INFORMATION TO USERS

This was produced from a copy of a document sent to us for microfilming. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help you understand markings or notations which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure you of complete continuity.

2. When an image on the film is obliterated with a round black mark it is an indication that the film inspector noticed either blurred copy because of movement during exposure, or duplicate copy. Unless we meant to delete copyrighted materials that should not have been filmed, you will find a good image of the page in the adjacent frame. If copyrighted materials were deleted you will find a target note listing the pages in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed the photographer has followed a definite method in "sectioning" the material. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For any illustrations that cannot be reproduced satisfactorily by xerography, photographic prints can be purchased at additional cost and tipped into your xerographic copy. Requests can be made to our Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases we have filmed the best available copy.
STUDIES IN KOREAN SYNTAX:
ELLIPSIS, TOPIC AND RELATIVE CONSTRUCTIONS

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

DOCTOR OF PHILOSOPHY
IN LINGUISTICS
AUGUST 1981

By
Sung-Yun Bak

Dissertation Committee

Roderick A. Jacobs, Chairman
Ho-Min Sohn
Laurence C. Thompson
Patricia A. Lee
Richard Day
Peter Lee
To the memory of my Father
ACKNOWLEDGEMENTS

I would like to thank each member of my thesis committee for his guidance and encouragement throughout. Special thanks go to Professor Roderick A. Jacobs. As program advisor, Professor Jacobs has given me valuable suggestions and insightful comments concerning the contents and form of the thesis, and warm encouragement whenever I needed it. I am grateful for his good humor and patience.

I cannot thank anyone else more than Professor Ho-Min Sohn. He has helped me both academically and non-academically in every way he could. He has been generously unsparing of his valuable time and spent many hours with me. Without his insightful criticisms, suggestions, and the discussions I had with him, this thesis would have been much poorer than it is. Much of my earlier confusion and many of my mistakes have been clarified and corrected through the discussions I had with him. Of course, all the errors and defects still in the thesis are my own. He also kindly gave me an opportunity to present my thesis to his course (EAIA 451, The Structure of Korean) when the final draft was completed. I would like to take this opportunity to thank the participants of the class for their constructive criticisms and suggestions: Professors Dong-Jae Lee and Sung-Ock Won, Mr. Won-Shik Nam, Miss Young-Ae Hong, and Ms. Pong-Soon Shim.

I am deeply indebted to Professors Patricia Lee, Laurence Thompson, Jeanette Gundel, Richard Day and Peter Lee
for their comments on the contents of the thesis, as well as support and understanding.

My heartfelt thanks go to Professor Dong-Jae Lee for the long discussions I had with him, for moral support, and particularly for 'sashimi'. He has kindly offered me a variety of substantial help on so many occasions. Without his help (both academic and non-academic), my life in Hawaii would have been much more difficult than it was. For this, I will remain thankful to him.

My deep gratitude goes to my teachers in Korea: Professors Sang-Buom Cheun, Maeng-Sung Lee, Yong Moon, Juck-Ryoon Hwang, and Wang-Rok Chang, who first introduced me to linguistics and have encouraged me to continue my studies. Professor Chin-W. Kim of the University of Illinois also deserves deep appreciation for encouragement.

Among my friends in Hawaii, I would like to thank Sue and Orlando Archibeque (now in Denver, Colorado) for their unforgettable friendship and encouragement. Sue and Orlando have made my life here much more enjoyable than it would have otherwise been. I also thank Joel Bradshaw for comments on the thesis, careful editing, and warm jokes.

The Culture Learning Institute of the East-West Center has provided me with a four-year scholarship, which covered a large portion of my expenses while I was studying in Hawaii. I would like to thank Mr. Larry Smith and Ms. Lyn Anzai for their support. I also would like to thank the Linguistic Society of America for the grant at the University of
Illinois at Urbana-Champaign in 1978.

Finally, I thank my mother and my wife, Jin-Young, without whose patience and assistance this thesis would not have been possible.

Sung-Yun Bak
Honolulu, Hawaii
August, 1981
This thesis deals with three aspects of Korean syntax: discourse ellipsis, topic and relative constructions. The general approach adopted here is not a "formal-sentential" one, where sentences are studied as the maximal unit of linguistic analysis in isolation from the context, but is rather a "functional and discourse-oriented" one, in which discourse as well as situational contexts, language use, and other pragmatic considerations are taken into account in linguistic analysis.

Chapters 1 and 2 concern discourse ellipsis. Chapter 1 reviews two earlier approaches to discourse ellipsis: the "syntactic deletion" approach and the "direct interpretation" approach of Shopen (1972a). It is argued that neither of these analyses provides a plausible account of ellipsis phenomena.

Chapter 2 proposes a new analysis. It presents a theoretical model that can account for the syntax and semantics of incomplete sentences involving discourse ellipses. This model is composed of two mechanisms: the "generation process" that identifies zero anaphora and reconstructs the syntactic structure underlying an incomplete sentence; and the "interpretation process" that determines the value of the anaphora. It is argued that interpretations cannot be done in purely linguistic terms and that various non-syntactic factors such as inference and pragmatic knowledge should be
considered. Coordinate ellipsis and Gapping are discussed as well.

Chapter 3 is about the notion of "topic." It proposes that topic is a gradient notion along the dimension of "aboutness." After arguing that topic, as a semantico-grammatical category in Korean, is a constituent of underlying structure, we discuss various semantic, syntactic, and discourse factors involved in topic choice. The form-function relationship of the particles ka and nun are also discussed.

Under the assumption that a relative head is the topic of the underlying embedded sentence, Chapter 4 examines various types of relative clauses in terms of topic-comment articulation. A linguistic explanation based on topic-comment structure is proposed to account for the incomprehensibility of self-embedding. A hypothesis that relative constructions are a syntactically frozen version of text is presented. Finally, it is argued that particle deletion in topics and relative constructions should be explained by the "recoverability" condition.
TABLE OF CONTENTS

ACKNOWLEDGEMENTS ........................................... iii
ABSTRACT ....................................................... vi
LIST OF ABBREVIATIONS ........................................ x

CHAPTER I. DISCOURSE ELLIPSIS: INTRODUCTION AND REVIEW

1. Introduction ............................................... 1
2. Classification and Data .................................... 11
   2.1 Classification ........................................ 11
   2.2 Data .................................................. 17
3. Review of Earlier Works ................................... 25
   3.1 Syntactic Deletion Approach ......................... 25
   3.2 Interpretive Analysis: Shopen (1972a) ................ 53
Notes to Chapter I ........................................... 61

CHAPTER II. DISCOURSE ELLIPSIS: A PROPOSED ANALYSIS

1. Introduction ............................................... 65
2. The Generation Process ................................... 77
   2.1 Nominal Ellipsis ...................................... 78
   2.2 Verbal Ellipsis ....................................... 86
3. The Interpretation Process ................................ 93
   3.1 Universe-of-Discourse: A Unified Concept of Context .... 93
   3.2 Data-based Interpretation Rules ...................... 102
      3.2.1 Nominal Ellipsis ................................. 102
      3.2.2 Verbal Ellipsis .................................. 126
   3.3 Inference-based Interpretation Rules .................. 128
      3.3.1 Definitization ................................... 130
      3.3.2 Conjunction ..................................... 136
      3.3.3 Association ...................................... 138
      3.3.4 Abstraction ..................................... 142
4. Ellipsis in Coordinate Structure ......................... 163
Notes to Chapter II ........................................... 191

CHAPTER III. TOPIC

1. Introduction ............................................... 200
CHAPTER IV. RELATIVE CONSTRUCTIONS

1. Introduction ........................................... 341
2. The Thematic Structures of Relative Clauses .......... 348
3. Particle Deletion in Topics and Relative Clauses ........ 387
Notes to Chapter IV ...................................... 400

