imai (1983) and Saito (1985)). The other one assumes that the topic phrase is base-generated in its surface position (e.g., Kuno (1973); Hoji (1985) and Moon (1994)). Before examining how long-distance topicalisation is explained in HPSG, we would like to make clear that it should be derived by extraction as opposed to base-generation.

Saito (1985) provides the following descriptive generalisation of topicalisation with respect to scrambling in Japanese:

- (7.35) a. Topicalisation, but not scrambling, allows resumptive pronouns.  
 b. Scrambling, but not topicalisation, is subject to subjacency.

(Saito 1985, p.325)

The possibility of resumptive pronouns in topicalisation (7.35a) is also attested in Korean. Consider the following examples which illustrate long-distance scrambling and long-distance topicalisation respectively:

- (7.36) a. John<sub>i</sub>-**lul** Bill-i Mary-ka {\*ku<sub>i</sub>-lul/e<sub>i</sub>} silhehanta ko  
 John-**OBJ** Bill-NOM Mary-NOM he-**OBJ** dislike COMPL  
 mitko-issta  
 believing-is  
 ‘John<sub>i</sub>, Bill believes that Mary dislikes {\*him<sub>i</sub>/e<sub>i</sub>}’
- b. John<sub>i</sub>-**nun** Bill-i Mary-ka {ku<sub>i</sub>-lul/e<sub>i</sub>} silhehanta ko  
 John-**TOP** Bill-NOM Mary-NOM he-**OBJ** dislike COMPL  
 mitko-issta  
 believing-is  
 ‘(Speaking of) John<sub>i</sub>, Bill believes that Mary dislikes {him<sub>i</sub>/e<sub>i</sub>}’

(Moon 1994, p.442, ex.10)

(7.36a) shows that long-distance scrambling does not allow resumptive pronouns, whereas long-distance topicalisation in (7.36b) can have the original position of the topic phrase replaced by the resumptive pronoun. This shows that long-distance topicalisation does involve extraction as resumptive pronouns are only allowed in the position from which an element is extracted. Recall also that this contrasts with the impossibility of resumptive pronouns in short-distance topicalisation, which indicates that short-distance topicalisation does not involve extraction (see section 7.3.3). For Saito (1985), the equivalent example to (7.36b) is ambiguous in that the topic phrase may be base-generated in that position as proposed in Kuno (1973), in which case the empty element will be *pro*. In fact, Moon's (1994) example in (7.36b) is provided as the evidence that topicalisation is derived by base-generation. Thus, testing in terms of resumptive pronouns does not fully prove that long-distance topicalisation involves extraction. Then the other test we can rely on is whether topicalisation is or is not subject to subjacency as stated in (7.35b). Moon (1994), following Saito's (1985) arguments for Japanese, provides the following examples in favour of her argument that topicalisation is a case of base-generation. If base-generation is the right derivation for long-distance topicalisation, the relevant constructions should not be sensitive to Island Constraints (Ross 1967) whereas the scrambling cases should be affected by Island Constraints:

- (7.37) a. \*Mary<sub>j</sub>-**lul** John-i [e<sub>i</sub> e<sub>j</sub> chacassten] salam<sub>i</sub>-ul poassta  
 Mary-ACC John-NOM was-looking-for person-ACC saw  
 '\*Mary<sub>j</sub>, John saw a person who was looking for e<sub>j</sub>'
- b. Mary<sub>j</sub>-**nun** John-i [e<sub>i</sub> e<sub>j</sub> chacassten] salam<sub>i</sub>-ul poassta  
 Mary-TOP John-NOM was-looking-for person-ACC saw  
 '(Speaking of) Mary<sub>j</sub>, John saw a person who was looking for e<sub>j</sub>'
- (Moon 1994, p.445, ex.18)

The ungrammaticality of (7.37a) indicates that long-distance scrambling is indeed derived by extraction, whereas its grammatical topicalisation counterpart in (7.37b) shows that it is a case of base-generation, thus with no extraction, though we would like to point out that Moon's (1994) grammaticality judgements for sentences like

(7.37b) are questionable. However, even if we admit that (7.37b) may exhibit a varying degree of grammaticality depending on dialectal factors, the following examples still contradict the base-generation hypothesis for long-distance topicalisation.

- (7.38) a. Nwu-ka [Mary-ka John-ul paypanhay-ss-ta-nun] somwun-ul  
 who-NOM Mary-NOM John-ACC betray-PAST-DC-COMPL rumour-ACC  
 mit-ess-ni?  
 believe-PAST-Q  
 ‘Who believed the rumour that Mary betrayed John?’
- b. \*Mary<sub>i</sub>-nun nwu-ka [e<sub>i</sub> John-ul paypanhay-ss-ta-nun]  
 Mary-TOP who-NOM John-ACC betray-PAST-DC-COMPL  
 somwun-ul mit-ess-ni?  
 rumour-ACC believe-PAST-Q  
 ‘As for Mary<sub>i</sub>, who believes the rumour that e<sub>i</sub> betrayed John?’

(7.38b) is derived from (7.38a) by topicalising *Mary* from subject position within the complex NP clause. This example turns out to be totally ungrammatical as expected under the extraction approaches. Even if (7.37b) was grammatical as suggested in terms of base-generation in Moon (1994), the base-generation hypothesis would not be able to account for the ungrammaticality of (7.38b) as a base-generated construction should not be sensitive to Island Constraints. The extraction hypothesis, however, can easily accommodate such sentences. (7.38b) not only violates the Wh-Island constraint, but also the Complex NP constraint (Ross 1967) due to the extraction of *Mary* out of the complex NP clause. This multiple violation of the Island Constraints then explains the ungrammaticality of (7.38b), and provides further confirmation that long-distance topicalisation indeed involves extraction. Also, as far as the test of subjacency is concerned, as long-distance scrambling is also sensitive to subjacency, thus involves extractions, we suggest that it should be analysed on a par with long-distance topicalisation. Let us now move on to examine how HPSG accounts for constructions like long-distance topicalisation.

In section 6.4, we have seen the mechanism of extraction in HPSG and provided the appropriate modifications to the representation of DEPS. The same representation



of DEPS applies to long-distance topicalisation in Korean, as an instance of an unbounded dependency construction. Let us recall the examples of long-distance topicalisation we have looked at, repeated below:

- (7.39) a. Mary<sub>i</sub>-ka [John<sub>j</sub>-i Bill<sub>k</sub>-ul cohahanta]-ko sayngkakhanta  
 Mary-NOM John-NOM Bill-ACC like-COMPL thinks  
 ‘Mary thinks John likes Bill’  
 DEPS ⟨ Mary, S⟨ John, Bill ⟩ ⟩
- b. Bill<sub>k</sub>-un Mary<sub>i</sub>-ka [John<sub>j</sub>-i e cohahanta]-ko sayngkakhanta  
 Bill-TOP Mary-NOM John-NOM like-COMPL thinks  
 ‘Bill, Mary thinks John likes’

From example (7.39a), when we topicalise the lower object *Bill*, the topic marked lower object *Bill* is extracted across the complement clause shown in (7.39b). The representation of DEPS of the matrix head ‘*sayngkakhanta (thinks)*’ is shown below:

$$(7.40) \quad \text{DEPS} \left\langle \text{MARY}_{[nom]i}, \text{S} \left\langle \text{JOHN}_j \text{ gap}_k \left[ \begin{array}{l} \text{LOCAL } \boxed{1} \\ \text{SLASH } \boxed{1} \end{array} \right] \right\rangle \right\rangle$$

The topic *Bill* as a filler does not appear in the DEPS, but its local information (indicated by  $\boxed{1}$ ) is carried by *gap*-ss which stays in its original position in the DEPS.

So far we have shown that short-distance and long-distance topicalisation are accounted for differently. The former is an instance of the *Head-comp Schema* which allows free word-order in languages like Korean. On the other hand, the latter as an instance of an unbounded dependency construction is explained by the *Filler-head Schema*. We have also shown what the representation of DEPS should be in each type of topicalisation. In the next subsection, we will observe how the representations of DEPS for the two kinds of topicalisation provided so far will predict the different topic-binding relations.

### 7.3.4 Predictions: Topic-Binding and DEPS

#### *Topic-Binding in Short-Distance Topicalisation*

In section 7.3.3, we have seen that scrambling does not alter the order of elements in DEPS. As we argued that short-distance topicalisation is a case of scrambling, we would predict that the following case of short-distance topicalisation would have the accompanying representation of DEPS given below:

- (7.41) a. John-un Mary-ka cohahanta  
John-TOP Mary-NOM likes  
'As for John, Mary likes him'  
b. DEPS  $\langle$  Mary, John  $\rangle$

(7.41a) is a short-distance topic construction where the topic-marked object *John* is fronted. The representation of DEPS in (7.41b) shows that the external argument *Mary* comes as the first element and the topicalised Theme as the second, just like what is assumed in scrambling. Let us now consider the following where the subject *Mary* is replaced by *caki*:

- (7.42) John-<sub>i</sub>-un caki-<sub>i</sub>-ka ceil cohahanta  
John-TOP self-NOM the most likes  
'As for John<sub>i</sub>, self<sub>i</sub> likes the most'  
DEPS  $\langle$  caki<sub>i</sub>, John<sub>j</sub>  $\rangle$

In (7.42), the object *John* is fronted over the subject *caki*. The corresponding DEPS is given such that the external argument *caki* comes first, followed by the topicalised object *John* as in (7.41). However, the given DEPS list does not predict the correct binding as *caki* cannot be bound by a lower element in DEPS. This is also the case when the dative PP is topic fronted as follows:

- (7.43) a. Bill<sub>j</sub>-ekey-nun John<sub>i</sub>-i [Mary<sub>k</sub>-ka caki<sub>i/j/k</sub>-lul silhehanta]-ko  
 Bill-DAT-TOP John-NOM Mary-NOM self-ACC hate-COMPL  
 malhayssta  
 said  
 ‘To Bill<sub>j</sub>, John<sub>i</sub> said that Mary<sub>k</sub> hates self<sub>i/j/k</sub>’
- b. DEPS < John<sub>i</sub>, S:< Mary<sub>k</sub>, caki<sub>i/j/k</sub> >, PP[Bill]<sub>j</sub> >

As short-distance topicalisation is a case of scrambling, DEPS would be the same as before the topicalisation. Thus, as in (7.43b), the dative PP should be the last element in DEPS. However, this representation of DEPS does not correctly predict the given binding relation where *Bill* binds *caki* because *Bill* is lower than *caki*. The problem we face in the data concerned comes from the fact that, even though short-distance topicalisation is a case of scrambling, the prominence of topic and its position preceding the commenting clause is not actually realised in DEPS in any way. Therefore, in order to make sure that in short-distance topicalisation the topic phrase appears as the first element in the argument structure, we postulate a lexical rule as follows:

(7.44) SHORT-DISTANCE TOPICALISATION LEXICAL RULE (SDTLR)

$$\left[ \begin{array}{l} \text{HEAD} \left[ \begin{array}{l} \textit{verb} \\ \text{DEPS} \langle \dots \boxed{n} \dots \rangle \end{array} \right] \\ \text{VAL} \langle \dots \boxed{n}_{[-nun]i} \dots \rangle \\ \text{CONX} [\text{TOPIC } \boxed{\quad}] \end{array} \right] \Rightarrow \left[ \begin{array}{l} \text{HEAD} \left[ \begin{array}{l} \textit{verb} \\ \text{DEPS} \langle \boxed{n}_{[-nun]i}, \dots \rangle \end{array} \right] \\ \text{VAL} [\langle \dots \boxed{n} \dots \rangle] \\ \text{CONX} [\text{TOPIC } \boxed{n}] \end{array} \right]$$

The input of the rule takes a lexical entry with an empty TOPIC value. From this input, short-distance topicalisation takes place when any member of the VALENCE list becomes a TOPIC by being marked with ‘-nun’. The lexical rule ensures that

the topic phrase is preposed in the output. This lexical rule will also block the generation of multiple topics, which are disallowed in Korean and Japanese (see Kuno (1973); for a different view, see Kuroda (1988, p.134)). Given the lexical rule above, the representation of DEPS in (7.43a) should be like the following:

(7.45) DEPS < PP[Bill]<sub>[top]j</sub>, John<sub>i</sub>, S:< Mary<sub>k</sub>, caki<sub>i/j/k</sub> > >

Via the Short Distance topicalisation Lexical Rule (SDTLR) we have provided in (7.44), *Bill* appears as the leftmost item in DEPS. This enables the topic *Bill* to bind *caki* together with other higher elements *John* and *Mary*. Let us also look at DEPS in (7.42) as the outcome of our lexical rule, given below:

(7.46) DEPS < John<sub>[top]i</sub>, caki<sub>i</sub> >

Here, *John* is the topic-marked object which is fronted and comes as the first element in DEPS, where *caki* is correctly bound by the higher element *John*.

This way, we provide a proper treatment for the case of topic-binding in short-distance topicalisation. We started from the proposal that short-distance topicalisation should be treated by the head-complement schema, which licenses clausal structure in free word order languages. This can be analysed as scrambling which allows the free ordering of constituents in a clause, thus does not involve extraction. As scrambling alone cannot guarantee the sentence initial position of topic in short-distance topicalisation, and thus is problematic for topic-binding, we provided a lexical rule which can properly promote the topic phrase to the first element in DEPS.

Short-distance topicalisation is similar to the double nominative construction in that both constructions express the topic phrase without involving extraction. Thus, we will call both topic constructions *gapless topic constructions*. However, it should be stressed that I classify short-distance topicalisation as gapless topic construction, unlike what is generally assumed in the literature where only the cases like double

nominative topic constructions are gapless topic construction; where a topic does not appear to correspond to any argument in the sentence.

As only the outer nominative in the double nominative constructions can be topic-marked, its representation of DEPS, like short-distance topicalisation, will have the topic phrase as the first element, which is another property of gapless topic constructions. This property of gapless topic constructions will be discussed again in section 8, together with the gapped topic constructions which we shall discuss in the next section. Let us now move onto the cases of long-distance topicalisation.

### *Topic-Binding in Long-Distance Topicalisation*

Consider first the example below in terms of the representation of DEPS:

- (7.47) Bill<sub>j</sub>-un Mary<sub>i</sub>-ka [caki<sub>i/\*j</sub>-ka e cohahanta]-ko sayngkakhanta  
 Bill-TOP Mary-NOM caki-NOM like-COMPL thinks  
 ‘Bill<sub>j</sub>, Mary<sub>i</sub> thinks caki<sub>i/\*j</sub> likes’

$$\text{DEPS} \left\langle \text{MARY}_{[nom]i} \text{ S} \left\langle \text{CAKI}_{[nom]i/*j}, \text{gap}_j \left[ \begin{array}{l} \text{LOCAL } \boxed{1} \\ \text{SLASH } \boxed{1} \end{array} \right] \right\rangle \right\rangle$$

In (7.47), the topic-marked lower object *Bill* is extracted across the complement clause. The accompanying DEPS in (7.47) contains *Mary*, the matrix subject, as the first item and *caki* as the first item in the embedded DEPS. The *gap-ss* marks the position of the extracted argument and carries its local value. *Caki* is bound by *Mary* and generalised topic-binding is not allowed. This is expected as the topic is not available at the relevant level of DEPS. Likewise, it is not that the generalisation on topic-binding breaks down for this particular example, but that the operation of the *Filler-head Schema* naturally makes the topic as a filler unavailable as a binder in the relevant level of DEPS. Therefore, we can draw an important distinction that in gapped topic constructions (long-distance topicalisation), a topic itself is not a part of DEPS, whereas in gapless topic construction (short-distance topicalisation), a topic always appears in DEPS.

However, this explanation does not extend to examples like (7.48):

- (7.48) John<sub>j</sub>-un [Mary<sub>i</sub>-ka [e caki<sub>i/\*j</sub>-lul silhehanta]-ko malhayssta]  
 John<sub>j</sub>-TOP Mary<sub>i</sub>-NOM self-ACC hate-COMPL said  
 ‘As for John<sub>j</sub>, Mary<sub>i</sub> said that e hates self<sub>i/\*j</sub>’

$$\text{DEPS} \left\langle \text{MARY}_{[\text{nom}]i}, \text{S} \left\langle \text{gap}_j \left[ \begin{array}{l} \text{LOCAL } \boxed{1} \\ \text{SLASH } \boxed{1} \end{array} \right], \text{CAKI}_{[\text{acc}]i/*j} \right\rangle \right\rangle$$

In (7.48), the topic is extracted from the lower subject position, which leaves the *gap-ss* behind in DEPS. Thus, we have two possible binders of *caki*: *Mary* and *gap-ss*. However, as the indexing shows, binding between *gap-ss* and *caki* is ruled out. In short, it is naturally expected that *caki* is not bound by the topic in this construction. But in DEPS, the *gap-ss* is present, representing the topic’s LOC value in the legitimate binder position for *caki*. How can binding between *gap-ss* and *caki* be blocked? As the binding relation in the DEPS list of (7.48) shows, *gap-ss* behaves as if it is there but excluded from determining the binding relations. This suggests two possible explanations.