BIBLIOGRAPHY .............................................. 404
**LIST OF ABBREVIATIONS**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agt</td>
<td>Agent</td>
</tr>
<tr>
<td>Com</td>
<td>Comitative</td>
</tr>
<tr>
<td>Decl</td>
<td>Declarative Ending</td>
</tr>
<tr>
<td>DSC</td>
<td>Double Subject Construction</td>
</tr>
<tr>
<td>Gl</td>
<td>Goal</td>
</tr>
<tr>
<td>Ins</td>
<td>Instrument</td>
</tr>
<tr>
<td>Loc</td>
<td>Locative</td>
</tr>
<tr>
<td>#</td>
<td>The marker of a Discourse-initial Sentence.</td>
</tr>
<tr>
<td>OM</td>
<td>Object Marker</td>
</tr>
<tr>
<td>Obj</td>
<td>Object</td>
</tr>
<tr>
<td>Pat</td>
<td>Patient</td>
</tr>
<tr>
<td>PP</td>
<td>Prepositional/Postpositional Phrase</td>
</tr>
<tr>
<td>PS</td>
<td>Phrase Structure</td>
</tr>
<tr>
<td>Quot</td>
<td>Quotative Marker</td>
</tr>
<tr>
<td>Ques</td>
<td>Question Marker</td>
</tr>
<tr>
<td>SM</td>
<td>Subject Marker</td>
</tr>
<tr>
<td>Src</td>
<td>Source</td>
</tr>
<tr>
<td>Tm</td>
<td>Time</td>
</tr>
<tr>
<td>TOP</td>
<td>Topic Marker</td>
</tr>
<tr>
<td>$\emptyset$</td>
<td>A Zero Pro-form</td>
</tr>
</tbody>
</table>
CHAPTER I

DISCOURSE ELLIPSIS: INTRODUCTION AND REVIEW

1. Introduction

1.1 Whatever their theoretical orientation, linguists generally assume that linguistics seeks to account for the correspondence between linguistic form and meaning. If this assumption is correct, the following text from a drama is of considerable interest.

(1) e ike khelkhelhakwun. 1
   oh well thirsty
   'Oh, Ø thirsty.'

(2) yengkaminim, cohun yangcwu-lul han pyeng
    old man good foreign liquor a bottle
    kaciko-issmunteypsyo.
    have
    'Sir, Ø have a bottle of good foreign liquor.'

(3) cincca-i-pcyo.
    real be
    'Ø a real good one.'

(4) ttak han can-ssik hapsita.
    just one glass each have
    'Let's have just one glass for each of us.'

Sentences (1-4) attract our attention because, although deficient in syntactic form, they do succeed in conveying their intended meanings. For example, every sentence lacks a grammatical subject and (4) omits a grammatical object. But these omissions cause no problem. We easily understand the meanings intended by the speaker. We might say that (1-4) are grammatically deficient but semantically
well-formed. How is this possible? How can a sentence deficient in form function just as well as a full sentence? What kinds of factors are involved in the way we understand language that permit (1-4) to be understood as conveying complete propositions? This is the question we are going to investigate in the following two chapters.

Sentences (1-4) will be called "incomplete" or "elliptical" sentences. Postponing a more exact definition of the incomplete sentence until 1.3 of Chapter 2, we will here define it roughly as a sentence in which some understood semantic material in its propositional make-up is left unsaid in its surface form. The process of leaving some understood features unsaid will be called "ellipsis." For example, the predicate 'be-real' in (3) requires a patient argument in its underlying proposition, but its grammatical reflex, the subject, is missing in the surface form of (3). Thus, (3) is an incomplete sentence. We will call the kind of elliptical process involved in (3) "discourse ellipsis" because the elliptical patient finds its referent across sentence boundaries. The elliptical argument in such cases will be called a "zero (discourse) anaphor."

Incomplete sentences and discourse ellipsis are discussed in Chapters 1 and 2. Chapter 1, which is an introduction and a review of previous works, consists of four sections. Section 1 discusses previous attitudes toward incomplete sentences and discourse ellipsis, and tries to motivate and justify the investigation of this subject in
linguistics. Section 2 classifies incomplete sentences and presents the data to be analyzed. Section 3 reviews the syntactic deletion analyses that have generally been accepted as a theory of incomplete sentences. We will show that these analyses are not only theoretically undesirable but also untenable on empirical grounds. Section 4 examines an alternative interpretive analysis proposed by Shopen (1972a). In Chapter 2 we will propose a new analysis.

1.2 Incomplete sentences and discourse ellipsis have until recently been largely neglected in generative grammar. It is even doubtful whether incomplete sentences have ever been treated as an integral part of linguistic theory. They are usually regarded as stylistic phenomena or as sloppy and lazy expressions not worthy of serious linguistic investigation. This general lack of concern about incomplete sentences reflects the underlying assumptions of generative grammar.

From the earliest stages of its development, generative grammar proclaimed that linguistic theory is primarily concerned with an ideal speaker-listener and his pure linguistic competence, which is

unaffected by such grammatically irrelevant conditions as memory limitations, distractions, shifts of attention and interest, and errors (random or characteristic) in applying his knowledge of the language in actual performance. (Chomsky 1965:3)

Though Chomsky did not claim that any utterance of less than a complete sentence is an error of some sort, this has been
the usual conclusion drawn from his remarks, and incomplete sentences have been dismissed as random performance.

As Lyons (1977:586) correctly points out, the "idealization process" by which Chomsky intended to eliminate performance variables from grammar includes at least three kinds. One is "regularization," which eliminates variables attributable to malfunctioning of physiological or psychological mechanisms involved in speech production such as slips of tongue, coughs, hesitations, and so on. Another is "standardization," which ignores such socio-culturally determined variables as accent, dialectal differences, and other speech variations related to sex and occupation. The third is "decontextualization," which eliminates all context-dependent features of a sentence. The latter include both linguistic discourse and the non-linguistic situation in which a sentence is uttered.

One basic property of incomplete sentences and discourse ellipsis is that they are all context-dependent phenomena, whether that context is discoursal or situational. Without appropriate context, most incomplete sentences cannot accomplish their communicative functions. Thus, it is a natural consequence of the idealization process of generative grammar that incomplete sentences and discourse anaphora have been neglected on a principled basis.

According to the conception underlying generative grammar, the essence of language--linguistic competence--is basically a mathematical and logical system, and this system
should be amenable to formal treatment. A grammar characterizing linguistic competence must consist of a set of formal rules which can generate all and only the grammatical sentences of a language. The generative system identified as the syntactic component of the language is an autonomous formal structure operating independently of semantics and language use. Thus, Chomsky (1972:198) says:

Grammars contain a substructure of perfectly formal rules operating on phrase markers in narrowly circumscribed ways...these rules [are] independent of meaning or sound in their function.

Therefore, the focus of linguistic research in generative grammar has been the search for the formal structure of syntax, which is believed to consist of a set of abstract symbols (S, NP, VP, Comp, etc.) and symbol-manipulating rules such as formation (or PS) rules and transformations. The former determines how symbols are put together to make a sentence, while the latter determines the relationships between sentences.

This assumption in generative grammar has resulted in a static and "non-communicative" view (Matthews 1979:88) of language. It considers language an abstract formal system which has little to do with communication. Thus, Chomsky again tells us:

a generative grammar is not a model for a speaker or a hearer...When we say that a sentence has a certain derivation with respect to a particular generative grammar, we say nothing about how the speaker or hearer might proceed in some practical or efficient way to construct such a derivation. (Chomsky 1965:9)

Given this kind of view of the goal of linguistics,
incomplete sentences would seem totally negligible, because they contribute almost nothing to clarifying the structure of the formal system and very little to revealing formal operations.

The third reason may be that discourse ellipses are not very amenable to formal treatment. The ellipsis rule is not sensitive to constituency and applies across sentence boundaries. Unlike structurally determined deletion rules, such as Equi-NP Deletion, Passive by-agent Deletion, and Imperative Subject Deletion, the ellipsis rule must be totally unconstrained if formulated within the rule format of generative grammar. The nature of pronominalization rules has been subject to controversy because they defy any easy formalization. Ellipsis presents even more difficulty in formalization. 6

1.3 Should incomplete sentences and discourse ellipses really be regarded as marginal phenomena pertaining only to linguistic performance? We believe not. Rather, we argue that they are legitimate objects of linguistic research and should be included as an integral part in the grammars of individual languages. In other words, the use of incomplete sentences and discourse ellipses is a part of our linguistic competence. Below we will try to justify this view.

First, incomplete sentences are subject to many of the same kinds of analysis as complete sentences. Chomsky argued that the grammar of a language must recognize the
grammaticality or ungrammaticality of any string in the language (Chomsky 1957:137). Consider the following incomplete sentences:

(5) eti-ey ka-si-pnikka? where to go HON 'Where are you going?'

(6)a *ey hakkyo ka-pnita. to school go
b *hakkyo-ey ka-si-pnita. school to go HON
c hakkyo-ey ka-pnita. school to go 'Going to school.'

As (6a-c) show, incomplete sentences are judged grammatical or ungrammatical just like full sentences. The same is observed in (8a-d) and (8e), where the former sentences are grammatical but the latter is not. We must ask why (8e) is ungrammatical (see Chap 2: 1.3) and its ungrammaticality must be accounted for by the grammar. Therefore, by Chomsky's own criterion, incomplete sentences must be dealt with in grammar.

The grammar must also account for synonymy relations among sentences and ambiguities of a sentence (Chomsky 1957). Consider the following sentences:

(7) Chelswu-nun Yenghi-lul chacawassta (kuriko),
    TOP OM visited and
    'Chelswu visited Yenghi (and),'

(8)a Chelswu-nun Yenghi-eykey semmwul-ul cwuessta.
    to present OM gave
    'Chelswu gave Yenghi a present.'
Sentences (8a-d) are all synonymous. Thus, just like full sentences, incomplete sentences again show synonymy relations among themselves. Since synonymy relations must be accounted for by the grammar—i.e. a part of linguistic competence—incomplete sentences like (8a-d) must be accounted for by the grammar.

In addition, just like full sentences, incomplete sentences can be ambiguous. They are also infinite in number. All these facts indicate that the speaker's ability to produce and interpret incomplete sentences is as much a matter of linguistic competence as his ability of dealing with full sentences. Therefore, incomplete sentences must be included in grammar even by Chomsky's own definition.

Second, the competence-performance distinction may be justifiable for "regularized" or "standardized" variables, because they are either physiological or sociological in nature and bear little significance for grammar. However, the relationship between context and grammar is very significant and cannot be relegated to the status of random performance. That relationship is often predictable and shows interesting
regularities, as we will see in Chapter 2, Section 3. Furthermore, there are a number of linguistic phenomena that cannot be explained adequately without recourse to context: pronominalization (Sanders 1970; Bolinger 1976), the definite article, word order (all functional sentence perspective theories and Contreras 1976), sentence stress (Gundel 1974; Schmerling 1975), sentence connectives (R. Lakoff 1971), and so on. Thus, it seems clear that at least certain kinds of context should be included in the description of grammar (Sanders 1970; Chang 1972). If so, incomplete sentences and discourse ellipses must be, too, because they are the most typical examples of the interaction between the formal grammar and the discoursal as well as semantico-pragmatic aspects of the context. Chapter 2, Section 3 will substantiate this argument.

Third, much effort has been expended in digging out the formal properties of language within the bounds of generativists' non-communicative hypothesis of language. This hypothesis has been quite successful in certain respects, but has been less so in others--e.g., it has done very little to show how language is actually used in communication. This is a "deplorable state of affairs" (Kuno 1979) if we take a communicative view of language (cf. Dik 1978; Kuno 1979; Matthews 1979), in which language is viewed principally as a means of communication governed primarily by the communicative functions it serves and only secondarily by its formal properties.
Thus, what now seems required, and what we believe is as important as an account of the formal properties of language, is a theory of language use which accounts, in a theoretically plausible manner, for the ways in which the formal and semantico-pragmatic knowledge of the speaker interact in actual communicative situations. If language is viewed from a communicative perspective, incomplete sentences and discourse ellipses take on a new significance. They are no longer marginal linguistic phenomena but crucial examples which show, more clearly than any other syntactic forms, the interaction between the formal aspects and the semantico-pragmatic aspects of grammar in actual communication. Further, we will try to show that an explanatory model based on language use provides a better description of incomplete sentences than does a formal theory (see Section 3 and Chapter 2).

Finally, study of incomplete sentences is particularly important to Korean, because such phenomena constitute a basic feature of the Korean language. In a study of what he called "rapport deletion" in English--identical to our "exophoric ellipsis" (see 2.1) --Thrasher (1974:7-10) proposed three constraints on ellipsis in English. First, ellipsis is confined to initial position; second, it occurs only in informal conversation; and third, the deleted elements are a small set of forms consisting of auxiliary verbs, subjects, articles, possessives, and if. As we will see later, none of these constraints are applicable to Korean, where
ellipsis can apply to any position in a sentence; to any style of speech, including the most formal and planned speech or writing; and to any constituent of the sentence. Kuroda (1965) has argued that ellipsis in Japanese [and Korean] is equivalent to pronominalization in English, but Korean ellipsis actually covers the functional range of both pronominalization and ellipsis in English.

All these facts clearly show that the functional load carried by ellipsis in Korean is much higher than in English, and that elliptical expressions are by no means sloppy expressions in Korean, even though they may be so in English. Incomplete sentences are more frequent than full sentences, even in formal writings. Though it is not clear how much weight should be given to this statistical fact, this phenomenon is an important feature of Korean, one which calls for a principled explanation (see Chap 2: 3.2) and must be included in the grammar of Korean. On the basis of these assumptions, we will take it as established that incomplete sentences and discourse ellipsis are legitimate targets of linguistic investigation.

2. Classification and Data

2.1. Classification

Incomplete sentences and discourse ellipsis have received little attention in Korean linguistics. Except for a few sketchy accounts given by Choi (1975:798-300), there has
been no systematic account of incomplete sentences in Korean. Choi classified incomplete sentences into seven types on the basis of several criteria: grammatical relations, persons, parts of speech, and sentence-vs.-discourse. The following is his taxonomy.

Choi's (1975:798-799) taxonomy

(I-III) Subject ellipsis

(I) First person subject ellipsis

(9) ø nayil caney-lul chacakakeyssney.
   tomorrow you OM visit-will
   'ø will visit you tomorrow.' (ø: na 'I')

(II) Second person subject ellipsis

(10) ø anyenghi kasipsiyo.
    in-peace go
    'Goodbye.' (ø: ne 'you')

(III) Third person subject ellipsis

(11) ø kicha-lul cosimhala.
    train OM watch out
    'Watch for the train.' (ø: nwukwutunci yeki-lul cinanun ca 'anyone who passes this place')

(IV) Ellipsis of verb

(12) ppacin tokki calwu ø.
    dislocated axe handle
    'ø a dislocated axe handle.' (ø: kathta 'looks like')

(V) Object ellipsis

(13) ø kwukyenghale wassupnita.
    to-see came
    'I came to see ø.' (ø: yeki 'here')

(VI) Ellipsis of adverbials

(14) enni -ka tolawassupnita.
    sister SM returned
    'My sister returned ø' (ø: cin-ev 'home')

(VII) Ellipsis in discourse
Choi adds two brief explanatory notes for the seven types of ellipses. He argues that incomplete sentences such as those in (9-14) are derived from corresponding full sentences by deleting some elements for "simplicity or emphasis," while, in (15-16), some elements are deleted to "avoid repetitions."