The first approach is to assume that *gap-ss* does not bind anaphors. This seems simple and convenient. However, this certainly is not universal. Consider, for example, the following example of topicalisation in English:

- (7.49) Mary<sub>i</sub>, [John thinks [ e likes herself<sub>i</sub>.]]

$$\text{DEPS} \left\langle \text{JOHN}_{[\text{nom}]}, \text{S} \left\langle \text{gap}_i \left[ \begin{array}{l} \text{LOCAL } \boxed{1} \\ \text{SLASH } \boxed{1} \end{array} \right], \text{HERSELF}_{[\text{acc}]i} \right\rangle \right\rangle$$

(7.49) shows that in English the reflexive can be bound by a long-distance topic, unlike Korean. To allow such topic binding means that *gap-ss* should be able to bind the anaphor. Thus, such an explanation is not adequate cross-linguistically.

The second possibility is that *gap-ss* may not be visible for binding in DEPS. However, there is evidence confirming that *gap-ss* should be explicitly visible in DEPS

to account for binding of pronominals as well as topicalised anaphors. Observe first the following example:

- (7.50) \*Ku<sub>i</sub>-nun [Mary-ka [Bill<sub>i</sub>-i e chohahanta]-ko sayngkakhanta]  
 Him-TOP Mary-NOM Bill-NOM likes-COMPL thinks  
 ‘\*Him<sub>i</sub>, Mary thinks that Bill<sub>i</sub> likes’

The representation of DEPS of *chohahanta* (*likes*) is as follows:

- (7.51) *likes*  
 DEPS ⟨ Bill[nom]<sub>i</sub>, gap-ss<sub>i</sub> ⟩

In (7.50), the pronoun itself has been long-distance topicalised from the lower object position and leaves *gap-ss* which is also pronominal as it structure-shares with the filler’s LOCAL value. The local DEPS of the head *chohahanta* (*likes*) is presented in (7.51). The ungrammaticality of binding in (7.50) proves that *gap-ss* should be visible for binding in DEPS in order to predict the violation of Principle B. The same applies when an anaphor itself is topicalised as shown below<sup>8</sup>. Note that in (7.52), the provided indexing shows that *caki* is bound by the matrix subject but not by the local subject. This point will be addressed in Section 7.3.5.:

- (7.52) Caki<sub>i/\*j/\*k</sub>-nun [Mary<sub>i</sub>-ka [Bill<sub>j</sub>-i e chohahanta]-ko sayngkakhanta]  
 self-TOP Mary-NOM Bill-NOM likes-COMPL thinks  
 ‘Self<sub>i/\*j/\*k</sub>, Mary<sub>i</sub> thinks that Bill<sub>j</sub> likes’

The representation of DEPS of *sayngkakhanta* (*think*) is as follows:

---

<sup>8</sup>Recall that in section 7.2.4, *caki* itself cannot be *-nun* marked due to the rigid word order in the double nominative construction. Contrary to such examples, however, in the gapped topic constructions as (7.52) above, the lower object *caki* is topicalised. This looks like a counterexample to our argument that *caki* is not discourse bound. It should be noted that the two topic constructions, gapless topic constructions (e.g. topic-marked outer nominative in the double nominative constructions and short-distance topicalisation) and gapped topicalisation (e.g. long-distant topicalisation) are two distinct constructions, and the impossibility of *-nun* marked *caki* only applies to gapless topic constructions. Another apparent reason that *caki* is not discourse bound but sentence internally bound comes from the fact that the binding relation is determined in DEPS. Given the A-binding principle, DEPS only contains the arguments and the adjuncts of the predicate, thus disallowing any discourse entity. All the more, as the indexing shows, even if it is bound by some discourse entity, such binding is not allowed.



(7.53) *thinks*  
 DEPS  $\langle \text{NP}[\textit{nom}]_i, \text{S}:\langle \text{NP}[\textit{nom}]_j, \textit{gap-ss}_{i/*j/*k} \rangle \rangle$

In (7.52), *caki* itself is long-distance topicalised out of the complement clause. In the representation of DEPS in (7.53), the *gap-ss*, being anaphoric, should be visible for binding to explain the binding between the matrix subject *Mary* and *gap-ss*. Therefore, the assumption on the invisibility of *gap-ss* turns out to be wrong.

Having seen that the two rather general assumptions fail to cover the basic data, we will look more closely for possible parametric explanations in the next section.

### 7.3.5 Parametric Possibilities

As we have seen in the last section, topicalisation in languages like English allows topic-binding with *gap-ss* binding the reflexive. This implies that the suggestion that *gap-ss* cannot bind *caki* may only be relevant to Korean, or possibly other languages with long-distance anaphors. In this section, we present two possible parametric explanations: One is based on the properties of *gap-ss* itself, which behaves like an unrealised pronominal, *pro* (Tsoulas 1999). The other is more to do with the property of long-distance anaphors themselves, such as *caki* and *zibun* in Japanese, which are prohibited from being bound by *gap-ss*.

#### ***Tsoulas 1999***

Looking closely at the binding relations in (7.52) above, we observe that although *gap-ss* can be bound by the subject *Mary*, it resists binding by the local subject *Bill*. This is clearly problematic in terms of *gap-ss* as pointed out earlier since *gap-ss* being anaphoric, like the filler *caki* in (7.53), should be able to be bound by the local subject *Bill*. A similar set of data in terms of long-distance scrambling in Korean as well as Japanese receives a different analysis in Tsoulas (1999). His analysis is cast within a minimalist framework and he argues against a movement account and in favour of a “base-generation” account of long distance scrambling. In his theory, the thematic positions of long distance scrambled elements are filled

not by traces but by empty pronominals (*pro*). Thus, the following examples are explained in his theory as violations of Principle B:

- *Multiple topicalisation/scrambling*

- (7.54) a. Mary<sub>i</sub>-ka [Bill<sub>j</sub>-i caki<sub>i/j</sub>-lul chohahanta]-ko malhayssta  
 Mary-NOM Bill-NOM self-NOM like-COMPL said  
 ‘Mary said that Bill likes self’
- b. Bill<sub>j</sub>-un/i caki<sub>i/\*j</sub>-lul [Mary<sub>i</sub>-ka [e e chohanata]-ko malhayssta]  
 Bill-TOP/NOM self-ACC Mary-NOM like-COMPL said

(Tsoulas 1999)

In (7.54b), the two arguments in the lower clause are both long-distance scrambled/topicalised, *caki* originating in the lower object position. The given binding relations of *caki* is the same as in (7.52). That is, *caki* are not bound by the local subject topic *Bill* while it can be bound by the higher subject *Mary*. If we assume with Tsoulas (1999) that empty pronominals occupy the positions marked *e*, then both the ungrammatical binding of *caki* to *Bill* in (7.54b) and the binding pattern of (7.54b) are independently explained. While this account adequately explains examples such as (7.52) and (7.54b), the analysis in terms of *gap-ss* still faces problems with such data. For (7.54b), the representation of DEPS would be as follows:

- (7.55) *said*  
 DEPS ⟨ NP[*nom*]<sub>i</sub>, S:⟨ *gap-ss*<sub>j</sub>, *gap-ss*<sub>i/\*j</sub> ⟩ ⟩

The second *gap-ss* being anaphoric like its filler *caki*, nothing blocks its binding to the first *gap-ss*. However, the given binding relation shows the opposite.

Even though Tsoulas’s (1999) analysis accounts for a significant part of the data, crucially it fails to explain the case where the lower subject is topicalised as in (7.48), which was also problematic for the explanation in terms of *gap-ss*. The example is repeated below:

- (7.56) John<sub>j</sub>-un [Mary<sub>i</sub>-ka [e caki<sub>i/\*j</sub>-lul silhehanta]-ko malhayssta]  
 John<sub>j</sub>-TOP Mary-NOM self-ACC hate-COMPL said  
 ‘As for John<sub>j</sub>, Mary<sub>i</sub> said that e hates self<sub>i/\*j</sub>’

In (7.56), the lower subject empty element binds *caki*. According to Tsoulas (1999), the lower subject position is filled by an empty pronominal which in turn should be able to bind *caki*. However, this contradicts the given binding relation.

Leaving aside cases like (7.56) above, accommodating Tsoulas’s (1999) account within HPSG poses significant problems. There are two options. One is that if we keep our analysis using the filler-gap constructions, we must be able to explain why *gap-ss* should behave like a pronominal. However, this is not simple because the property of *gap-ss* as a nominal object is entirely dependent on that of the filler. Unless it is proved that *gap-ss* is locally free in independent constructions, it would be hard to limit the property of *gap-ss* just to pronominals. The other option is to assume long-distance topicalisation as a case of left-dislocation, as argued for long-distance scrambling by Tsoulas (1999). As left-dislocation does not involve movement in GB terms, this implies that analysing long-distance topicalisation as a filler-gap construction may be questionable. However, there have been recent works that attempt to explain cases of left-dislocation as instances of extraction in HPSG. For example, Alexopoulou (1999) argues that clitic left dislocation in Greek is an instance of adjunct extraction and she explains the phenomenon in the domain of the filler-gap construction. Either of these options may lead to a more complete picture for *caki-binding* in long-distance topicalisation. We will not discuss those possible explanations to their full extent. However, for the data not yet completely covered, we will move on to provide another possible explanation, focusing on the properties of *caki* itself.

### ***Anaphors and $\phi$ -features***

One of the similarities among long-distance anaphors in many languages is that they lack a full specification of  $\phi$ -features such as number, person and gender. For

example, *caki* in Korean is third person singular but gender unspecific. In addition to *caki*, Korean has another long-distance anaphor *casin*, which can be used on its own and has the same distribution as *caki*. The difference, however, is that *casin* can form another kind of anaphor by combining with a personal pronoun, illustrated below:

- (7.57) Ku/Kunyo + *casin* = Kucasin/Kunyocasin...  
 He/Her... self Himself/Herself...

*Casin* is also third person singular and gender-unspecific. However, when it combines with the third person feminine pronoun *kunyo*, yielding the form *kunyecasin*, it becomes fully specified in  $\phi$ -features (including gender). Crucially, it can then be bound by a long-distance topicalised (or scrambled) element, unlike *caki*:

- (7.58) a. Mary<sub>j</sub>-nun/ka [Bill<sub>i</sub>-i [*e casin*<sub>\*j/i</sub>-ul chohahanta]-ko sayngkhanta]  
 mary-TOP/NOM Bill-NOM self-ACC like-COMPL thinks  
 ‘Mary<sub>j</sub>, Bill<sub>i</sub> thinks likes self<sub>\*j/i</sub>’  
 b. Mary<sub>j</sub>-nun/ka [Bill<sub>i</sub>-i [*kunyecasin*<sub>j/\*i</sub>-ul chohahanta]-ko sayngkhanta]  
 Mary-TOP/NOM Bill-NOM herself-ACC like-COMPL thinks  
 ‘Mary<sub>j</sub>, Bill<sub>i</sub> thinks likes herself<sub>j/\*i</sub>’

In Japanese and Hindi, where long-distance anaphors are also attested, the same observation concerning scrambling and Binding is made by Saito (1992) and Mahajan (1989). Interestingly, the anaphors in these languages also uniformly lack a full  $\phi$ -feature specification. Observe also that the pattern in (7.57) is also observed with *zisin* in Japanese:

- (7.59) Mary-wa/ga [Bill-ga [kanozyozisin-ga sukida] to omotteiru]  
 Mary-TOP/NOM Bill-NOM herself-NOM like COMP think  
 ‘Mary, Bill thinks likes herself’

To have a clear picture of the binding relations with respect to the  $\phi$ -specification, a summary of the binding relations in the data we have seen is given below:

- (7.60) Filler, DEPS of the Head  $\langle \rangle$
- a. (7.58b) and (7.49)  
 $NP_{\square i}, \langle gap_i \rightarrow refl[\phi] \rangle$
- b. (7.48) and (7.58a)  
 $NP_{\square i}, \langle gap -\times \rightarrow caki[?] \rangle$
- c. (7.52)  
 $Caki_{\square i}, \langle NP[\phi] \rightarrow gap_i \rangle$

Each item above illustrates the filler and DEPS with  $\phi$ -specifications (whether fully specified or underspecified) for the corresponding examples as indicated. (7.60a) shows the case of English and Korean topicalisation where the reflexives are fully specified in  $\phi$ -features. The binding relations show that *gap-ss* can bind the reflexives. (7.60b) illustrates that *gap-ss* cannot bind *caki* which is not fully specified in  $\phi$ -features (indicated by ‘?’). Lastly, (7.60c) indicates that *gap-ss* can be bound by an NP which is fully specified in  $\phi$ -features.

From the above, the following empirical generalisation emerges.

- (7.61) Binding is established when one of the two nominals (the binder or the bindee) has a fully-specified  $\phi$ -feature matrix.

In HPSG terms, we can state that *gap-ss* cannot be the binder or the bindee of an NP which is not fully specified in  $\phi$ -features. However, it is not straightforward to implement this generalisation in HPSG, where coindexing is established by structure-sharing of the INDEX value between the binder and the bindee. INDEX has the features PER (person), NUM (number) and GEN (gender) as its value (see section 4.2.1). As *caki* is underspecified for GEN, it is compatible with either a [GEN:*masculine*] or a [GEN:*feminine*] binder. This means that the underspecification for GEN of *caki* does not stop it being coindexed with any other nominal which has a fully specified INDEX value. This also implies that *gap-ss* is free to bind *caki* as the former can have a fully specified index value structure-shared with the filler.

In order to surmount the technical problem and to capture the generalisation provided in (2.44), we first propose the following internal structure of *gap-ss*:

$$(7.62) \quad \left[ \begin{array}{l} \textit{gap-ss} \\ \\ \text{LOCAL} \mid \text{CAT} \mid \text{HEAD} \dots \boxed{\square} \end{array} \left[ \begin{array}{l} \text{INDEX} \left[ \begin{array}{l} \text{AGR} \left[ \begin{array}{l} \text{PER} \quad \textit{per} \\ \text{NUM} \quad \textit{num} \\ \text{GEN} \quad \textit{gen} \end{array} \right] \\ \text{INST} \quad \textit{i} \end{array} \right] \end{array} \right] \right] \right]$$

The feature structure of *gap-ss* above has an additional feature INST which is distinct from the ARG feature (for agreement, or  $\phi$ -feature) and takes an index as its value. We suggest that in Korean, it is only the INST value of *gap-ss* (indicated by ‘*i*’) that is relevant to establishing a binding relation in extraction constructions, but not the INDEX value (indicated by ‘ $\square$ ’). That is, in this context, *gap-ss* becomes a radically underspecified element for binding, and this is what relates to its suppressed binding relations with *caki* which is not fully specified in the ARG value. For *caki* to obtain a fully specified value of GEN, the binder must be fully specified in the ARG value. It then follows that *gap-ss* cannot be a possible antecedent because it is only the INST value, (an indexing) that can be offered to *caki*, but not the ARG value.

This explains not only cases where a single element is long-distance topicalised as in (7.60), but also cases where two elements are topicalised as in (7.54b). When there are two *gap-ss* elements in consideration for the binding relation, binding is not licensed as neither of them can fulfill the full specification of the  $\phi$ -features.

So far, we have examined a complicated problem dealing with binding relations between *gap-ss* and *caki*. From the observation of extensive data, we proposed that the key to the explanation essentially lies in the properties of *caki* with respect to  $\phi$ -features such that *gap-ss* does not establish a binding relation with an NP with impoverished  $\phi$ -features. In order to achieve a proper implementation of this generalisation in HPSG, we have modified the internal structure of *gap-ss* for binding in Korean to render its  $\phi$ -features (in ARG) irrelevant to binding.