Choi's taxonomy and explanations are unsatisfactory on several counts. The whole purpose of a taxonomy is to select and organize data, usually according to certain shared properties, so that grammatical and semantic generalizations can be formulated and a systematic explanation provided. From this point of view, Choi's taxonomy is undesirable.

First, grammatical relations such as subject (I-III), object (V), and non-terms such as PP's (VI) are used to classify (9-11) and (13-14). However, grammatical relations do
not seem to play a significant role in explaining ellipsis as a general phenomenon. It might be argued that at least the category of the subject is significant because subject ellipsis is much more common than ellipsis of other grammatical relations. This is true cross-linguistically (Thrasher 1974). But subject ellipsis is more common than other ellipses simply because the subject which normally occupies sentence-initial position is higher in topicality than other role relations (Chap 3: 3.2.3.3), not because subject itself is fundamentally different from other grammatical relations.

Nominals refer to entities, and nominal ellipses are appropriate depending upon whether or not the entities to which they refer are identifiable in the context. In this respect, the subject nominal is not different from nominals having other grammatical relations. Consider

(17) na-nun Yenghi-lul coahanta. waynyahamyen, I TOP OM like because 'I like Yenghi, because,'

(18)a ø yeyppuki ttaymwun-ita. pretty because is 'ø is pretty.' (ø: Yenghi-ka 'Yenghi')

b hyengnim-i ø coahaki ttaymwun-ita. brother SM like because is 'My brother likes ø.' (ø: Yenghi-lul 'Yenghi')

c nay-ka chwum-ul ø kathi chwueesski ttaymwun-ita. I SM dance OM together dance because is 'I danced with ø together.' (ø:Yenghi-wa 'with Yenghi')

The grammatical relations which the elliptical nominal Yenghi
has in (18) are different: it is subject in (18a); object in (18b); and comitative, a non-term argument, in (18c). However, these differences in grammatical relations do not affect their ellipses.

We will argue here that a distinction which crucially affects ellipsis is to be made not among the grammatical relations of nominals, but between the two grammatical categories of nominals and verbals. Nominals and verbals exhibit a clear and fundamental difference in the nature of the phenomena they denote. Nominals typically refer to entities like Chelswu, a school, etc., but verbals represent actions or states of affairs such as being beautiful, play, etc. Since there are no concrete entities to refer to in verbals, and since what they represent must be understood from the situation, the latter reference is much more unclear and difficult to interpret than the former. This difference is very significant and certain generalizations can be formulated in terms of this distinction (see Chap 2: 2.2; 3.2.2; 4.2).

Thus, we propose that one axis of classification be "nominal-vs.-verbal." This will group Choi's I-III, V-VI, and some of VII together under nominal ellipsis, and group Choi's IV and some of VII under verbal ellipsis.

Note that subtypes I-III for subject ellipsis are based on the criterion of "person." Yet there is no reason why this criterion should not apply in the case of object (V) and non-term (VI) ellipsis. If Choi's taxonomy is to be consistent, object ellipsis should be sub-classified into three
types: first person object ellipsis, second person object ellipsis, and third person object ellipsis. The same goes for non-term ellipsis. But it is clear that this does not contribute to an adequate explanation of ellipsis.

Type VII is distinguished from the other types depending on whether ellipsis occurs within or across a sentence boundary. Let us call this criterion "sentence-vs.-discourse." However, this criterion is also somewhat misleading. Note that the surrounding sentences in the discourse provide the antecedents for the elliptical elements in (15-16), but this is not the case in (9-14) because they are all simple sentences. Thus, what Choi seems to have in mind is whether an antecedent is available in the text or not. However, whether or not an antecedent is available in the text does not necessarily correspond to the sentence-vs.-discourse distinction. For example, an elliptical element of a simple sentence can find its antecedent within the same sentence; conversely, a discourse consisting of more than one sentence can have an elliptical element without any antecedent within the discourse:

(19)a apeci -ka tola-ko-ku _sessta. kulayse
父亲 SM back come -ing so
'Father was coming home, so, '

b ø ellun ttwie-nakassta.
quickly run out
'Ø ran out quickly.' (Ø: na 'I', Chelswu, John)

In (19b) the zero anaphor does not have its antecedent in the text. We therefore propose that the sentence-vs.
discourse criterion be replaced by another distinction—the availability or unavailability of the antecedent (or referent) either in the discourse context, as in (15-16), or in the situational context, as in (9-14). We will call the former "anaphoric ellipsis" and the latter "exophoric ellipsis," terms borrowed from Halliday and Hasan (1976).

Finally, Choi's claim that incomplete sentences are derived from full sentences by deleting elements for "simplicity or emphasis" (for (9-14)) or for "avoiding repetitions" (for (15-16)) is simplistic, as we will show in Section 3. It also should be mentioned that the kind of ellipsis involved in (15) is somewhat different from the others, and will be discussed under the separate heading of "coordinate ellipsis" in Chapter 2, Section 4.

To summarize, we have proposed two criteria for classifying ellipsis in Korean: nominal vs. verbal, and anaphoric vs. exophoric. They will cross-classify ellipsis as follows:

(20) A proposed taxonomy of Korean ellipsis

<table>
<thead>
<tr>
<th>Category</th>
<th>Nominal</th>
<th>Verbal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exphoric</td>
<td>Exphoric nominal</td>
<td>Exphoric verbal</td>
</tr>
<tr>
<td>Anaphoric</td>
<td>Anaphoric nominal</td>
<td>Anaphoric verbal</td>
</tr>
</tbody>
</table>

2.2 Data

2.2.1 Nominal Ellipsis

A nominal ellipsis is the surface non-realization of a nominal (with an optional postposition). This is a semantic
classification cutting across grammatical relations, and the information to be retrieved is the referential meaning of an entity.

**Exophoric nominal ellipsis**

In exophoric nominal ellipsis, reference is made to an entity existing in the non-linguistic situational context, which may temporarily be considered the immediate perceptual field of the discourse participants. The most typical example is elliptical reference to speaker and addressee:

(21)a  \( \emptyset \) eti kasio?
    where go
    'Where \( \emptyset \) going?' (\( \emptyset \): you)

b  \( \emptyset \) cip -ey kao.
    house to go
    '\( \emptyset \) going home.' (\( \emptyset \): I)

c  \( \emptyset \) encey tasi naosio?
    when again come out
    'When \( \emptyset \) coming out again?' (\( \emptyset \): you)

d  \( \emptyset \) 10 si -ey nao-keyss-o.
    o'clock at come will
    '\( \emptyset \) will come out at ten.' (\( \emptyset \): I)

(22) \( \emptyset_1 \) ka-si-l ttay, \( \emptyset_2 \) \( \emptyset_3 \) pwuluseyyo.
    go HON when call
    '\( \emptyset_1 \) call \( \emptyset_3 \), when \( \emptyset_1 \) go.' (\( \emptyset_1 \): you, \( \emptyset_2 \): you, \( \emptyset_3 \): I)

Due to their obviousness in the physical perceptual field, these two roles in a speech act are always presupposed and can be ellipsed as in (21-22).

Reference can be made to the set of the two speech roles, or reference to a third party is sometimes combined with that of one of the speech roles:
(23) Ø nayil kucang-ey kaca. tomorrow cinema to go
'Let's go to the movies tomorrow.' (Ø: uri 'we')

(24) Ø nayil -pwuthe phaep -ey tulekanta. tomorrow from strike at enter
'Ø will enter the strike tomorrow.' (Ø: uri 'we')

In (24), the set referred to by the zero anaphor includes the
speaker and a third party—a group of laborers—whose iden-
tity is known to the discourse participants. Note that the
third party is not present in the immediate perceptual field:
this shows that the situational context to which exophoric
reference is made is not necessarily a perceptual field ex-
isting in physical space, but rather an abstract or concep-
tualized situation in which the needed identification can be
made (see Chap 2: 3.1).

Third persons who do not take speech roles can also be
the referents of zero anaphora. They may be physically pre-
sent in the immediate perceptual field,

(25) Ø kwukeyenghale wassupnita. see to came
'I came to see Ø.' (Ø: cip 'house', yeki 'here')

(26) Ø an-ulo tulyeta-nohala. inside bring put
'Bring Ø inside.' (Ø: chayksang 'desk', hwapwun
'pot')

or, the third party can be an entity in the abstract or con-
ceptualized context, just as in (24). For example, suppose
that John and Mary are waiting for Chelswu, who has not
shown up for more than an hour. John says to Mary:
(27) # han-sikan-inə ə kitalyeto ə o-ci ahnney.
   an hour as-long-as wait but come not
   'Even though we waited for ə an hour, but ə
does not show up.' (ə: Chelswu)

It is obvious both to John and Mary who is in question, and
so we might say that Chelswu is an entity existing in the
conceptual world they share. In fact, a set of entities that
the speaker assumes to be adequately specified by the con-
text can be referred to by means of zero:

(28) ə canti-lul palp-ci masio.
   grass OM tread not
   'Do not step on the grass.' (ə: nwukwutenci 'who-
ever')

(29) ə kicha-lul cosimhala.
   train OM watch
   'Watch out for the train.' (ə: anyone who cross-
es this railway crossing)

Finally, there is what we may call "conventional ellip-
sis," in which the ellipted element can be inferred from the
conventions of language use. Some of the examples which
Choi gives under different types of ellipsis are relevant
here:

(30) enni -ka ə tola-wassupnita.
    sister SM back came
    'My sister returned ə.' (ə: cip-ey 'home')

(31) 12 si -ka toyessuni ə mekko-kapsita.
    o'clock became so eat and go
    'Since it is already noon, let's go after eating
    ə.' (ə: pap 'rice')

(32) ə twutulila, kulemyen ə yellil-kes-ita.
    knock then open will
    'Knock on ə. Then, ə will open.' (ə: mwun 'door')

Anaphoric nominal ellipsis
In anaphoric nominal ellipsis, reference is made to an entity which can be identified in the structure of the text. The direction of reference can be either forward or backward, but forward (or cataphoric) reference is quite rare. The situations in which anaphoric nominal ellipsis takes place are of three kinds. The most common is found in the answers of question-answer pairs:

(33)a Chelswu-nun com ettehsupnikka? TOP well how Ques 'How is Chelswu?'

b Ø kwaynchansupnita. fine 'Ø is fine.' (Ø: Chelswu)

(34)a encey hankwuk-ey kasipnikka? when Korea to go Ques 'When do you go to Korea?'

b nayil Ø kapnita. tomorrow go 'I will go Ø tomorrow.' (Ø: hankwuk-ey 'to Korea')

Anaphoric nominal ellipsis is also found in questions about the statements which appear in earlier text. These are often echo-questions used to elicit confirmation or show surprise:

(35)a ce-nun maykcwu-lul cohahapnita. I TOP beer OM like 'I like beer.'

b cengmal Ø cohahaseyyo? really like 'Really like Ø.' (Ø: maykcwu 'beer')

(36)a Chelswu-ka nayil Pusan-ey kapnita. SW tomorrow to go 'Chelswu is going to Pusan tomorrow.'
(36)b nwu-ka ꞌ ꞌ ka-yo?
who SM go
'Who goes ꞌ ꞌ ?' (ꞌ ꞌ : Pusan-ev 'to Pusan')

The third situation is when the speaker adds one or more sentences to a statement already made earlier in the conversation by any of the interlocutors:

(37)a Chelswu-nun yenghwa-lul cohahapnita.
TOP movies OM like
'Chelswu likes movies.'

b ꞌ ꞌ yenkuk-to cohahaciyo.
drama also like
'ꞌ ꞌ likes dramas, too.' (ꞌ ꞌ : Chelswu)

When two or more sentences are added, a so-called "topic chain" (Tsac 1977) is produced:

(38)a nailon-un cilkipnita.
nylon TOP tough
'Nylon is tough.'

b ꞌ ꞌ sayk -to kop -ko
color also beautiful and
'The color is pretty, and' (ꞌ ꞌ : nylon)

c ꞌ ꞌ kwukyeci-ci-to anhsupnita.
wrinkled also not
'ꞌ ꞌ is not wrinkled, either.' (ꞌ ꞌ : nylon)

d tto ꞌ ꞌ swipkey ppalacipnita.
also easily washable
'ꞌ ꞌ is also washed very easily.' (ꞌ ꞌ : nylon)

2.2.2 Verbal Ellipsis

Verbal ellipsis refers to cases in which the verbal predicate of a sentence is not present on the surface.

Exophoric verbal ellipsis

Exophoric verbal ellipsis is usually found in
illocutionarily marked sentences such as orders, promises, or proposals:

(39) ca ili -lo ø.
    well this-way to
    'ø this way, please.' (ø: osio 'come', sesio 'stand')

(40) ani ettehkey yeki-l ø?
    well how here om
    'How come you are here?' (ø: wassci 'came')

(41) (A messenger comes to inform a mother that her son has been killed in a battle.)
    ani wuli ay -ka ø?
    my dear our son sm
    'My dear, is my son ø?' (ø: cwukesseyo 'killed')

(42) ye po. nay cikap ø.
    darling my wallet
    'Darling, ø my wallet?' (ø: cuwusio 'give',
    odiy isso 'where is')

(43) tangchi-anh-un soli ø.
    unreasonable sound
    'ø unreasonable.' (ø: ita 'is', ha-ci mala 'do not say')

Exophoric verbal ellipsis is sometimes observed in idiomatic expressions, such as proverbs:

(44) panul totwuk-i so totwuk ø.
    needle thief sm cow thief
    'He who steals a pin will ø a thief of a cow.'
    (ø: townta 'become')

(45) chen -li kil -to han kelum-pwuthe ø.
    thousand li journey also one step from
    'A journey of one thousand miles ø from one step.'
    (ø: sicak-ita 'begins')

Anaphoric verbal ellipsis

The situations in which anaphoric verbal ellipsis normally occurs are identical to those for anaphoric nominal
ellipsis. It occurs in answers in question-answer pairs,

(46)a emeni -nun eti -ey kasyessni?
  mother TOP where to went
  'Where has Mother gone?'

  b hakkyo-ey ø yo.
  school to
  'ø to school.' (ø: kasseyo 'went')

(47)a nwukwu-ka ilen cis -ul hayssni?
  who SM this thing OM did
  'Who did this kind of thing?'

  b cey-ka ø yo.
  I SM
  'I ø.' (ø: haysseyo 'did')

in questions following up a given statement, either to elic-
it confirmation or to show surprise:

(48)a ce kimchi-lul tamka-wasseyo.
  well kimchi OM made-brought
  'Well, I brought kimchi for you.'

  b kimchi-lul ø yo?
  kimchi OM
  'ø kimchi?' (ø: tamka-wasseyo 'brought')

(49)a ese kapsita.
  quickly go
  'Let's go now.'

  b eti -lo ø yo?
  where to
  'Where ø?' (ø: kayo 'go')

and sometimes in expansions of a previous sentence:

(50)a i os -un maywu ssakwunyo.
  this clothes TOP very cheap
  'These clothes are very cheap.'

  b ce os -to ø yo.
  that clothes also
  'Those, too.' (ø: ssayo 'cheap')
3. Review of Earlier Works

3.1 Syntactic Deletion Approach

3.1.1 Though no explicit rule has been proposed, incomplete sentences have been assumed to be derived from corresponding full sentences by deleting some elements, especially those identical with elements previously mentioned in the discourse (Chomsky 1965: 145, 182; Choi 1975). We will call this kind of analysis the "syntactic deletion analysis."