### 7.3.6 Conclusion and Open Problems

In this chapter, we have provided a structural account for cases of *caki*-binding which have been considered to be subject to discourse conditions only. We have proposed that the data in question, topic-binding, can be naturally explained by appealing to the analysis of a different Korean Construction: Double nominative Constructions and short/long-distance topic constructions. Formal analysis for topic-binding licensed in such constructions was implemented by using lexical rules: Double Nominative Lexical Rule (DNLR) and Short-distance topicalisation Lexical Rule (SDTLR). We have also established that long-distance scrambling and long-distance topicalisation should be treated on a par. A subset of the binding facts that we have observed can be explained by the introduction of TOPIC to the argument structure list. The main difficulty lies in blocking binding of *caki* by *gap-ss*. We have suggested a possible explanation based on the property of *gap-ss* being a pronominal and the lack of full specification in *caki*'s  $\phi$ -features. We showed that each of these possibilities reflects a significant part of empirical reality. The question of better and more complete formal implementation of one or more of these possibilities, or part thereof, is left open.

The next chapter will be devoted to providing a more unified account for topic-binding using the multiple inheritance hierarchy. We will categorise topic constructions in two groups: Gapped topic constructions and Gapless topic constructions. To the current dimension of phrasal and clausal types, we will add an independent dimensional constraint for topicalisation, TOPICALITY. The subtypes of TOPICALITY will be cross-classified with the subtypes of *phrase* and CLAUSE. We will argue that this type hierarchy will capture both the universality and parametric variations of topicalisation cross-linguistically, thus will provide a more unified account of topicalisation. Based on this type hierarchy, we will propose that topic-binding in different topic constructions in Korean can be explained in a more consistent way without having to partly rely on lexical rules.



## Chapter 8

# Topic-Binding in The Type-Hierarchy

In the last chapter, we looked at three different topic constructions: DNC topic constructions in section 7.2 and short and long-distance topicalisation in section 7.3. In accounting for *caki*-binding, we proposed two different lexical rules: DNCLR (Double Nominative Constructions Lexical Rule) and SDTLR (Short-Distance Topicalisation Lexical Rule). In this section, we will provide an explanation for *caki*-binding phenomena in long-distance topicalisation using lexical rules. Overall, we have seen that the account for each case of topic-binding is adequate in that we did not need to postulate any addition to or modification of the binding principle. In spite of that, the accounts of different *caki*-binding in different topic constructions do not seem very consistent. For example, two of the three cases of topic-binding need lexical rules (DNCLR, SDTLR) whereas the other case (topic binding in long-distance topic constructions) does not need any lexical rule. Could there be a more unified account for the three cases of topic-binding?

Recall that the three types of topic constructions can be reduced to the two main groups: gapless and gapped topic constructions. While long-distance topicalisation belongs to the gapped topic construction, the DNC topicalisation and short-distance topicalisation belong to the gapless topic construction because they do not involve any extraction, thus not leaving any *gap*. If we can make a further generalisation in terms of the two-way distinction in types of topicalisation and further provide a

similar way of accounting for *caki*-binding in the two different topic constructions, we will be able to provide a more generalised and elegant explanation of topic-binding.

In this chapter, we will provide a unified treatment of topic-binding in terms of a type-hierarchy. All the instances of *caki*-binding in question involve topic clauses. Furthermore, each instance of *caki*-binding in different topic constructions is associated with a particular syntactic structure: gapped topic constructions with the *hd-filler-schema* and gapless topic constructions with the *hd-comp-schema* and double nominative constructions. Therefore each topic construction is cross-classified with a *clausal* and *phrasal* type. Each topic construction will inherit the constraints from its super types (*phrase* and *clause*), which will also yield a particular representation of DEPS. Thus, the DEPS list in each type of topicalisation will be given in terms of a multiple inheritance hierarchy rather than lexical rules.

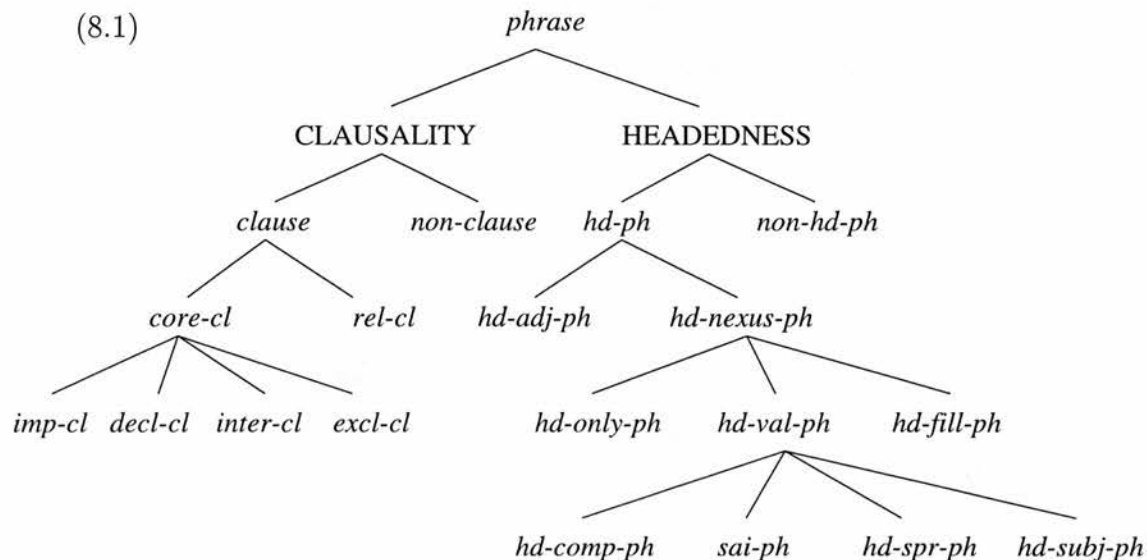
In the following section, we will begin by reviewing the standardly accepted hierarchy of the types *clause* and *phrase*, to which we will propose relevant modifications. Then we will have a close look at the constraints on the relevant types and attempt to modify some constraints to suit Korean topic constructions better and which will lead to correct predictions for topic-binding. Lastly, we will discuss the merits of the account based on the multiple inheritance hierarchy in comparison with the approach using lexical rules.

## 8.1 Type-Hierarchy: Phrases and Clauses

In section 4.1.2, we reviewed the basics of a type hierarchy in HPSG. *Signs* are organised into a type hierarchy where more general supertypes, are classified into more specific subtypes and a subtype must obey all the constraints on its supertype. We illustrated the type hierarchy of type *word* and, as an example, we saw a multiple inheritance hierarchy for the lexical verb *adore*. Likewise, phrases and clauses can be modelled in the type-hierarchy in terms of whether they are headed or not, or what kinds of daughters are involved, etc. The type hierarchy of such phrasal structure

information is further extended with the informational dimension of clausality to express generalisations about the shared properties of diverse types of phrases.

The types *phrase* and *clause* are classified in HPSG, as below, based on common phrasal structures in English and many other languages (Ginzburg & Sag to appear).



Seen as above, phrases are classified as either *headed-phrase* (*hd-ph*) or *non-headed-phrase* (*non-hd-ph*), each type exhibiting a variety of subtypes. Among the headed-phrases, a distinction is drawn between *head-adjunct-phrase* (*hd-adj-ph*) and *head-nexus-phrase* (*hd-nexus-ph*), the latter kind being broken down into the three subtypes *head-filler-phrase* (*hd-filler-ph*), *head-valence-phrase* (*hd-val-ph*) and *head-only-phrase* (*hd-only-ph*). Finally, *hd-val-ph* has four subtypes: *head-subject-phrase* (*hd-subj-ph*), *head-complement-phrase* (*hd-comp-ph*), *head-specifier-phrase* (*hd-spr-ph*), and *subject-auxiliary-inversion* (*sai-ph*), as indicated. On the side of the phrase type-hierarchy, moreover, this classification allows us not only to recognise a distinction between clauses and non-clauses, but also to identify at least the following subtypes of the type *clause*: *decl-cl* (*declarative-clause*), *inter-cl* (*interrogative-clause*), *imp-cl* (*imperative clause*), *excl-cl* (*exclamative-clause*), *core-cl* (*core-clause*) and *rel-cl* (*relative clause*).

Just as in the case of the lexicon, phrasal and clausal types also obey type-specific constraints. For example, the type *phrase* on the top of the hierarchy is assigned the most general constraint as follows:

$$(8.2) \quad \textit{phrase} \implies \left[ \text{SS} \mid \text{LOC} \mid \text{CAT} \mid \text{COMP} \quad \langle \rangle \right]$$

This constraint is a formally stated version of the Empty COMPS Constraint (i.e., all complements are saturated), saying that all phrases have the empty list as their COMPS value. From this super type *phrase*, its subtypes come to have more specific and restricted constraints. For instance, the following constraint has to be satisfied for the type *hd-ph*: this constraint is a formally restated version of the Head-Feature Principle (HFP) that we have already seen in section 4.1.2:

$$(8.3) \quad \text{Head Feature Principle} \\ \textit{hd-ph} \implies \left[ \begin{array}{l} \text{SS} \mid \text{LOC} \mid \text{CAT} \mid \text{HEAD} \quad \boxed{1} \\ \text{HD-DTR} \left[ \text{SS} \mid \text{LOC} \mid \text{CAT} \mid \text{HEAD} \quad \boxed{1} \right] \end{array} \right]$$

Now we will look at other subtypes of *hd-ph* which are directly relevant to the explanation of topic-binding. Firstly, a lexical head and the zero or more complements that it selects may form a phrase of type *hd-comp-ph* (it is a subtype of *hd-ph*, *hd-nexus-ph* and *hd-val-ph*, thus should satisfy all the constraints of those supertypes). Furthermore, the type *hd-comp-ph* is subject to the following constraint:

$$(8.4) \quad \text{Constraint on Head-Complement-phrase (standard)} \\ \textit{hd-comp-ph} \implies \left[ \begin{array}{l} \text{HD-DTR} \left[ \begin{array}{l} \textit{word} \\ \text{SS} \mid \text{LOC} \mid \text{CAT} \mid \text{COMPS} \langle \boxed{1}, \dots, \boxed{n} \rangle \end{array} \right] \\ \text{NON-HD-DTR} \left\langle \left[ \text{SS} \quad \boxed{1} \right], \dots, \left[ \text{SS} \quad \boxed{n} \right] \right\rangle \end{array} \right]$$

This just says that any number of non-head daughters are allowed and their SYNSEM values are identified with the value of COMPS on the head daughter that selects them.

Similarly, the following constraint is relevant to the type of phrase where a verb phrase is combined with a subject:

(8.5) Constraint on *hd-subj-ph*

$$hd\text{-subj-ph} \implies \left[ \begin{array}{l} SS | LOC | CAT | SUBJ \langle \rangle \\ HD\text{-DTR} \left[ \begin{array}{l} phrase \\ SS | LOC | CAT | SUBJ \langle \boxed{1} \rangle \end{array} \right] \\ NON\text{-HD} | DTRS \langle [ \boxed{1} ] \rangle \end{array} \right]$$

The difference here from the constraint for the type *hd-comp-ph* is that only one non-head daughter is allowed and the SYNSEM value of such a non-head daughter structure-shares with the SUBJ value of the head daughter. Also note that the phrase itself has an empty value for the corresponding feature (i.e., SUBJ). Although it is the case that most languages allow only one subject per clause, this is not quite universal as we have seen from cases of double nominative constructions in Korean. We will propose a subtype of *hd-subj-ph* to suit this parametric variation in the next section.

Another relevant subtype of *hd-ph* is type *hd-fill-ph*, which is associated with unbounded dependency constructions and which is subject to the following constraint we have seen in section 6.4:

(8.6) Constraint on Head-Filler-Phrase

$$hd\text{-fill-ph} \implies \left[ \begin{array}{l} SS \quad \left[ \begin{array}{l} LOC | CAT | HEAD \textit{verb} \\ SLASH \boxed{2} \end{array} \right] \\ HD\text{-DTR} | SS \quad \left[ \begin{array}{l} SLASH \{ \boxed{1} \} \uplus \boxed{2} \end{array} \right] \\ NON\text{-HD}\text{-DTRS} \quad \left\langle \left[ \begin{array}{l} SS | LOC \boxed{1} \end{array} \right] \right\rangle \end{array} \right]$$

The constraint in (8.6) states that the phrase should be a verbal projection. Also, among the members of the SLASH set of the head-daughter, whatever is identified

with the local value of the non-head daughter is removed (discharged) and the rest is passed up to the phrase's SLASH set. A typical example of such a type in English would be in topicalisations such as the following:

- (8.7) a. The cake, John likes.  
 b. Mary, John believes Bill likes

To account for the English type of topicalisation as an instance of an unbounded dependency construction, another type called *top-cl* (topicalised clause) is posited as a subtype of both *hd-filler-ph* and *core-cl*. Thus, the type *top-cl* inherits the constraints from both supertypes, *hd-filler-ph* and *core-cl*, and also has to obey the following further constraints: (Ginzburg & Sag to appear):

(8.8) Constraint on Topic clausal type (standard for English)

$$\begin{array}{l}
 \text{top-cl} \implies \left[ \begin{array}{l}
 \text{SS | LOC} \\
 \text{HD-DTR} \\
 \text{NON-HD-DTRS}
 \end{array} \right] \left[ \begin{array}{l}
 \text{CAT | HEAD} \\
 \text{CONT} \\
 \left[ \begin{array}{l}
 \text{verb} \\
 \text{INV -} \\
 \text{IC +}
 \end{array} \right] \\
 \boxed{3} \\
 \left[ \text{SS | LOC | CONT } \boxed{3} \right] \\
 \left\langle \left[ \text{WH } \langle \rangle \right] \right\rangle
 \end{array} \right]
 \end{array}$$

While the type *hd-filler-ph* is responsible for the mechanism of discharging and percolating the SLASH members, the constraint in (8.8) above prevents the phrase from being inverted (indicated by [INV -]) and also distinguishes topicalisation from wh-question extraction. Furthermore, the head daughter's CONT is identified with that of the topicalised clause.

The other dimension of the phrasal types is a type hierarchy of CLAUSALITY. Type *clause* is characterised by the following constraint:

$$(8.9) \quad \textit{clause} \Rightarrow \left[ \text{SS} \mid \text{LOC} \mid \text{CONT } \textit{message} \right]$$

The constraint in (8.9) says that all instances of type *clause* must specify the semantic type *message* whose subtypes consist of *proposition*, *directive* etc. Type *clause* is further divided into two subtypes: *core-cl* and *rel-cl*. To distinguish from type *rel-cl*, instances of type *core-cl* are subject to the following constraint:

$$(8.10) \quad \textit{core-cl} \Rightarrow \left[ \text{SS} \mid \text{LOC} \mid \text{CAT} \mid \text{HEAD} \left[ \begin{array}{l} \textit{verb} \\ \text{MOD } \textit{none} \end{array} \right] \right]$$

The condition in (8.10) guarantees that the clause is prohibited from the function of modifying (indicated by [MOD *none*]) and this naturally stops overgenerating relative clauses (for example, *\*the man I like his shoes*). As for the subtypes of type *core-cl*, they have distinctive values for the features MOOD and CONTENT. For example, the following constraint is imposed on instances of type *decl-cl*:

$$(8.11) \quad \textit{decl-cl} \Rightarrow \left[ \begin{array}{l} \text{SS} \mid \text{LOC} \left[ \begin{array}{l} \text{CAT} \mid \text{HEAD} \mid \text{MOOD } \textit{decl} \\ \text{CONT} \left[ \begin{array}{l} \textit{proposition} \\ \text{SOA } \boxed{1} \end{array} \right] \end{array} \right] \\ \text{HD-DTR} \mid \text{SS} \mid \text{LOC} \mid \text{CONT } \boxed{1} \end{array} \right]$$

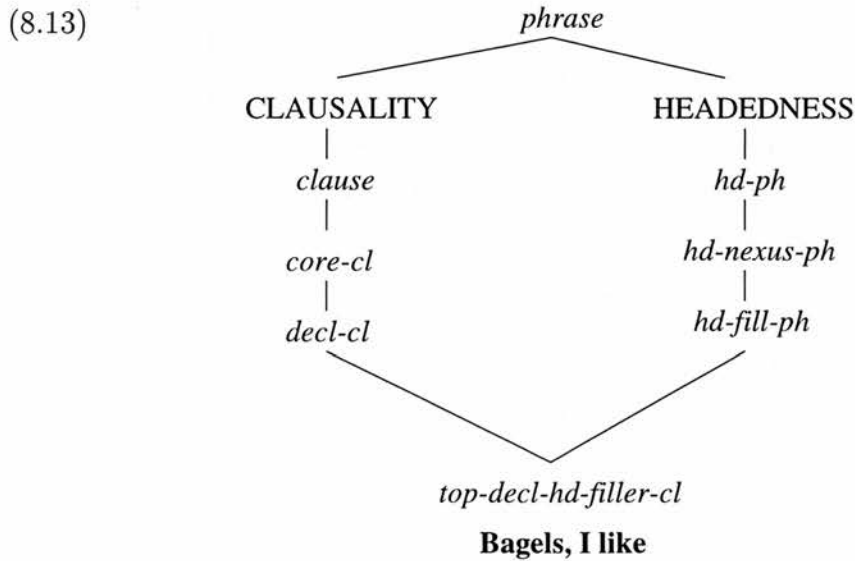
Similarly, other core clausal subtypes would have different values for MOOD and CONTENT. For instance, the constraint for type *inter-cl* will specify the value of MOOD and CONTENT as *indicative* and *question* respectively as follows:



(8.12)

$$inter-cl \Rightarrow \left[ SS \mid LOC \left[ \begin{array}{l} CAT \mid HEAD \mid MOOD \textit{ indic} \\ CONT \textit{ question} \end{array} \right] \right]$$

Having observed briefly the inner structure of the phrasal hierarchy and some constraints on the relevant subtypes, let us have a look at how an instance of topicalisation in English such as ‘Bagels, I like’ is accounted for in terms of its cross-classification and the constraint inheritance given below:



(8.13) above illustrates the case where the topicalised sentence inherits constraints starting from the supertype *phrase* down to the most specific subtype *top-decl-hd-fill-cl*. Identifying the proper position of the type *top-decl-hd-fill-cl* in the phrasal hierarchy through constraint inheritance will have the result that any instance of this type should have the following properties:

(8.14)

<i>top-decl-hd-fill-cl</i>	
SYNSEM	<i>canon-ss</i> SLASH { }
	DEPS < ..., [ <i>canon-ss</i> LOC 5 SLASH { }], ... >
	WH < >
	CONT [0] [ <i>proposition</i> SOA 2]
	HEAD [4] [IC + INV - VFORM <i>fin</i> ]
	SUBJ < >
HD-DTR	<i>phrase</i> SLASH { 5 }
	WH < >
	CONT [0]
	HEAD [4]
	SUBJ < >
NON-HD-DTRS	[LOC 5]
	WH < >

(8.14) shows all the details of the inherited constraints of the relevant types. It declares the relevant features and specifies the correct values. Further, as we have emphasised earlier, all this detailed information does not have to be restated in the lexical entry of each lexical item, for example, *like* in (8.13). Via constraint inheritance, all we have to do is declare the right type in the right position in the hierarchy, and this will avoid restating the information in (8.14) repeatedly in the lexical entry.