For illustration, let us consider the following discourse:

(52)a Chelswu-nun eti -ey kassni?
  TOP where to went
  'Where did Chelswu go?'

b hakkyo-ey kassyeo.
school to went
  '(He) went to school.'

In the syntactic deletion analysis: (52b) is generated by going through the following derivations.

Step 1: PS rules of the grammar generate the underlying phrase marker (53):

(53)
Step 2: Lexical insertion rules insert actual lexical items under the pre-terminal nodes:

\[(54)\]

Step 3: A syntactic deletion transformation, which we will call the "Ellipsis rule" deletes Chelswu of (54) under identity with Chelswu of (52a), thus changing (54) to (55):

\[(55)\]

(55) shows the surface structure of (52b).

The ellipsis rule operating in Step 3 of the derivation may be formulated as follows:

\[(56)\]

Ellipsis

\[
\begin{array}{cccc}
X & A_1 & Y & A_2 & Z \\
1 & 2 & 3 & 4 & 5 \\
1 & 2 & 3 & \emptyset & 5 \\
\end{array}
\]

Condition: (a) $2 = 4$

(b) Index 3 can contain sentence boundaries.

The basic assumption of the syntactic deletion analysis is that any instance of zero anaphora was originally a fully specified lexical formative in underlying structure, as in Step 2 of the foregoing derivation, and that this formative is later deleted by the Ellipsis rule under identity
with an element in the previous discourse. Note that in this analysis the semantic interpretation of a zero anaphor is a matter of no particular importance because it can be done at Step 2 before the Ellipsis rule applies.

What is more important in the deletion analysis is the Ellipsis rule itself. How can we refine the rule, particularly the identity condition between Index 2 and Index 4, to such an extent that the rule both becomes descriptively adequate, that is, accommodates all of the discourse anaphora observed in the language, and remains compatible with transformational theory in general? The next section will show that the deletion analysis fails in both respects.

3.1.2 The most serious problem facing the syntactic deletion analysis is exophoric ellipsis, both nominal and verbal. We will show that under the theoretical assumptions of transformational theory, exophoric ellipses cannot be handled adequately by the deletion analysis. The first argument is based on the established assumption of the recoverability condition, and the other two arguments are related to other basic tenets of transformational grammar.

An implication of the general acceptance of the Katz-Postal proposal that transformations must preserve meaning is that the materials affected by deletion rules must be recoverable. This is a logical consequence of the proposal, because if deleted materials are not recoverable after deletion transformations have applied, the meanings before and
after the transformations will be different. This constraint, known as the Recoverability condition, stipulates that deleted items must be recoverable.

A convention has been suggested (Chomsky 1964, 1965; Hankamer 1979) that is supposed to constrain deletion transformations in such a way as to guarantee recoverability. Under this convention, deletions can apply in only two situations: (A) a deletion rule can delete a constituent under identity with some other constituent within the same sentence or in the surrounding text. The deleted element is recoverable because of the presence of the antecedent. Condition (A) is met in such structurally determined deletions as Equi-NP Deletion and in most kinds of anaphoric ellipses. (B) If deletion applies unconditionally, i.e., without a linguistic antecedent, then it must be limited to a specific element in a specified structural position (i.e., in the structural description of the rule). This deleted item is also recoverable because the rule deletes only very specific elements in specified positions. Thus, Chomsky says:

We are proposing the following convention to guarantee recoverability of deletion: a deletion can eliminate only a dummy element or a formative explicitly mentioned in the structural index (for example, you in imperatives) or the designated representative of a category (for example, the wh-question transformations that delete Noun Phrases are in fact limited to indefinite pronouns) or an element that is otherwise represented in the sentence in a fixed position." (Chomsky 1964:68-72; Chomsky 1965: 144-145)

Deletion rules sanctioned by Condition (B) are: Imperative you Deletion; Complementizer that Deletion; Passive by-agent
Deletion; and probably ko-ha Deletion in Korean.

Let us consider the following exophoric ellipses with regard to the above convention.

(57) # Ž kan pam -ey isanghan kkwum-ul kkwuesста.
    last night at strange dream OM dreamt
    'Last night Ž dreamt a strange dream.'
    (Ž: na 'I', Chelswu, Yenghi, John, ku salam
    'that man',)

(58) # na-nun Ž nayngmyen-ipnita. 12
    I TOP cold noodle is
    'As for me, Ž is noodle.'
    (Ž: mekuvenun kes
    'what I am going to eat', mantunun kes 'what
    I am making',)

(59) # ca ili -lo
    well this way to
    'Ž this way, please.'
    (Ž: osio 'come', sesio
    'stand', ancusio 'sit',)

(60) (Two men are discussing plans to kill Yenghi.
    One of them gives a stocking to the other, sug-
    gesting that he hang her with it. The other per-
    son says:)

    # ani ikel -lo Ž ?
    what this with
    'What, Ž with this?'
    (Ž: cwukipnikka 'kill',
    salhayhapnikka 'kill', mok-ul maymnikka 'hang',)

Sentences (57-60) can occur as discourse-initial sentences.
Thus, deletions in (57-60) must meet Condition B, which stip-
ulates that an unconditional deletion must be one of the fol-
lowing four types: a dummy element, a specified element in
the structural index of the rule, a designated representative
of a category, or an element occurring in a fixed position.
However, none of these constraints apply to the deletions in-
volved in (57-60). The deleted elements in (57-60) are nei-
ther dummy elements nor specified elements (e.g., that in
in Complementizer that Deletion). They cannot be a designated representative of a category because they can be almost any noun, verb, PP, and so on. Their positions are not fixed, either, because they may be subjects, objects, datives, etc. Therefore, unless the recoverability condition is given up, the deletions in (57-60) cannot be justified in transformational grammar. The deleted material, whatever it may turn out to be, is not recoverable once eliminated by the Ellipsis rule.

Now, let us concentrate on (57), although our comments will apply to (53-60) as well. Whether or not the deleted element is recoverable, there is another problem in setting up an underlying structure for (57). Note that there are many possible referents of the deleted element in (57), some of which are given in parenthesis. But which one out of these multiple candidates has actually been deleted? We cannot choose any one of them arbitrarily, as Choi (1975:798) did in choosing na 'I', and there is no way to tell which one was there originally. So, we have no option but put them all in the underlying structure.

Since the set of possible referents of the deleted element is infinite, this means that we have to posit an infinite number of underlying structures for a single surface form like (57). For example, (57) can be given the following set of underlying structures:
The list in brackets can be indefinitely long.

Furthermore, each item in the list can in turn be expanded indefinitely by recursion. As an illustration, let us take *ku salam* 'the man' of (61):

(62)a *ku salam*
  b *[[khi-ka khun] ku salam]*
    c *height tall*
      d *'the tall man'*
  -i
  e *[[nay-ka mannan[khi-ka khun]] ku salam]*
    f *I SM met*
      g *'the tall man whom I met'*
  kan pam-ey isanghan kkwum-ul kkwuessupnita.