Before moving onto the developments necessary to account for Korean topic-binding, in the following section we will revisit the current statement of type *topic-cl* and point out some problems in applying it to languages like Korean. We will revise the current hierarchy to accommodate other types of topicalisations than the English examples, on the basis of which we will account for Korean topic-binding.

## 8.2 A Revised Phrasal Hierarchy:

### TOPICALITY

In the last section, we have seen that English topicalisation is explained as an instance of subtypes of *core-cl* and *hd-fill-ph* in the phrasal type-hierarchy. In addition to the inherited constraints on such supertypes, there are extra constraints on type *topic-cl* which prevent inversion and distinguish it from Wh-questions (see the constraint in (8.8)). Therefore, there is no type of topic clauses independently declared in the phrasal type hierarchy, and topicalisation in English, for example, will be basically handled by an extra constraint which is imposed on the subtypes of *core-cl* and *hd-fill-ph*. This way of relating topicalised clauses to general topic constructions raises two questions.

Firstly, extraction admittedly is a very often used mechanism yielding topic constructions in many languages. This is nonetheless not the only way to realise topicalisation cross-linguistically. In Korean, for example, a topic phrase is morphologically marked with *-nun* which should be also fronted to the sentence initial position. However, we have seen that fronting a topic phrase does not necessarily involve extraction. That is, type *topic-cl* is not always associated with type of *hd-fill-ph*. It can be associated with other phrase structures such as type *hd-subj* where the subject is topic-marked as given below:

- (8.15) John-un ece swukce-lul miri cechwulhayssta  
 John-TOP yesterday homework-ACC in advance submitted  
 ‘As for John<sub>i</sub>, (he<sub>i</sub>) submitted the homework yesterday in advance’

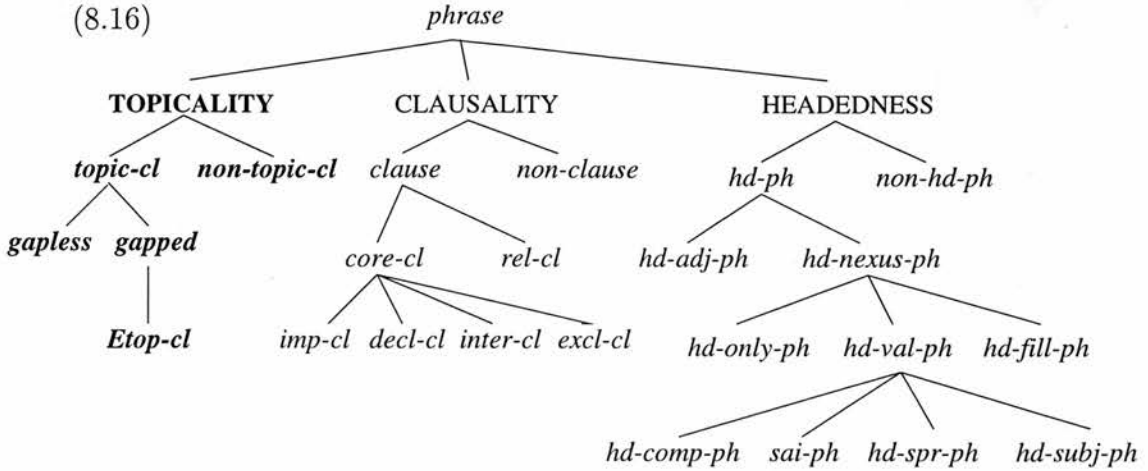
The fact that topicalisation can be associated with phrase structures other than the head-filler phrase is all the more apparent if we recall that, in Korean, the topicalisation types are determined depending on what syntactic structures are associated with them. For example, we saw that the Double Nominative Construction and the head-complement phrase can underly the gapless topic construction and that the head-filler-phrase yields a English kind of topicalisation (gapped topicalisation). Thus, imposing the constraint of *top-cl* only on the subtypes of *core-cl* and *hd-fill-ph* is too restrictive to adequately cover the general topic-constructions cross-linguistically.

Secondly, we would like a single declaration of what characterises ‘topic’, rather than an analysis based on the combination of existing phrasal subtypes. The constraint for *topic-cl* is added to the type resulting from the combination of the subtypes of *core-cl* and *hd-fill-ph*. If we follow the same route to explain Korean topic constructions, we will need at least two extra types of topic constructions: some constraints for *topic-cl* imposed i) on the combination of subtypes of *core-cl* and some types of double nominative constructions (we have not yet proposed what type and what position these should take in the phrasal type-hierarchy), and ii) on the combination of the subtypes of *core-cl* and *hd-comp-ph* for short-distance scrambling topicalisations. However, if we construe different topic constraints for different topic constructions just by combining different subtypes of *clauses* and *phrases*, we run the risk of proposing a number of different topic constraints cross-linguistically and thus failing to capture the universal properties of topicalisation.

As a result, we may also have a number of different types of topic constructions, which neglects the common properties of topicalisation cross-linguistically and hence will certainly miss generalisations about topicalisation.

In order to surmount the problems raised in the current phrase type-hierarchy in accounting for topicalisation in languages other than English, we argue that another partition called TOPICALITY should be posited in parallel with HEADEDNESS and CLAUSALITY. This will add another independent dimension of the constraint on

topicalisations to the phrasal type-hierarchy; the resulting form with the additional partition and its subtypes is shown below:



The first subtypes in TOPICALITY are *top-cl* and *non-top-cl*. These two types are distinguished simply depending on whether there is an empty or nonempty value in TOPIC. The constraints on these two types we state respectively below:

(8.17) Constraint on Non Topic clausal type

$$non-top-cl \Rightarrow \left[ SS | LOC | CONX \left[ TOPIC [ ] \right] \right]$$

(8.18) Constraint on Topic clausal type

$$top-cl \Rightarrow \left[ \begin{array}{l} \left[ \begin{array}{l} CAT | HEAD \quad verb \\ SS | LOC \quad \left[ \begin{array}{l} CONT \quad [2] \\ CONX \quad \left[ TOPIC [1]_{[marking]} \right] \end{array} \right] \end{array} \right] \\ HD-DTR \quad \left[ SS | LOC | CONT [2] \right] \end{array} \right]$$

The constraint on type *top-cl* in (8.18) above simply says that an instance of this type has to be a verbal projection, the CONTENT value should be the same as in the head daughter and most importantly, the TOPIC in the CONX<sup>1</sup> should have one noun

<sup>1</sup>For the detailed status of TOPIC in CONX, see section 7.2.6.

phrase as its value. This last condition on the value of TOPIC contrasts with that in the constraint of type *non-top-cl* where TOPIC must have an empty list as its value. If we recall the constraint for *topic-cl* for English topicalisation we saw in (8.8) in the last section, one can notice the difference. Whereas both (8.18) and (8.8) are constraints on type *top-cl*, (8.18) is now in position as a subtype of TOPICALITY and should be the most general constraint on topicalisation (i.e., not only for English topicalisation). Therefore, the information on features like INV (in (8.8)) which is relevant only for English is not specified in the constraint of the general type *top-cl* in (8.18). The constraint for the English type of topicalisation will only be stated in a subtype specific for English Topicalisation. One last crucial thing to note in the constraints on type *top-cl* in (8.18) is that in CONX, the topic NP is required to be specified with a topic marking indicated as (NP<sub>[marking]</sub>). This constraint is underspecified at this level, which allows different ways the topic phrase itself is realised cross-linguistically. For example, in languages where the topic phrase is morphologically marked, an appropriate topic marker for such languages will be specified as a constraint: [-*nun*] for Korean and [*wa*] for Japanese. On the other hand, in languages where morphological marking is not required, such as in English, the value of [marking] can be simply null and the topic phrase will appear without any morphological marking.

As for the subtypes of type *top-cl*, we propose two: *gapless* and *gapped*. The purpose of having the two types is to distinguish the topicalisation types without the extraction mechanism (*gapless*) from those which involve extraction as in English (*gapped*). Depending on the parametric properties cross-linguistically, these subtypes are further restricted using other dimensional constraints. However, focusing on Korean topicalisation, we will keep to two subtypes: *gapless* and *gapped*.

In the constraints on types *gapless* and *gapped*, the distinction lies in the involvement of extraction and this distinction can be realised in the SLASH value. We propose the following constraints for types *gapless* and *gapped* respectively as below:

(8.19) Constraint on Gapless Topic Clause (Preliminary)

$$\text{gapless-top-cl} \Rightarrow \left[ \begin{array}{l} \left[ \begin{array}{l} \text{HEAD} \quad \textit{verb} \\ \text{CONX} \quad \left[ \text{TOPIC} \quad \boxed{\text{NP}} \right] \\ \text{SLASH} \quad \{ \} \end{array} \right] \\ \text{SS} \mid \text{LOC} \mid \text{CAT} \\ \text{HD-DTR} \mid \text{SS} \mid \text{LOC} \mid \text{CAT} \mid \text{SLASH} \{ \} \end{array} \right]$$

(8.20) Constraint on Gapped Topic Clause

$$\text{gapped-top-cl} \Rightarrow \left[ \begin{array}{l} \left[ \begin{array}{l} \text{CONX} \quad \left[ \text{TOPIC} \quad \boxed{\text{NP}} \right] \\ \text{SLASH} \quad \{ \} \end{array} \right] \\ \text{SS} \mid \text{LOC} \mid \text{CAT} \\ \text{HD-DTR} \quad \left[ \text{SS} \mid \text{LOC} \mid \text{CAT} \mid \text{SLASH} \quad \boxed{\text{I}} \right] \\ \text{NON-HD-DTRS} \quad \left\langle \left[ \text{SS} \mid \text{LOC} \quad \boxed{\text{I}} \right] \right\rangle \end{array} \right]$$

The constraint on type *gapless* in (8.19) guarantees that in its head projection, the SLASH value should be empty and this will exclude all instances with extraction. On the other hand, the constraint on type *gapped* in (8.20) makes sure that the SLASH value is non empty and hence that instances of this type can only be associated with the type *hd-filler-ph*. The constraint in (8.20) is too general to realise both English and Korean types of topicalisation. While the constraint on type *gapped* will suffice in accounting for Korean gapped topicalisation, there are languages where additional constraints are required. For example, as we have seen from the original constraint for English topicalisation as in (8.8), English will need specification of INV and so on. These parametric variations results in two subtypes of type *gapped*. Let us call a subtype of *gapped* for English Topicalisation type *Etop-cl*, the constraint on which should simply be stated as follows:



(8.21) The Constraint for type *Etop-cl*

$$\left[ \begin{array}{l} \text{SS | LOC} \\ \text{NON-HD-DTRS} \end{array} \left[ \begin{array}{l} \text{CAT | HEAD} \\ \langle \left[ \text{WH } \langle \rangle \right] \rangle \end{array} \left[ \begin{array}{l} \textit{verb} \\ \text{INV -} \\ \text{IC +} \end{array} \right] \right] \right]$$

As the conditions for general topic clauses are taken care of already in both types *top-cl* and *gapped-cl*, now the constraint on type *Etop-cl* is minimal in comparison with that in (8.8).

Now let us move onto the types of topicalisation where the mechanism of extraction is not involved. Type *gapless* is the other subtype of type *top-cl*. As the constraint on type *gapless* in (8.19) shows, we simply exclude all instances with extraction by declaring an empty SLASH value on the head daughter. The instances of this type include topicalisation in languages where topicalisation is realised not by extraction, but rather by other methods, for example, morphological marking and fronting as in Korean. Therefore, the constraint for type *gapless* should make sure that whatever is topicalised must appear as the first element in DEPS. In order to implement this idea properly on the constraint on *gapless-top-cl*, we revise the constraint on *gapless-top-cl* in (8.19) as follows:

(8.22) Constraint on Gapless Topic Constructions (Revised)

$$\text{gapless-top-cl} \Rightarrow \left[ \begin{array}{l} \text{SS | LOC | CAT} \\ \text{HEAD} \left[ \begin{array}{l} \text{verb} \\ \text{ARG-ST} \left\langle \dots, \boxed{\text{NP}}_2, \dots \right\rangle \\ \text{DEPS} \left\langle \boxed{\text{NP}}_2, \dots \right\rangle \end{array} \right] \\ \text{CONX} \left[ \text{TOPIC NP}_2 \right] \\ \text{SLASH} \{ \} \end{array} \right] \\ \left[ \text{HD-DTR | SS | LOC | CAT | SLASH} \{ \} \right]$$

Gapless topic constructions involve two types of constructions: short-distance topicalisation and Double Nominative Constructions. Supposing that the topicalised clauses are declarative clauses, in terms of our new phrasal type hierarchy, short-distance topicalisation can be explained by declaring it as type *top-decl-hd-comp-cl*, inheriting constraints from its supertypes *gapless-cl*, *decl-cl* and *hd-comp-ph* (recall that as we have argued that short-distance topicalisation is a case of scrambling, just a matter of word-order swapping, it is treated within the Head-Complement schema; see the section 7.3.3). Furthermore, as pointed out in section 6.3, in (8.22), ARG-ST exists in parallel with DEPS. This is to capture the reordering the topicalised argument to a sentential initial position..

What about the DNC topicalisation? This can be problematic because we have not yet defined what type the DNCs can be. Assuming that the double nominative constructions are a subtype of *hd-val-ph*, there are two options. One is to declare the DNCs as a subtype of *hd-val-ph* in parallel with type *hd-subj-ph* and *hd-comp-ph*, etc. The other is to posit the DNCs as subtype of *subj-hd-ph*. Since the DNCs have very much in common with the *hd-subj-ph* apart from having two elements in the SUBJ value, the latter option is more reasonable and is more consistent with the current type-hierarchy. Let us now call the new subtype of *hd-subj-ph* *dnc-ph*; the

constraint that it is subject to is given as follows:

(8.23) Constraint on DNC phrase (Gill 1999))

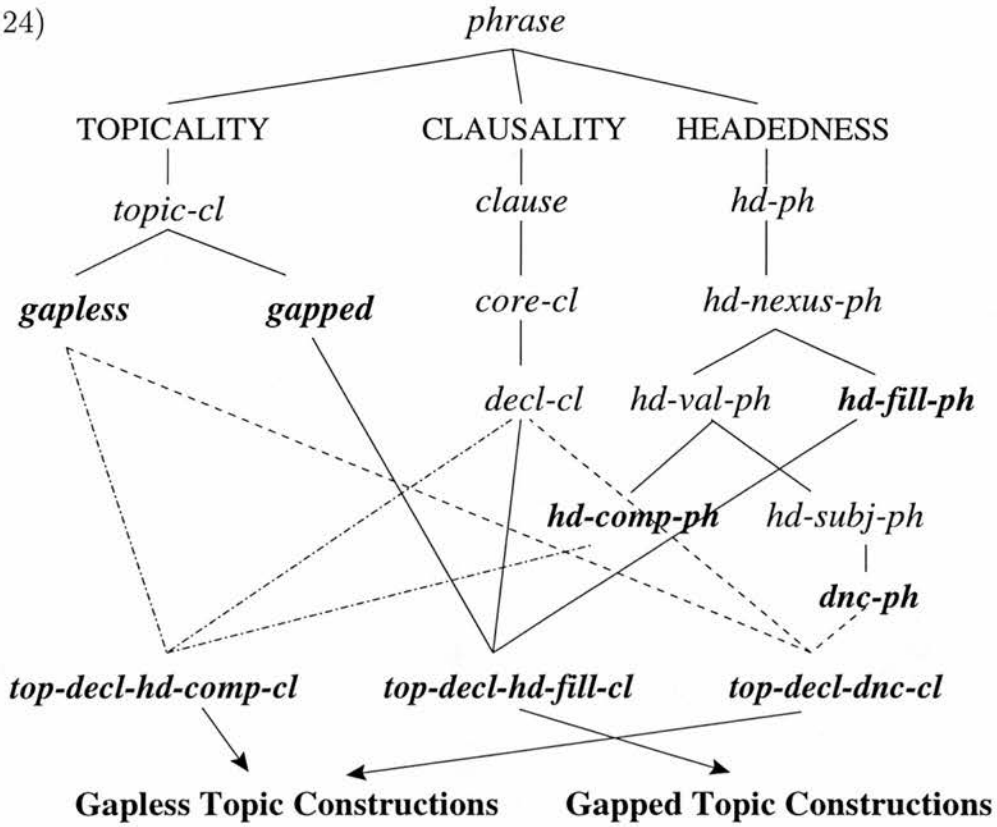
$$\begin{array}{l}
 \left[ \begin{array}{l}
 \text{SS | LOC | CAT | SUBJ } \langle \rangle \\
 \text{HD-DTR } \left[ \begin{array}{l}
 \textit{phrase} \\
 \text{SS | LOC } \left[ \begin{array}{l}
 \text{HEAD } \left[ \begin{array}{l}
 \textit{verb} \\
 \text{DEPS } \langle \boxed{1}_2, \boxed{3}_4 \oplus (\boxed{5}) \rangle \rangle \\
 \text{CAT } \left[ \text{SUBJ } \langle \boxed{1} \rangle, \langle \boxed{3} \rangle \right] \\
 \text{CONT } [ \text{R}[\boxed{2}, \boxed{4}] ]
 \end{array} \right] \\
 \text{NON-HD-DTRS } \langle [ \boxed{1}, \boxed{3} ] \rangle
 \end{array} \right]
 \end{array} \right]
 \end{array}
 \end{array}
 \Rightarrow \text{dnc-ph}
 \end{array}$$

The constraint guarantees that the head daughter looks for two elements in its SUBJ value, and accordingly, there are two non-head daughters whose synsem values are structure-shared with the SUBJ values on the head. The CONT ensures that the two subject NPs hold the R relation. Comparing (8.23) with the constraint on its supertype *hd-subj-ph* in (8.5), one can notice that the only difference is in the number of the non-head-daughters and the SUBJ value and the semantic restrictions holding between the two subject NPs.

On the other hand, long-distance topicalisation can be analysed as a type called *top-decl-hd-fill-cl*, which is the subtype of *gapped-cl*, *decl-cl* and *hd-fill-ph*.

Having added another dimension of constraints on TOPICALITY to the phrasal type-hierarchy and posited a new subtype of *hd-subj-ph*, *dnc-ph*, we propose the final version of the revised phrasal type-hierarchy, omitting some irrelevant types, as follows:

(8.24)



The following three examples demonstrate an instance of each type of topic construction in Korean and the binding relations of *caki*:

(8.25) a. **Type** *top-decl-hd-comp-cl*

Mary<sub>i</sub>-nun caki<sub>i</sub>-ka ceil cohahanta

Mary-TOP self-NOM the most like

‘As for Mary<sub>i</sub>, self<sub>i</sub> likes (her) the most’

b. **Type** *top-decl-dnc-cl*

John<sub>i</sub>-un caki<sub>i</sub>-ka Mary-ekey ku saken-ul selmyenghayssta

John-TOP self-NOM Mary-DAT the accident-ACC explained

‘As for John<sub>i</sub>, self<sub>i</sub> explained the accident to Mary’

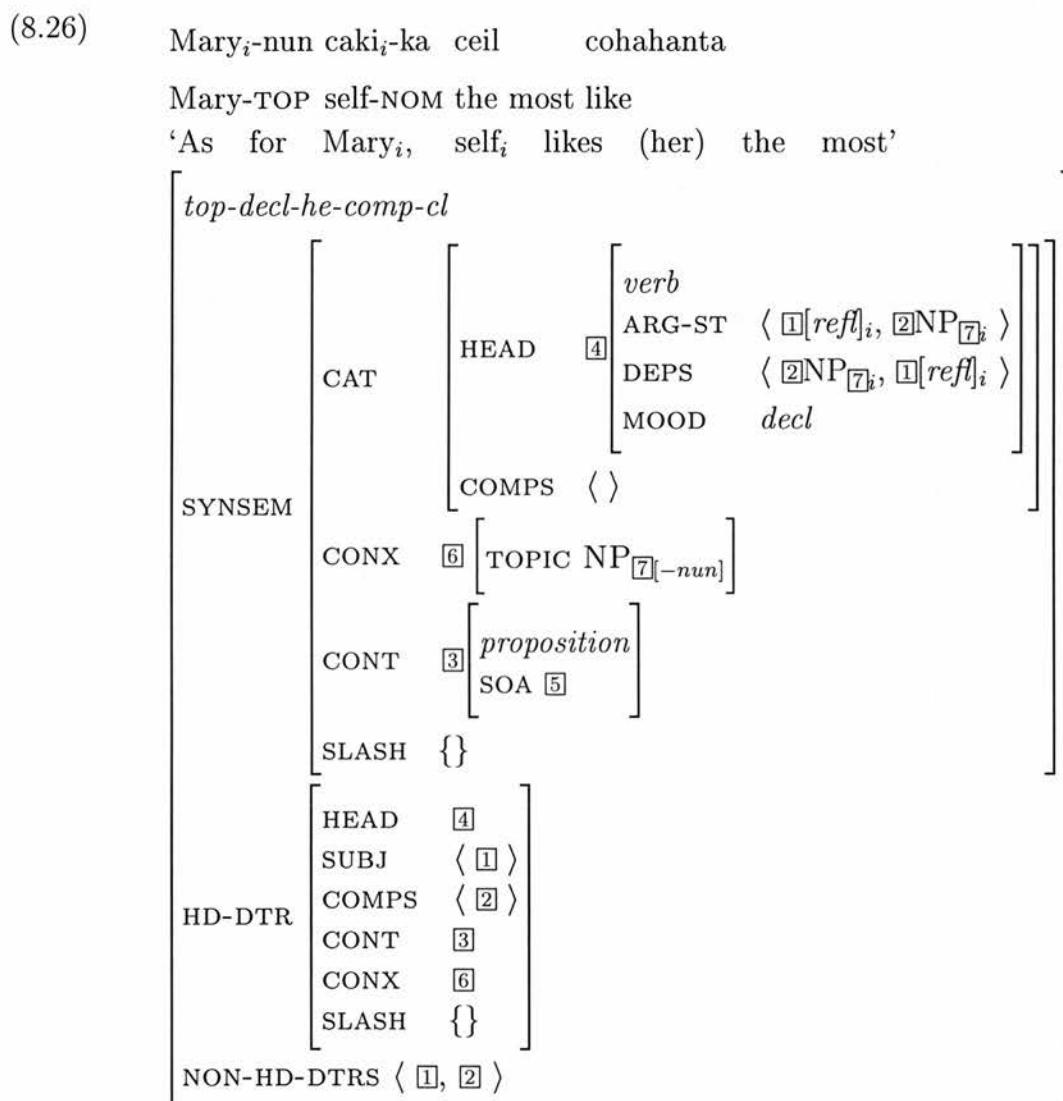
c. **Type** *top-decl-hd-fill-cl*

Mary-nun [John-i [caki-ka cohahanta-ko] malhayssta

Mary-TOP John-NOM self-NOM like-COMPL said

‘Mary, John<sub>i</sub> said that self<sub>i</sub> likes (her)’

For each instance of topicalisation in (8.25), identifying the corresponding positions in the phrasal hierarchy through the path of constraint inheritance indicates that each case above should have the properties shown below, from which we will also check whether each representation provides a correct prediction of the given binding relations. The feature structure below is that of (8.25a):



Since the sentence (8.25a) is an instance of type *gapless*, the SLASH value is empty both on the head daughter and at the phrasal level, to be distinguished from the *gapped* type of topicalisation. Also, as an instance of type *decl*, its MOOD and CONT values are properly marked as *decl* and *proposition* respectively. The constraint on type *sd-top-cl* deals with the representation of the DEPS of the sentence, where the



In the DEPS of the sentence, from the constraint on *dnc-top-cl*, the two SUBJ values appear together as the first elements when *caki*, the second element is correctly bound by the higher element *John*. Therefore, the DEPS value which follows from the multiple inheritance hierarchy predicts the correct binding relations.

Unlike both *sd-top-cl* and *dnc-top-cl*, the Korean types of gapped topicalised clause, Korean gapped topicalisation does not have as restricted a constraint. The constraint on type *gapped* just makes sure that the topicalisation in question involves extraction and the rest of the constraints are handled by type *hd-fill-ph* that the topicalised clause is associated with, as shown below:





### 8.3 Summary and Conclusion

In this chapter, we have provided an account of topic-binding in terms of the multiple inheritance hierarchy. In order to allow the type hierarchy proposed by Ginzburg & Sag (to appear) to accommodate more comprehensive constructions in Korean, we have made the appropriate modifications which are itemised as follows:

- Adding a new type *dnc-ph* as a subtype *hd-subj-ph*
- Adding another dimensional constraint for TOPICALITY in parallel with HEAD-EDNESS and CLAUSALITY.

Firstly, to account for the double nominative constructions, rather than adding a type *dnc-ph* as a subtype of *hd-ph*, we have posited it as a subtype of *hd-subj-ph*. This is because, considering that *dnc-ph* has much in common with *hd-subj-ph*, most of the constraints will be inherited from type *hd-subj-ph* and the constraints imposed on type *dnc-ph* can be minimal. Secondly, rather than having constraints for topicalised clauses declared on the subtypes of *core-cl* and *hd-fill-ph* as proposed for topicalised clauses in English, we have made a modification such that topicalisation has an independent hierarchical constraint which can be cross-classified with other subtypes of *phrase*. From a cross-linguistic view, we argue that our new type-hierarchy provides a more efficient and unified account of topicalisation. Different languages share common properties in their topicalisation such as extraction, but still exhibit different ways of realising topics. Thus we capture the common property of topicalisation across languages and state them as a constraint on a supertype of topicalisation, *top-cl*, which is broken down into different subtypes to handle properties of topicalisation in different languages. For Korean topicalisation, we have two most specific types: *gapless-cl* and *gapped-cl*. These are associated with different syntactic structures from *hd-ph* and account for three types of topic constructions in Korean. We have also seen that the constraint on the DEPS in different topicalised clauses predicts the correct binding relations of *caki*.

Comparing the analysis of topic-binding in terms of lexical rules as proposed in section 7.2.4 and 7.3.3, as far as predicting correct binding relation is concerned,

the multiple inheritance hierarchy is as efficient as the lexical rule. However, the former achieves the following three advantages:

- (8.29)
- a. It provides a unified account for different topic constructions in Korean
  - b. It can be generalised cross-linguistically
  - c. It is more economical

In treating topic-binding using lexical rules, we have different analyses for different kinds of topicalisation. For example, we need lexical rules for short and DNC topicalisation, whereas we can do without them for long-distance topicalisation. On the other hand, with the multiple inheritance hierarchy, the three different topic constructions are explained in a more consistent way by cross-classifying subtypes of *topic-cl* and *hd-ph*. On the same note, the account based on the type-hierarchy contributes to defining universal properties of topicalisation. For example, whereas the lexical rules may be useful in characterising and accounting for the specific types of Korean topicalisation, the type-hierarchy provides a clearer picture of the universal properties of topicalisation, by way of grouping common properties of topicalisation cross-linguistically. This further helps us to capture a truly minimal parametric variation correctly.

From the point of view of generality, the approach with the multiple inheritance hierarchy can be generally applicable to languages other than English and Korean. For example, capturing the similarities and differences of topicalisation across languages, we have provided the two main types *gapped* and *gapless* to handle topicalisation with or without extraction. On the other hand, this is not the case for the analysis with lexical rules, which does not say, or is not relevant to, anything concerning topicalisation in languages other than Korean. In this sense, using the type-hierarchy is a better way to approach topic-binding than using lexical rules.

Lastly, we have seen that lexical entries are assumed to contain a large amount of information. However, in the approach using the multiple inheritance hierarchy, the information that the three different types of topicalisation have to specify becomes

minimal, because the majority of information is already stored in their supertypes. Thus, the information that the three types of topic construction must contribute is the most specific and truly idiosyncratic information of their own type. As for the lexical rule, on the other hand, it procedurally takes a lexical entry as input, and gives another lexical entry as output, thus increases the size of the lexicon. This imposes a considerable burden in the lexicon in comparison with the approach using the multiple inheritance hierarchy, thus, in this sense, the latter provides a more economical and efficient account of topic-binding.

## 8.4 Conclusion

In most approaches to long-distance anaphora, the tendency has been to combine syntactic and discourse factors either conjunctively or disjunctively to explain non-subject binding cases. The reason is that instances of non-subject binding fall outside the straightforward syntactic approaches which are generally designed for subject-binding. As for the case of *caki*-binding, we have excluded the possible involvement of discourse factors by showing that *caki* is not bound at the discourse level, but rather within the sentential domain. In this chapter, we covered instances of topic-binding. We proposed that the part of the data which used to be considered an instance of discourse-binding should be analysed instead as a case of syntactic binding. Even though the notion of topic is generally viewed as a discourse notion, a closer examination of topicalisation in Korean shows that topic-binding can be treated on a par with previously analysed cases of *caki*-binding. Thus it should consistently satisfy our principle at the level of DEPS. This brought together seemingly different patterns of *caki*-binding to be explained with a single principle, which, therefore, provides a more simple but unified account of *caki*-binding.

## Chapter 9

# Summary and Final Remarks

The overarching goal of this thesis was to provide a unified account of Long-distance Anaphora in Korean. We especially focused on the behaviour of the long-distance anaphor *caki*.

There were two main objectives in this work. The first objective was empirical and consisted in establishing the core operative factors licensing *caki*-binding. This objective was achieved by observing the distribution of *caki* occurring in various structural environments, followed by a critical review of approaches to date from the point of view of *caki*-binding. The major findings with respect to the data were that *caki* in simple and complex sentences behaves in a similar way and more importantly, it does not take discourse antecedents. The analysis also revealed that there are regular patterns in both subject and non-subject binding of *caki*. On the other hand, after the critical review of the approaches for long-distance anaphora, we proposed that most approaches are not adequate for *caki*-binding because they fail to provide a unified account for both subject and non-subject binding instances of *caki* at the same time. Constraints of discourse approaches do not seem directly relevant to *caki*-binding as *caki* does not take a discourse antecedent. We concluded that the major factors are structural and not related to discourse.

The second objective of this thesis was theoretical, and consisted in the formulation of a uniform account of the different binding phenomena characteristic of *caki*. The syntactic framework assumed was HPSG where binding relations are determined in a single level of representation within which elements are arranged in terms of obliqueness. For this objective, we provided three major proposals. First, we argued that it is not the notion of obliqueness, but the lexical argument hierarchy (Kiss 1991) that is crucial in determining *caki*-binding patterns. Employing the notion of lexical argument hierarchy provided a unified account for both subject and non-subject binding of *caki*.