It is clearly wrong to claim that (57) is infinitely ambiguous, as the underlying structures (61-62) would imply. If we consider the recoverability condition, we have to say that (57) is derived from (61-62) by deleting an infinite number of elements in parenthesis without any hope of their being recoverable. This kind of unconstrained deletion is surely unacceptable for any grammar claiming to be of any theoretical interest (cf. Culicover 1970).

Finally, consider sentences like (58) and (60). Suppose that (58) is uttered in a restaurant when a waitress has come to take an order and that (60) represents a plan to kill *Yenghi*. Although we know the specific contexts in which
these sentences are used, it is difficult to tell exactly what lexical material has actually been deleted in each of these sentences. For example, what has been deleted in (58) might be phrased in at least two ways: makulyenun kes 'what I am going to eat' or owumwunhalyeko hanun kes 'what I am going to order'. The same is true of (60): cwukipnikka or salhayhapnikka 'kill', or epsaypelipnikka 'eliminate'. In either case, the understood semantic material is rather clear, but there is no way to tell the precise syntactic manifestation in which the semantic concept was realized. This is a problem for a purely syntactic rule of deletion and for the deletion analysis in general.

In sum, we see no plausible way to explain either nominal or verbal exophoric ellipsis within the framework of a syntactic deletion analysis. Three specific problems have been discussed in this regard: the problem of insuring recoverability, the potentially infinite number of underlying structures, and the problem of specifying the syntactic shape of the deleted material. The syntactic deletion analysis must be rejected on the basis of exophoric ellipsis.

3.1.3 Exophoric ellipsis is problematic because no linguistic antecedents are available in the surrounding text to function as deletion triggers and to insure recoverability. In anaphoric ellipsis, on the other hand, a linguistic antecedent is usually available in the text. Recoverability is therefore guaranteed and the deletion analysis seems adequate.
But the actual facts are not so simple. There are a number of cases in which a simple syntactic deletion does not work. Before we examine another set of counterexamples, let us review what "identity" means in the deletion analysis.

Several parameters should be considered in defining the identity condition: referential, structural, lexical identity, and identity of logical form (Sag 1977). They are all crucial to the operation of the Ellipsis rule. Some will be discussed in later sections; we concentrate here on lexical identity.

Lexical identity that is required for discourse ellipsis may be defined as a strict word-for-word equivalence between the antecedent and the victim. Consider the following:

(63) a Chelswu₁-ka wuli-lul chacawassta.
    SM us OM visited
    'Chelswu₁ came to visit us.'

   b kulena amwu-to ku sił-epsmun nyesek₁-ul panki-
      but nobody the silly guy OM welcome
      ci-anassta.
      did not
      'But nobody welcomed the silly guy₁.'

Even though the two underlined NP's in (63a-b) are coreferential, the NP in (63b) cannot be deleted, as in (64):

(64) *kulena amwu-to Ø panki-ci anassta.
     but nobody welcome did not
     'But nobody welcomed Ø.'

(64) is not acceptable in its intended sense, because it no longer carries the meaning that Chelswu is a silly guy.
Therefore, Ellipsis rule (56) can apply only when lexical identity obtains between the antecedent and the deleted victim, otherwise elements that add meaning may be lost.

However, the lexical identity defined in terms of a strict verbatim equivalence is too strong. This was first observed by Chomsky (1965: 180), who gives the following examples:

(65) These men are more clever than Mary $\phi$.
($\phi$: *is* clever)

(66) I have a friend $\phi$ from England.
($\phi$: *the* friend is)

(65-66) show that deletion rules like Comparative Deletion and Relativized NP Deletion disregard such differences as that of number or definiteness between the antecedent and the victim. To accommodate facts like these, Chomsky (1965) suggested that identity in deletion rules be understood as "non-distinctness" between the antecedent and the victim in essential semantic features. He says:

The general principle for erasure operation, then, is this: a term X of the proper analysis can be used to erase a term Y of the proper analysis just in case the inherent part of the formative X is not distinct from the inherent part of the formative Y. (p. 182)

This proposal means that all transformationally introduced features such as number and definiteness can be ignored in the syntactic operation of deletion.

Even though lexical identity between the antecedent and the victim is weakened to "non-distinctness", it seems that we have not arrived at a descriptively adequate formulation
of the identity condition which can accommodate all anaphoric ellipses. Consider the following texts:

(67) a Chelswu-nun mayil cipwung-ul kochinun kwun.  
TOP everyday roof OM mend 
"Chelswu mends the roof everyday."

b kulem Ø sip-nyen-cen-ey ciessnuntey tangyenhaci.  
sure ten year ago at built so natural  
"Of course, Ø was built ten years ago and so it is natural." (Ø: Chelswu-ka cipwung-ul kochinun ku cip 'the house whose roof Chelswu is mending')

(68) a ccey Chelswu-ney cip-ey kassessnuntey,  
yesterday of house to went and 
'I went to Chelswu's house yesterday, and'

b Ø pap-ul mek-ko issstela.  
rice OM eating  
'Ø were eating rice.' (Ø: Chelswu-ney sikkwu  
'Chelswu's family members')

(69) a Yenghi-ka cwungmay-lul setallako hanuntey,  
SM match-making arrange asked and 
'Yenghi asked me to arrange a match, and'

b Ø pantusi uysa-iewa hantay.  
by-all-means doctor must-be  
'Ø must be a doctor, she specified.' (Ø: Yenghi ka kvelhon-halyen salam 'the man whom Yenghi is going to marry')

(70) a seymeynt han photay-ey molay twu-thong-ul nethko,  
cement one pack in sand two bucketful put  
'Put two bucketfuls of sand in a pack of cement,'

b Ø₁ mwul-ul pwue, Ø₂ cal sekusio.  
water OM pour well mix  
'then, pour water and mix well.'  
(Ø₁: seymeynt-wa molay 'cement and sand'  
Ø₂: seymeynt-wa molay-wa mwul 'cement, sand and water')

The deletions involved in (67b-70b) are all anaphoric, but no linguistic antecedent which satisfies the lexical identity condition in terms of non-distinctness is found for them in the surrounding text. Deletions like these constitute
serious counterexamples to any syntactic analysis of deletion, because their antecedents are not overtly mentioned in the discourse.

Ellipses of this nature are permitted because of reasoning or inference processes all human use in communicating. They raise the question whether ellipsis can really be explained in purely syntactic terms. Inferential ellipses of this nature are quite common in Korean. We will provide some more examples:

(71)a ecey sen-ul poassney.
yesterday marriage interview had
'Yesterday I had a marriage interview.'

b ø yeyppuki-nun hatekwuman, cipan-i an coha
beautiful but family SM not good
kumantwuessney.
gave up
'Ø was beautiful, but her family was not good,
so, I gave it up.' (Ø: sen-ul pon vece 'the
girl I met at the marriage interview')

(72)a akka kyomwusil-ey kassteni,
while ago teachers' room to went
'I went to the teachers' room a while ago.'

b machim ø hoyuy-lul hasi-ko istekwunyo.
just-in-time meeting have -ing
'Ø were having a meeting just at that time.'
(Ø: sensayngnintul 'teachers')

(73)a Kumseng TV-lul hana sassney.
OM one bought
'I bought a Kumseng TV.'

b ø cal naoten?
well coming
'Is Ø coming clear?' (Ø: hwamyen 'the picture')

(74)a na-nun 'The Grapes of Wrath'-lul ilkessmuntey,
I TOP CM read
'I read 'The Grapes of Wrath'.
The foregoing two arguments concern empirical problems the syntactic deletion analysis is faced with. In this section we argue that the deletion analysis is also undesirable from a theoretical point of view because of the enormous power the grammar must have if the theory is to work. Let us go back to the generative mechanism for deriving incomplete sentences we sketched in 3.1.1. There, we posited three steps: Step 1 for generating the underlying phrase marker; Step 2 for lexical insertion; and Step 3 for deletion. We will examine each of these three steps in the inverse order.