Secondly, we dealt with empirical and theoretical problems of ARG-ST. Empirically ARG-ST is not able to account for the mysterious binding patterns of *caki* in adjuncts. Theoretically, ARG-ST coexisting alongside DEPS seems superfluous. We proposed that the two problems can be simultaneously resolved by replacing ARG-ST by DEPS which is made a HEAD feature to account for long-distance binding instances of *caki*. This way, subject, non-subject and adjunct binding relations of *caki* can be represented in one level of representation, DEPS, avoiding any redundancy. Further, it is at the level of DEPS that we proposed our formal binding principle of *caki*-binding.

Thirdly, we focused on the set of data of *caki*-binding which have been believed to be subject to discourse factors. We argued that these data can in fact be explained in terms of structural factors rather than discourse factors. We showed that such binding instances are closely related with topic-orientedness of Korean and can be explained in terms of different topic related constructions such as double nominative constructions and short/long-distance topicalisation. We provided two necessary lexical rules to predict the correct binding relations and to mark the preceding position of topic at DEPS in double nominative constructions and short-distance topic constructions. This was followed by a even more unified account for topic-binding in terms of type-hierarchy. One major extension to the current type-hierarchy we proposed was that we added another dimension of TOPICALITY in

parallel with CLAUSALITY and HEADEDNESS. This not only helps in characterising topicalisation cross-linguistically but also provides an elegant and uniform account of topic-binding without using lexical rules.

There are various avenues for future research suggested by some residual problems as well as the findings themselves. First, in accounting for the topic binding instances in long-distance binding, we encountered a theory internal problem in blocking the binding between *caki* and *gap-ss*. We proposed some possible explanations: one is using Tsoulas's (1999) approach where in HPSG terms, *gap-ss* should be viewed as *pro*. The other is to do with the fact that *caki* lacks a full specification in  $\phi$ -features and the treatment we proposed was via the modification of the internal structure of *gap-ss*. Whatever the source of the problems may be, in both possible explanations, we tried to amend the properties of *gap-ss* and there are reasons why this may not be ultimately desirable. A part of the reason is that *gap-ss* itself in the current HPSG still remains a primitive notion. On the other hand, most of the work on long-distance anaphora only focuses on the antecedents and how they are determined rather than the properties of the anaphoric elements themselves. In a lexical framework, this makes sense though and we hope to undertake such a project in the future.

Second, at the outset of this thesis, we set a question of what universal principle underlies long-distance anaphora. In this thesis, we investigated Korean long-distance anaphora as an element of the answer to this question. We argued that as far as *caki*-binding is concerned, it is not subject to any discourse factors and proposed that *caki*-binding can be explained by structural and formal constraints without relying on involving any discourse approaches. However, long-distance anaphora in other languages like Japanese and Chinese may indeed involve such factors. In this thesis, there were several places where we compared *caki*-binding patterns with *zibun*-binding. From the observation of *zibun*-binding patterns, it seems that *zibun* is bound within a larger, discourse, domain and is crucially subject to discourse conditions. The differences can be summarised as follows:



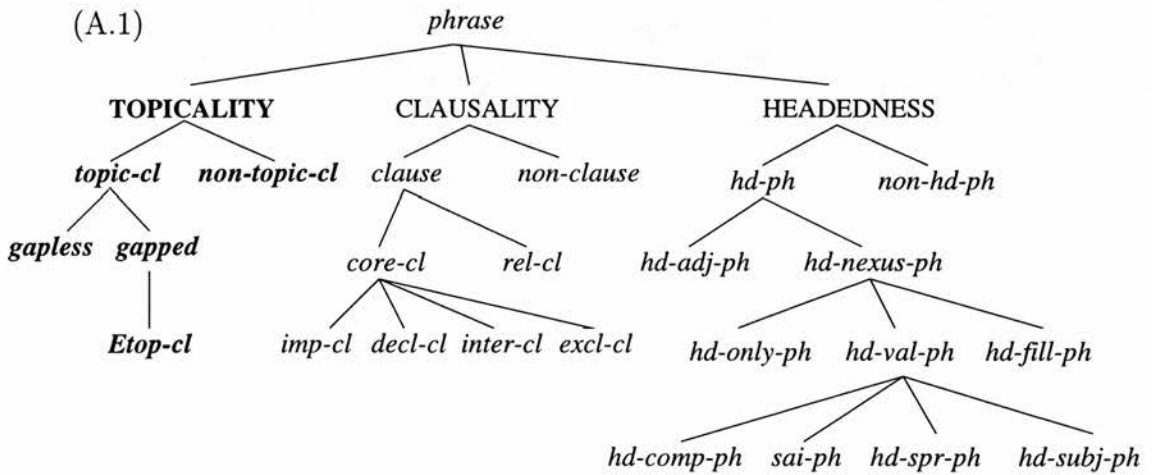
- (9.1) a. *Zibun* can be bound by a speaker or by a first person pronoun as shown in (3.44) and (3.49)
- i. **Zibun<sub>i</sub>**-ga ikimasu  
 self-NOM go  
 ‘I will go’
- b. Pragmatics seems to play a role in determining the binding patterns of *zibun*. Pragmatic force seems to play a role here as for example, concessive and temporal adjunct clauses display different binding patterns (a fact which has been shown to be irreducible to structural differences). We have seen examples in (2.22) and (3.53)
- c. Functional categories make difference in *zibun*-binding relation. Different complementizers, tense and aspect markers seem to interact with *zibun*-binding.

Apart from the differences between *caki* and *zibun* to do with their lexical properties it is interesting to observe that most of the factors active in *zibun*-binding seem inert in *caki*-binding. *Why?* If I had more time, I would like to pursue this question as the next logical step from here. However, time is running out and I have to leave this question “for further study”. I hope I will soon return to it.



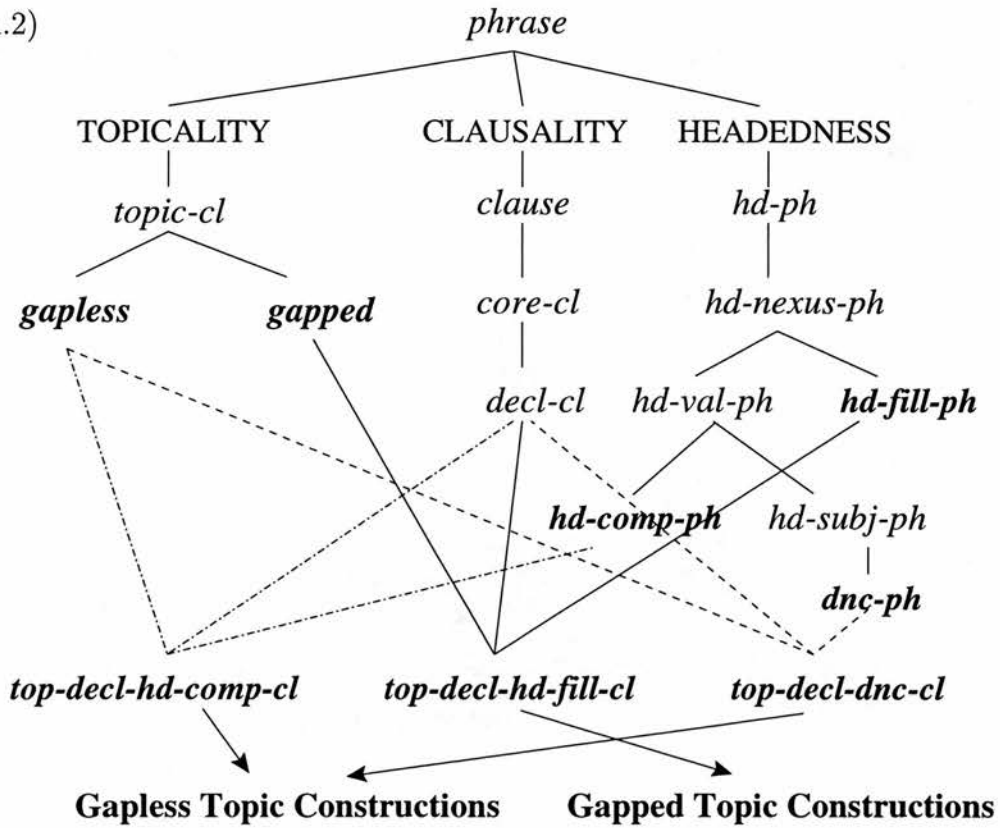
# Appendix A

## List of Type-Hierarchy and Constraints



(As in 8.16)

(A.2)



(As in 8.24)

(A.3)

Constraint on Non Topic clausal type

$$non-top-cl \Rightarrow \left[ SS \mid LOC \mid CONX \left[ TOPIC \square \right] \right]$$

(As in 8.17)

(A.4) Constraint on Topic clausal type

$$top-cl \Rightarrow \left[ \begin{array}{l} \left[ \begin{array}{l} CAT | HEAD \quad verb \\ CONT \quad \boxed{2} \\ CONX \quad \left[ \begin{array}{l} TOPIC \boxed{1}_{[marking]} \end{array} \right] \end{array} \right] \\ HD-DTR \quad \left[ \begin{array}{l} SS | LOC | CONT \boxed{2} \end{array} \right] \end{array} \right]$$

(As in 8.18)

(A.5) Constraint on Gapless Topic Constructions

$$gapless-top-cl \Rightarrow \left[ \begin{array}{l} \left[ \begin{array}{l} \left[ \begin{array}{l} HEAD \quad \left[ \begin{array}{l} verb \\ ARG-ST \quad \langle \dots, \boxed{1}NP_{\boxed{2}}, \dots \rangle \\ DEPS \quad \langle \boxed{1}NP_{\boxed{2}}, \dots \rangle \end{array} \right] \\ CONX \quad \left[ \begin{array}{l} TOPIC NP_{\boxed{2}} \end{array} \right] \\ SLASH \quad \{ \} \end{array} \right] \\ HD-DTR | SS | LOC | CAT | SLASH \{ \} \end{array} \right]$$

(As in 8.22)

(A.6)

Constraint on Gapped Topic Clause

$$gapped-top-cl \Rightarrow \left[ \begin{array}{l} SS | LOC | CAT \quad \left[ \begin{array}{l} CONX \quad \left[ \begin{array}{l} TOPIC \quad \boxed{NP} \end{array} \right] \\ SLASH \quad \{ \} \end{array} \right] \\ HD-DTR \quad \left[ \begin{array}{l} SS | LOC | CAT | SLASH \quad \boxed{I} \end{array} \right] \\ NON-HD-DTRS \quad \left\langle \left[ \begin{array}{l} SS | LOC \quad \boxed{I} \end{array} \right] \right\rangle \end{array} \right]$$

(As in 8.20)

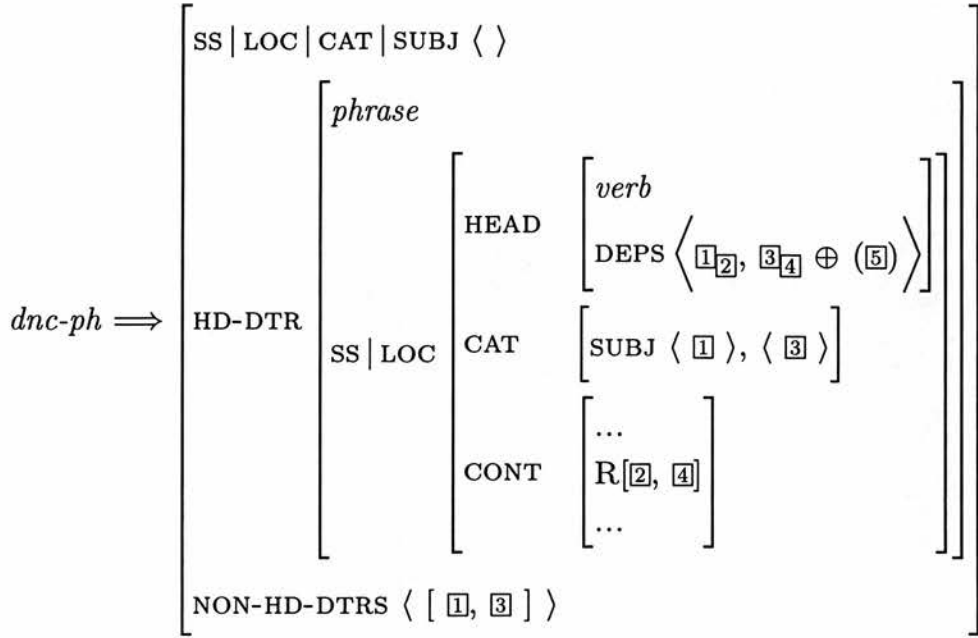
(A.7) The Constraint for type *Etop-cl*

$$\left[ \begin{array}{l} SS | LOC \quad \left[ \begin{array}{l} CAT | HEAD \quad \left[ \begin{array}{l} verb \\ INV \quad - \\ IC \quad + \end{array} \right] \end{array} \right] \\ NON-HD-DTRS \quad \left\langle \left[ \begin{array}{l} WH \quad \langle \rangle \end{array} \right] \right\rangle \end{array} \right]$$

(As in 8.21)

(A.8)

Constraint on DNC phrase



(As in 8.23)

## Examples with Feature Structures

(A.9)

Mary<sub>i</sub>-nun caki<sub>i</sub>-ka ceil cohahanta

Mary-TOP self-NOM the most like

'As for Mary<sub>i</sub>, self<sub>i</sub> likes (her) the most'

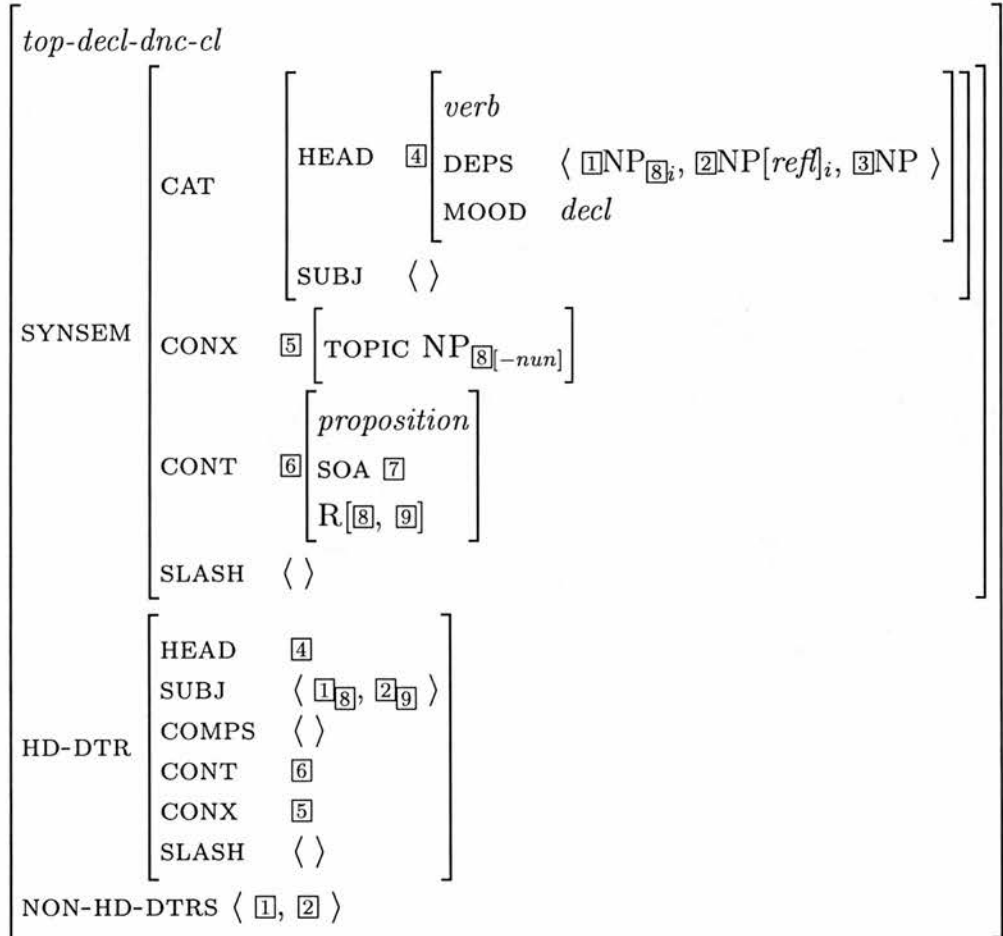
<i>top-decl-he-comp-cl</i>																																							
SYNSEM	<table style="border-collapse: collapse; margin-left: 20px;"> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">CAT</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;"> <table style="border-collapse: collapse; margin-left: 20px;"> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">HEAD</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;"> <table style="border-collapse: collapse; margin-left: 20px;"> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">[4]</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;"> <table style="border-collapse: collapse; margin-left: 20px;"> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;"><i>verb</i></td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;"></td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">ARG-ST</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">⟨ [1]<i>refl</i><sub>i</sub>, [2]NP<sub>[7]<sub>k</sub></sub> ⟩</td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">DEPS</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">⟨ [2]NP<sub>[7]<sub>k</sub></sub>, [1]<i>refl</i><sub>i</sub> ⟩</td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">MOOD</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;"><i>decl</i></td> </tr> </table> </td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">COMPS</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">⟨ ⟩</td> </tr> </table> </td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">CONX</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">[6] [TOPIC NP<sub>[7]<sub>k</sub></sub>[-<i>nun</i>]]</td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">CONT</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">[3] [<i>proposition</i> SOA [5]]</td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">SLASH</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">{ }</td> </tr> </table> </td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px; vertical-align: middle;">HD-DTR</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;"> <table style="border-collapse: collapse; margin-left: 20px;"> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">HEAD</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">[4]</td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">SUBJ</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">⟨ [1] ⟩</td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">COMPS</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">⟨ [2] ⟩</td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">CONT</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">[3]</td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">CONX</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">[6]</td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">SLASH</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">{ }</td> </tr> </table> </td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">NON-HD-DTRS</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">⟨ [1], [2] ⟩</td> </tr> </table>	CAT	<table style="border-collapse: collapse; margin-left: 20px;"> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">HEAD</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;"> <table style="border-collapse: collapse; margin-left: 20px;"> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">[4]</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;"> <table style="border-collapse: collapse; margin-left: 20px;"> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;"><i>verb</i></td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;"></td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">ARG-ST</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">⟨ [1]<i>refl</i><sub>i</sub>, [2]NP<sub>[7]<sub>k</sub></sub> ⟩</td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">DEPS</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">⟨ [2]NP<sub>[7]<sub>k</sub></sub>, [1]<i>refl</i><sub>i</sub> ⟩</td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">MOOD</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;"><i>decl</i></td> </tr> </table> </td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">COMPS</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">⟨ ⟩</td> </tr> </table> </td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">CONX</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">[6] [TOPIC NP<sub>[7]<sub>k</sub></sub>[-<i>nun</i>]]</td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">CONT</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">[3] [<i>proposition</i> SOA [5]]</td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">SLASH</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">{ }</td> </tr> </table>	HEAD	<table style="border-collapse: collapse; margin-left: 20px;"> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">[4]</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;"> <table style="border-collapse: collapse; margin-left: 20px;"> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;"><i>verb</i></td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;"></td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">ARG-ST</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">⟨ [1]<i>refl</i><sub>i</sub>, [2]NP<sub>[7]<sub>k</sub></sub> ⟩</td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">DEPS</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">⟨ [2]NP<sub>[7]<sub>k</sub></sub>, [1]<i>refl</i><sub>i</sub> ⟩</td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">MOOD</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;"><i>decl</i></td> </tr> </table> </td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">COMPS</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">⟨ ⟩</td> </tr> </table>	[4]	<table style="border-collapse: collapse; margin-left: 20px;"> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;"><i>verb</i></td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;"></td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">ARG-ST</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">⟨ [1]<i>refl</i><sub>i</sub>, [2]NP<sub>[7]<sub>k</sub></sub> ⟩</td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">DEPS</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">⟨ [2]NP<sub>[7]<sub>k</sub></sub>, [1]<i>refl</i><sub>i</sub> ⟩</td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">MOOD</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;"><i>decl</i></td> </tr> </table>	<i>verb</i>		ARG-ST	⟨ [1] <i>refl</i> <sub>i</sub> , [2]NP <sub>[7]<sub>k</sub></sub> ⟩	DEPS	⟨ [2]NP <sub>[7]<sub>k</sub></sub> , [1] <i>refl</i> <sub>i</sub> ⟩	MOOD	<i>decl</i>	COMPS	⟨ ⟩	CONX	[6] [TOPIC NP <sub>[7]<sub>k</sub></sub> [- <i>nun</i> ]]	CONT	[3] [ <i>proposition</i> SOA [5]]	SLASH	{ }	HD-DTR	<table style="border-collapse: collapse; margin-left: 20px;"> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">HEAD</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">[4]</td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">SUBJ</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">⟨ [1] ⟩</td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">COMPS</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">⟨ [2] ⟩</td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">CONT</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">[3]</td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">CONX</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">[6]</td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">SLASH</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">{ }</td> </tr> </table>	HEAD	[4]	SUBJ	⟨ [1] ⟩	COMPS	⟨ [2] ⟩	CONT	[3]	CONX	[6]	SLASH	{ }	NON-HD-DTRS	⟨ [1], [2] ⟩
CAT	<table style="border-collapse: collapse; margin-left: 20px;"> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">HEAD</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;"> <table style="border-collapse: collapse; margin-left: 20px;"> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">[4]</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;"> <table style="border-collapse: collapse; margin-left: 20px;"> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;"><i>verb</i></td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;"></td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">ARG-ST</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">⟨ [1]<i>refl</i><sub>i</sub>, [2]NP<sub>[7]<sub>k</sub></sub> ⟩</td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">DEPS</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">⟨ [2]NP<sub>[7]<sub>k</sub></sub>, [1]<i>refl</i><sub>i</sub> ⟩</td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">MOOD</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;"><i>decl</i></td> </tr> </table> </td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">COMPS</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">⟨ ⟩</td> </tr> </table> </td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">CONX</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">[6] [TOPIC NP<sub>[7]<sub>k</sub></sub>[-<i>nun</i>]]</td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">CONT</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">[3] [<i>proposition</i> SOA [5]]</td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">SLASH</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">{ }</td> </tr> </table>	HEAD	<table style="border-collapse: collapse; margin-left: 20px;"> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">[4]</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;"> <table style="border-collapse: collapse; margin-left: 20px;"> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;"><i>verb</i></td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;"></td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">ARG-ST</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">⟨ [1]<i>refl</i><sub>i</sub>, [2]NP<sub>[7]<sub>k</sub></sub> ⟩</td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">DEPS</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">⟨ [2]NP<sub>[7]<sub>k</sub></sub>, [1]<i>refl</i><sub>i</sub> ⟩</td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">MOOD</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;"><i>decl</i></td> </tr> </table> </td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">COMPS</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">⟨ ⟩</td> </tr> </table>	[4]	<table style="border-collapse: collapse; margin-left: 20px;"> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;"><i>verb</i></td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;"></td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">ARG-ST</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">⟨ [1]<i>refl</i><sub>i</sub>, [2]NP<sub>[7]<sub>k</sub></sub> ⟩</td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">DEPS</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">⟨ [2]NP<sub>[7]<sub>k</sub></sub>, [1]<i>refl</i><sub>i</sub> ⟩</td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">MOOD</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;"><i>decl</i></td> </tr> </table>	<i>verb</i>		ARG-ST	⟨ [1] <i>refl</i> <sub>i</sub> , [2]NP <sub>[7]<sub>k</sub></sub> ⟩	DEPS	⟨ [2]NP <sub>[7]<sub>k</sub></sub> , [1] <i>refl</i> <sub>i</sub> ⟩	MOOD	<i>decl</i>	COMPS	⟨ ⟩	CONX	[6] [TOPIC NP <sub>[7]<sub>k</sub></sub> [- <i>nun</i> ]]	CONT	[3] [ <i>proposition</i> SOA [5]]	SLASH	{ }																		
HEAD	<table style="border-collapse: collapse; margin-left: 20px;"> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">[4]</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;"> <table style="border-collapse: collapse; margin-left: 20px;"> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;"><i>verb</i></td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;"></td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">ARG-ST</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">⟨ [1]<i>refl</i><sub>i</sub>, [2]NP<sub>[7]<sub>k</sub></sub> ⟩</td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">DEPS</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">⟨ [2]NP<sub>[7]<sub>k</sub></sub>, [1]<i>refl</i><sub>i</sub> ⟩</td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">MOOD</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;"><i>decl</i></td> </tr> </table> </td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">COMPS</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">⟨ ⟩</td> </tr> </table>	[4]	<table style="border-collapse: collapse; margin-left: 20px;"> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;"><i>verb</i></td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;"></td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">ARG-ST</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">⟨ [1]<i>refl</i><sub>i</sub>, [2]NP<sub>[7]<sub>k</sub></sub> ⟩</td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">DEPS</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">⟨ [2]NP<sub>[7]<sub>k</sub></sub>, [1]<i>refl</i><sub>i</sub> ⟩</td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">MOOD</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;"><i>decl</i></td> </tr> </table>	<i>verb</i>		ARG-ST	⟨ [1] <i>refl</i> <sub>i</sub> , [2]NP <sub>[7]<sub>k</sub></sub> ⟩	DEPS	⟨ [2]NP <sub>[7]<sub>k</sub></sub> , [1] <i>refl</i> <sub>i</sub> ⟩	MOOD	<i>decl</i>	COMPS	⟨ ⟩																										
[4]	<table style="border-collapse: collapse; margin-left: 20px;"> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;"><i>verb</i></td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;"></td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">ARG-ST</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">⟨ [1]<i>refl</i><sub>i</sub>, [2]NP<sub>[7]<sub>k</sub></sub> ⟩</td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">DEPS</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">⟨ [2]NP<sub>[7]<sub>k</sub></sub>, [1]<i>refl</i><sub>i</sub> ⟩</td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">MOOD</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;"><i>decl</i></td> </tr> </table>	<i>verb</i>		ARG-ST	⟨ [1] <i>refl</i> <sub>i</sub> , [2]NP <sub>[7]<sub>k</sub></sub> ⟩	DEPS	⟨ [2]NP <sub>[7]<sub>k</sub></sub> , [1] <i>refl</i> <sub>i</sub> ⟩	MOOD	<i>decl</i>																														
<i>verb</i>																																							
ARG-ST	⟨ [1] <i>refl</i> <sub>i</sub> , [2]NP <sub>[7]<sub>k</sub></sub> ⟩																																						
DEPS	⟨ [2]NP <sub>[7]<sub>k</sub></sub> , [1] <i>refl</i> <sub>i</sub> ⟩																																						
MOOD	<i>decl</i>																																						
COMPS	⟨ ⟩																																						
CONX	[6] [TOPIC NP <sub>[7]<sub>k</sub></sub> [- <i>nun</i> ]]																																						
CONT	[3] [ <i>proposition</i> SOA [5]]																																						
SLASH	{ }																																						
HD-DTR	<table style="border-collapse: collapse; margin-left: 20px;"> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">HEAD</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">[4]</td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">SUBJ</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">⟨ [1] ⟩</td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">COMPS</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">⟨ [2] ⟩</td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">CONT</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">[3]</td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">CONX</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">[6]</td> </tr> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">SLASH</td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 5px;">{ }</td> </tr> </table>	HEAD	[4]	SUBJ	⟨ [1] ⟩	COMPS	⟨ [2] ⟩	CONT	[3]	CONX	[6]	SLASH	{ }																										
HEAD	[4]																																						
SUBJ	⟨ [1] ⟩																																						
COMPS	⟨ [2] ⟩																																						
CONT	[3]																																						
CONX	[6]																																						
SLASH	{ }																																						
NON-HD-DTRS	⟨ [1], [2] ⟩																																						

(As in 8.26)

(A.10) John<sub>i</sub>-un caki<sub>i</sub>-ka Mary-ekey ku saken-ul  
 John-TOP self-NOM Mary-DAT the accident-ACC  
 selmyenghayssta

explained

‘As for John<sub>i</sub>, self<sub>i</sub> explained the accident to Mary’



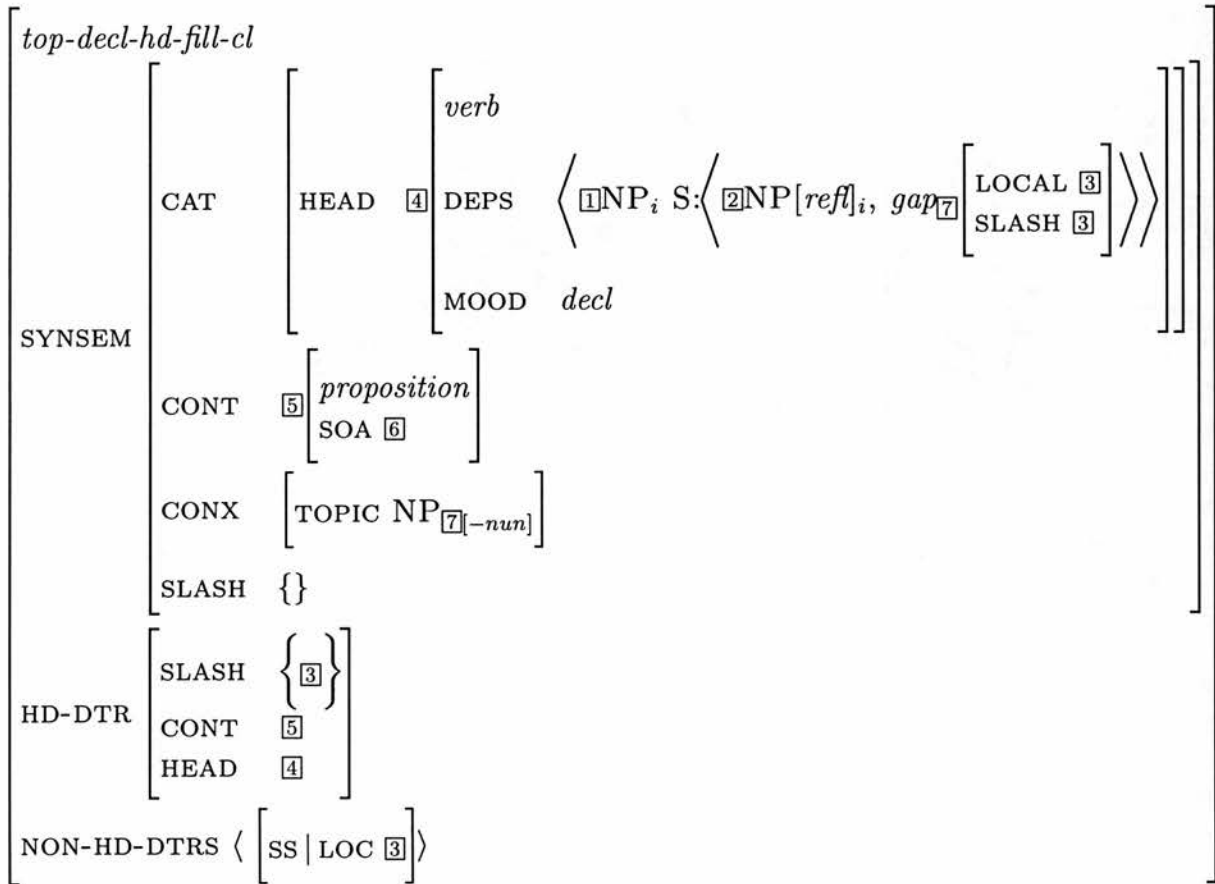
(As in 8.27)

(A.11)

Mary-nun [John-i [caki-ka cohahanta-ko]  
 Mary-TOP John-NOM self-NOM like-COMPL  
 malhayssta

said

'Mary, John<sub>i</sub> said that self<sub>i</sub> likes (her)'



(As in 8.28)



# References

- Akatsuka N (1971). Psych movment in Japanese and some crucialy related syntactic phenomena, *in Working Papers in Linguistics*, Vol. 10, Ohio State University, pp. 8–34.
- Alexopoulou T (1999), The Syntax of Discourse Functions in Greek: a Non-Configurational Approach, PhD thesis, University of Edinburgh.
- Alsina A (1993), Predicate Composition: A Theory of Syntactic Function Alternations, PhD thesis, Stanford University.
- Alsina A (1996). *The Role of Argument Structure in Grammar*, Stanford: CSLI.
- Arka I W (1998), From morphosyntax to pragmatics in Balinese, PhD thesis, University of Sydney.
- Banfield A (1982). *Unspeakable Sentence*, Boston: RKP.
- Belletti A & Rizzi L (1988). Psych-Verbs and  $\Theta$ -Theory, *Natural Language and Linguistic Theory* 6, 291–352.
- Borsley R (1987). Subjects and Complements in HPSG, Technical report, Stanford: CSLI. Technical Report no. CSLI-107-87.
- Bouchard D (1984). *On the content of empty categories*, Dordrecht:Foris.
- Bouma G, Malouf R & Sag I (1998). A unified theory of complement, adjunct and subject extraction, *in Proceedings of the joint conference on Formal Grammar, Head-Driven Phrase Structure Grammar, and Categorical Grammar*, Saabrucken, pp. 84–97.

- Bredekamp A (1996), Towards a Binding Theory for Head-driven Phrase Structure (G)rammar, PhD thesis, University of Essex.
- Bresnan J (1978). A Realistic Transformational Grammar, *in* M Halle, J Bresnan & G A Miller, eds, *The Linguistic Theory and Psychological Reality*, Cambridge, Mass.: MIT Press, pp. 1–59.
- Bresnan J & Kanerva J M (1989). Locative Inversion in Chicheŵa, *Linguistic Inquiry* **20**(1), 1–50.
- Bresnan J & Zaenen A (1990). Deep Unaccusativity in LFG, *in* P K. Dziwirek, Farrell & E Eejias-Bikande, eds, *Grammatical Relations: A Cross-Theoretical Perspective*, Stanford, CA: CSLI Publications, pp. 45–57.
- Chang S J (1995). *Korean*, Amsterdam: John Benjamins Publishing Company.
- Choi H W (1999). *Optimizing Structure in Context: Scrambling and Information Structure*, CSLI: Stanford.
- Chomsky N (1980). On Binding, *Linguistic Inquiry* **11**, 1–46.
- Chomsky N (1981). *Lectures on Government and Binding.*, Dordrecht:Foris.
- Chomsky N (1982). *Some Concepts and Consequences of the Theory of Government and Binding Theory*, Cambridge, Mass.: MIT Press.
- Chomsky N (1986). *Knowledge of Language: Its nature, Origin, and Use*, New York: Praeger.
- Chomsky N (1993). A minimalist program for linguistic theory, *in* K Hale & S J Keyser, eds, *The view from Binding 20*, Cambridge, Mass.: MIT Press.
- Chomsky N (1995). *The Minimalist Program*, Cambridge, Mass.: MIT Press.
- Chung C (1996), Prominence of An Antecedent and Its Effect on Korean Caki-Binding. Ms. Dongseo University, Korea.
- Cole P, Hermon G & Sung L M (1990). Principles and Parameters of Long-Distance Reflexives, *Linguistic Inquiry* **21**(1), 1–22.
- Cole P & Sung L M (1994), Principles and Parameters of long distance reflexives. Unpublished manuscript, University of Delaware.
- Copestake A, Flickinger D & Sag I A (1997), Minimal Recursion Semantics: an introduction. Available at <ftp://csli-ftp.stanford.edu/linguistics/sag/mrs.ps.gz>.
- Dalrymple M (1993). *The syntax of anaphoric binding*, Stanford, CA: CSLI.

- Davis A (1997), Lexical Semantics and linking in the hierarchical lexicon, PhD thesis, Stanford University.
- di Sciullo A M & Williams E (1987). *On the Definition of Word*, Cambridge, MA.: MIT Press. Linguistic Inquiry Monograph 14.
- Doron E & Heycock C (1999). Filling and Licensing Multiple Specifiers, in D Adger, S Pintzuk, B Plunkett & G Tsoulas, eds, *Specifiers: Minimalist Approaches*, OUP, pp. 69–89.
- Eynde F V (1995). An HPSG treatment of *it* extraposition without lexical rules, *CLIN VI: Papers from the sixth CLIN meeting* pp. 231–48.
- Fiengo R & Kim H (1981). Binding and (C)ontrol in (K)orean: (S)tructural (R)estrictions on (A)naphors in a (N)on-configurational (L)anguage, in *Ene*, Linguistic Society of Korea: Seoul.
- Fukui N (1984), Studies in Japanese Anaphora. Ms. Massachusetts Institute of Technology.
- Gazdar G, Klein E, Pullum G & Sag I (1985). *Generalized Phrase Structure Grammar*, Harvard University Press: Cambridge, Mass.
- Gerds D B (1985). Surface Case and Grammatical Relations in Korean: The evidence from Quantifier Float, in S K et al., ed., *Harvard Studies in Korean Linguistics*, Vol. 1, Seoul: Hansin Publishing Company, pp. 48–61.
- Ginzburg J & Sag I A (to appear). *English Interrogative Constructions*, CSLI Publications.
- Giorgi A (1991). Prepositions, Binding and  $\theta$ -marking, in J Koster & E Reuland, eds, *Long-distance anaphora*, Cambridge: Cambridge University Press, pp. 185–208.
- Givón T (1975). Serial Verbs and Syntactic Change: Niger-Congo, in *Word Order and Word Order Change*, Austin: University of Texas Press, pp. 47–112.
- Grimshaw J (1990). *Argument Structure*, MA: MIT Press.
- Gunji T (1987). *Japanese Phrase Structure Grammar: A Unification Based Approach*, Dordrecht: Reidal.
- Hasegawa N (1984), A Lexical Interpretive Theory with Emphasis on the Role of Subject. ms. University of Massachusetts, Amherst.

- Hellan L (1988). *Anaphora in Norwegian and the Theory of Grammar*, Foris Publications Holland: Dordrecht.
- Hellan L (1991). Containment and Connectedness Anaphors, in J Koster & E Reuland, eds, *Long-distance anaphora*, Cambridge: Cambridge University Press, pp. 27–48.
- Heycock C & Lee Y S (1990). Subjects and Predication in Korean and Japanese, in H Hoji, ed., *Japanese/Korean Linguistics*, SLA:Stanford, pp. 239–254.
- Hoji H (1985), Logical Form Constraints and Configurational Structures in Japanese, PhD thesis, University of Washington.
- Hong S (1985), A and A' Binding in Korean and English: Government-Binding Parameters, PhD thesis, The University of Connecticut.
- Howard I & Neyekawa-Howard A M (1976). Passivization, in M Shibatani, ed., *Syntax and Semantics 5: Japanese Generative Grammar*, New York: Academic Press, pp. 201–238.
- Huang C T J (1984). On the Distribution and Reference of Empty Pronouns, *Linguistic Inquiry* 15(4), 531–574.
- Huang Y (1994). *The Syntax and Pragmatics of Anaphora*, CUP: Cambridge.
- Iida M (1996). *Context and Binding in Japanese*, CSLI:Stanford.
- Imai T (1983). On the significance of 'Move-alph' and Empty Categories in Non-Configurational Languages, in Y Otsu, ed., ???, ???
- Jackendoff R (1972). *Semantic Interpretation in Generative Grammar*, MIT Press, Cambridge, Massachusetts.
- Jespersen O (1924). *Philosophy of Grammar*, London: George Allen and Unwin.
- Kameyama M (1984). Subject/Logophoric Bound Anaphor *Zibun*, in *Proceeding of CLS*, Vol. 20, pp. 228–238.
- Katada F (1990). The LF representation of Anaphors, *Linguistic Inquiry* 22(2), 287–313.
- Kathol A (1994). Passives without lexical rules, in J Nerbonne, K Netter & C Pollard, eds, *German in Head-Driven Phrase Structure Grammar*, CSLI Publications, pp. 237–272.

- Kim A H (1985). The Functions of Linear Order in Korean Syntax, *in* Kuno, ed., *Harvard Studies in Korean Linguistics I*, Hanshin Publishing Company, pp. 154–167.
- Kim S H (1993), Division of labor between Grammar and Pragmatics, PhD thesis, Yale University.
- Kiss K (1991). The Primacy Condition of Anaphora and Pronominal Variable binding, *in* K Koster & E Reuland, eds, *Long-Distance Anaphora*, CUP:Cambridge.
- Koenig J P & Jurafsky D (1995). Type underspecification and on-line type construction in the lexicon, *in Proceedings of the Thirteenth West Coast Conference on Formal Linguistics*, Stanford: CSLI.
- Krieger H U & Nerbonne J (1993). Feature-based inheritance networks for computational lexicons, *in* T Briscoe, V de Paiva & A Copestake, eds, *Inheritance, Defaults, and the Lexicon*, Cambridge: Cambridge University Press, pp. 90–136.
- Kuno S (1972). Pronominalization, Reflexivization and Direct Discourse, *Linguistic Inquiry* 3, 161–195.
- Kuno S (1973). *The Structure of the Japanese Language*, Cambridge, Massachusetts: MIT Press.
- Kuno S (1976). Subject, Theme, and the Speaker's Empathy, *in* C N Li, ed., *Subject and Topic*, New York: Academic Press.
- Kuno S (1978). *Danwa no Bunpoo ('Grammar of Discourse')*, Tokyo: Taisyuukan.
- Kuno S (1987). *Functional Syntax*, Chicago:University of Chicago Press.
- Kural M & Tsoulas G (1999), Pronouns in Syntax and Discourse. ms. InQuiZit Technologies and University of York.
- Kuroda S Y (1965), Generative Grammatical Studies in the Japanese Language, PhD thesis, MIT, Cambridge, Massachusetts.
- Kuroda S Y (1973). Where Epistemology, Style and Grammar Meet: A Case Study from Japanese, *in* S Anderson & P Kiparsky, eds, *A Festschrift for Morris Halle*, New York, Holt, Rinehart and Winston, pp. 377–391.

- Kuroda S Y (1988). Whether We Agree or Not: A Comparative Syntax of English and Japanese, in W J Poser., ed., *Papers from the Second International Workshop on Japanese Syntax*, CSLI.
- Langacker R (1969). On (p)ronominalization and (t)he (c)hain of (c)ommand, in D Reibel & S Schane, eds, *Modern studies in English*, Englewood Cliffs, N. J.: Prentice-Hall, pp. 160–186.
- Li C N & Thompson S A (1976). Subject and Topic: A New Typology of Languages, in C N Li, ed., *Subject and Topic*, New York: Academic Press.
- Mahajan A K (1989), On the A/A-bar Distinction: Scrambling and Weak Crossover in Hindi. Ms., MIT.
- Maling J (1984). Non-Clause-Bounded Reflexive in Modern Icelandic, *Linguistics and Philosophy* 7, 211–241.
- Manning C D (1994), Ergativity: Argument Structure and Grammatical Relations, PhD thesis, Stanford University.
- Manning C D (1996), Argument Structure as a locus of Binding Theory. Ms. University of Sidney.
- Manning C D & Sag I A (1998a). Argument Structure, Valence, and Binding, *Nordic Journal of Linguistics* pp. 107–144.
- Manning C D & Sag I A (1998b). Dissociations Between Argument Structure And Grammatical Relations, in G Webelhuth, J P Koenig & A Kathol, eds, *Lexical And Constructional Aspects of Linguistic Explanation*, CSLI Publications.
- Manning C D & Sag I A (1999). Dissociations between argument structure and grammatical relations, in G Webelhuth, J P Koenig & A Kathol, eds, *Lexical and constructional aspects of linguistic explanation*, Stanford: CSLI Publications.
- Manning C D, Sag I A & Iida M (1999). The Lexical Integrity of Japanese Causatives, in R Levine & G Green, eds, *Studies in Contemporary Phrase Structure Grammar*, Cambridge University Press, Cambridge and New York.
- \*<http://www.sultry.arts.su.edu.au/cmanning/papers/ManningSagIida.ps>
- Marantz A (1984). *On the Nature of Grammatical Relations*, Cambridge, Mass.: MIT Press. Linguistic Inquiry Monograph 10.

- McCawley N (1976). Reflexivization: (a) (t)ransformation (a)pproach, in M Shibatani, ed., *Syntax and Semantics 5: Japanese Generative Grammar*, Academic Press: New York, pp. 51–116.
- McConnell-Ginet S (1982). Adverbs and Logical Form: A Linguistically Realistic Theory, *Language* **58**, 144–184.
- Momoi K (1985). Semantic Roles, Variation, and the Japanese Reflexive, *University of Chicago Working Papers in Linguistics* **1**, 73–92.
- Moon G S (1994). Topic Constructions in Korean, in Y S Kim, B C Lee, K J Lee, Y K Yang & J Y Yoon, eds, *Explorations in Generative Grammar: A Festschrift for Dong Whee Yang*, Seoul:Hankuk Publishers.
- Pica P (1987). On the Nature of the Reflexivization Cycle, in J McDonough & B Plunkett, eds, *Proceedings of NELS 17*, pp. 483–99.
- Pickering M & Berry G (1991). Sentence Processing without Empty Category, *Language and Cognitive Processes* **6**, 229–259.
- Pollard C & Sag I (1987). *Information-Based Syntax and Semantics.*, Vol. 1 of *CSLI Lecture Notes Series*, Stanford:CSLI.
- Pollard C & Sag I (1994). *Head-Driven Phrase Structure Grammar*, The University of Chicago and CSLI Publications.
- Przepiórkowski A (1997). Adjuncts as Complements: Evidence from Case Assignment, in *Proceedings of the 4th International Conference on HPSG*.
- Pustejovsky (1988), Event Semantic Structure. Ms. Brandeis University.
- Rappaport G C (1986). On Anaphor Binding in Russian, *Natural Language and Linguistic Theory* **4**, 97–120.
- Reinders-Machowska E (1991). Binding in Polish, in J Koster & E Reuland, eds, *Long-distance anaphora*, Cambridge: Cambridge University Press, pp. 137–150.
- Reinhart T (1976), The Syntactic Domain of Anaphora, PhD thesis, MIT.
- Reinhart T & Reuland E (1993). Reflexivity, *Linguistic Inquiry* **24**(4), 657–720.
- Ross J R (1967). *Constraints Variables in Syntax*, Cambridge, Massachusetts: MIT Press.



- Sag I A (1997). English Relative Clause Constructions, *Journal of Linguistics* **33**(2), 431–484.
- Sag I A & Fodor J D (1994). Extraction without trace. ms.
- Sag I A & Miller P H (To appear). French clitic movement without clitics or movement, *Natural Language and Linguistics Theory* .
- Saito M (1985). Some Asymmetries in Japanese and Their theoretical Implications, PhD thesis, Massachusetts Institute of Technology.
- Saito M (1992). Long Distance Scrambling in Japanese, *Journal of East Asian Linguistics* **1**, 69–118.
- Sells P (1987). Aspects of Logophoricity, *Linguistic Inquiry* **18**(3), 445–81.
- Sohn H M (1994). *Korean: Descriptive Grammar*, London:Routledge.
- Thráinsson H (1991). Long-distance reflexive and the typology of NPs, in J Koster & E Reuland, eds, *Long-Distance Anaphora*, CUP: Cambridge, pp. 49–76.
- Toman F (1991). Anaphors in binary trees: an analysis of Czech reflexives, in J Koster & E Reuland, eds, *Long-distance anaphora*, Cambridge: Cambridge University Press, pp. 151–170.
- Tsoulas G (1999). On Long-Distance Scrambling in Korean, in I W Lee & S Kuno, eds, *Harvard Studies in Korean Linguistics VIII*, Department of Linguistics, Harvard University, Hanshin Publishing Company, pp. 217–302.
- Ueda M (1984). On Japanese Reflexive *zibun*. Ms. University of Massachusetts at Anherst.
- Vallduvì E (1992). *The Informational Component*, New York: Garland.
- van Noord G & Bouma G (1994). Adjuncts and the Processing of Lexical Rules, in *Proceedings of the 15th International Conference on Computational Linguistics (COLING)*, Kyoto, pp. 250–256.
- Wechsler S (1995). *The Semantic basis of argument structure*, Stanford:CSLI.
- Wilkins W (1988). Thematic Structure and Reflexivization, in W Wilkins, ed., *Syntax and Semantics 5: Thematic Relations*, Academic Press: San Diego, CA, pp. 191–213.
- Williams E (1980). Predication, *Linguistic Inquiry* **11**, 203–238.



- Williams E (1981). Argument Structure and Morphology, *Linguistic Review* **1**, 81–114.
- Xue P, Pollard C & Sag I (1994), A new perspective on Chinese *ziji*. from HPSG electronic papers archive, Ohio State University.
- Yang D W (1983), Extended Binding Theory of Anaphora. Ms. Harvard University and Seoul National University.
- Yang D W (1989). An Anaphor Movement, in *Proceeding of NELS*, Vol. 19, pp. 435–452.
- Yoon J H S (1987). Some queries concerning the syntax of multiple subject constructions in Korean, in S K et al., ed., *Harvard Studies in Korean Linguistics II*, Seoul:Hanshin Publishing Complanry.
- Yoon J & Yoon J M (1990). Chain Condition, Ambiguity of Government and Derivational Grammar, in *NELS21*.
- Zribi-Hertz A (1989). Anaphor Binding and Narrative Point of View: English Reflexive Pronouns in Sentence and Discourse, *Language* **65**, 695–727.