Let us consider sentences (52a-b) for Step 3. In order for Step 3 to work, the ellipsis rule must compare the underlying structure (54) with the antecedent sentence (52a) and delete the NP Chelswu of (54) under coreference and identity with Chelswu of (52a). This means that the transformational rule of Ellipsis must be able to make use of coreference relations between the two NPs. Transformations in their usual sense are formal operations working on category symbols like NP, V, Comp. They are purely syntactic operations, so that they are rarely sensitive to whether two given items refer to the same individual or not. If it is to be sensitive to such matters, the ellipsis rule in question must have exceptional power. But this kind of exceptional power allowed to
Ellipsis rule must be considered _ad hoc_ unless it is extensively justified on independent grounds. Even if it is justified, it will unquestionably increase the power of transformations enormously, and proportionately reduce the explanatory power of the transformational theory.

Much the same arguments were made by Jackendoff against a pronominalization rule in English:

Transformations generally cannot mention coreference relations: there is no rule, for example, that preposes a noun phrase if it is coreferential with some other noun phrase. Thus in stating pronominalization and reflexivization transformations that refer to intended coreference, we are implicitly granting transformations power which they do not in general possess. (Jackendoff 1972: 109)

Determining coreference is basically a semantic process and it is doubtful whether a purely syntactic operation like Ellipsis rule can perform the task of coreference determination.

Note that the ellipsis rule based on coreference relation is possible only if coreference between two NPs is marked in underlying structure. However, Lasnik (1976) pointed out another theoretical problem resulting from this practice. It is well-known that some pronouns must be base-generated (see 3.1.5.2). Consider

(75) # He claims that he is sick.

If (75) is used as a discourse-initial sentence, the pronouns _he_ must be base-generated. Note that the two _he's_ in (75) can be either co-referential or distinct. Thus, (75) must
be assigned two underlying structures (76-77):

(76) \textit{He} claims that \textit{he} is sick.
(77) \textit{He} claims that \textit{he} is sick.

(76-77) show that coreference marking must be done for pronouns in the underlying structure.

Now, let us consider another sentence:

(78) \textit{Schwartz} claims that \textit{he} is sick.

In a transformational analysis of pronominalization, (78) is derived from ((79):

(79) \textit{Schwartz} claims that \textit{Schwartz} is sick.

in which two fully specified NP's are marked for co- or distinct-reference. Then, we have to say that coreference marking is needed in underlying structure not only between two full NPs, as in (79), but also between two pronouns, as in (76-77). If so, as Lasnik points out, there is no reason why coreference marking should not be made between a full NP and a pronoun in underlying structure.

This has serious implications for any transformational analysis of pronominalization. First, the underlying structure of (78) does not have to be (79). Indeed, the structure (78) may be considered the underlying structure because coreference marking can occur between a full NP and a pronoun in such an underlying structure. A transformational analysis is thus unnecessary and superfluous.
Second and more important, a transformational analysis is faced with a contradiction. For example, a passive transformation may apply to (78), itself the underlying structure of the same sentence. The passive transformation changes (78) to (80):

(80) *That he₁ is sick is claimed by Schwartz₁.

The ungrammaticality of the derived structure (80) shows that coreference marking between a full NP and a pronoun cannot be made in underlying structure. That is, coreference marking must be made at a later stage of derivation (cf. Jackendoff 1972; Bresnan 1976:15-17). So, a contradiction arises.

The above discussion shows that marking coreference in underlying structure is problematic. Therefore, the deletion analysis whose ellipsis rule is based on coreference marking in underlying structure must also be problematic.

3.1.4.2 Now, let us move to Step 2, the lexical insertion process. Consider the following text from a drama in which two people are plotting:

(81)a sensayng -to asikeyssciman, ku-pwun ani-myen you too know may that man not if

nwu-ka ku-il-ul  hakeyssupnikka?
who SM the job OM do will
'As you may know, if it is not the man, who will do this kind of job?'

b ø yensey-ka myech-i-sipnikka?
age SM what be HON
'What is the age?'
What is significant here is (81b). In the syntactic deletion analysis (81b) is generated as follows: first, the base rules generate the underlying structure, whose pre-terminal string (when subcategorization rules such as Chomsky's (1965:85) have applied) will look like (82):

\[(82)\]
\[
\begin{array}{c}
NP \\
N_1 \\
\{+N\text{ animate}\} \\
\end{array} \quad \quad \quad \quad \quad \quad \begin{array}{c}
S \\
NP \\
\end{array} \quad \quad \quad \quad \quad \quad \begin{array}{c}
S' \\
NP \\
\end{array} \quad \quad \quad \quad \quad \quad \begin{array}{c}
VP \\
\end{array} \quad \quad \quad \quad \quad \quad \begin{array}{c}
V \\
\end{array} \\
\end{array}
\]

Then, the lexical insertion rules will insert lexical entries into the pre-terminal string. Lexical insertion is a kind of syntactic transformation rule which replaces "a complex symbol of a pre-terminal string" with a lexical item which is "not distinct" in syntactic and semantic features (Chomsky 1965:84-90). So, the lexical insertion rule will examine the lexicon and will allow the lexical entries vensey 'age', mvech 'what', and ita 'be' to be inserted under $N_2$, $N_3$ and $V$, respectively.

But how about $N_1$, which is the topic of (81b)?

In
this case, the lexical insertion rule cannot work in its usual way, because the lexical insertion here is context-sensitive. That is, the lexical insertion rule must examine the context to find out the lexical item to be inserted under $N_I$. The lexical rule finds from (81a) that there are two candidates for the topic NP: $ku\ pwun$ 'that man' and $sensayng$ 'you'. But which one is to be chosen?

Since (81b) seems to be ambiguous between these two readings, as is clear from the ensuing conversational exchanges (81c-d), it might be argued that (81b) is actually ambiguous in deep structure. Either (83) or (84) might be proposed as the deep structures of (81b):

(83) $ku\ pwun\un$ yensey-ka myech-ipnikka?
    'As for that man, what is his age?'
(84) $sensayng\un$ yensey-ka myech-ipnikka?
    'What is your age?'

However, (81b) is ambiguous only in surface structure as a result of ellipsis and this was not so in deep structure. (81d) clearly shows that what the speaker of this sentence intends is the meaning of (83), in which the referent of the zero anaphor is 'that man', not 'you'. The topic $ku\ pwun$ of the deep structure in (83) has presumably been deleted by the Ellipsis rule under identity with $ku\ pwun$ of (81a).

Thus, the lexical insertion rule must choose $ku\ pwun$ from among these two candidates for $N_I$. In order for the rule to do so, it must be able to find out what the speaker of (81b) had in mind when he uttered the sentence. Note
that the lexical insertion rule is a purely syntactic rule (Chomsky 1965:89-90) which has no access to the speaker's intention in discourse contexts. Therefore, if the deletion analysis is made workable, the lexical insertion rule must be given the power to divine the speaker's intention in discourse contexts. This will greatly increase the power of the grammar.

Because of this immense power allowed in the grammar, particularly with regard to the lexical insertion rule, the explanatory power of the syntactic deletion analysis is significantly reduced. This deletion analysis effectively prevents us from investigating many real and crucial questions about the use of incomplete sentences and discourse anaphora. Consider the following texts:

(85)a Chelswu-nun Yenghi-lul kosohayssta.  
TOP OM accused
'Chelswu accused Yenghi.'

b Ø ton -ul hwumchyess-ki ttaymwun-iessta.  
money OM stole because was
'Because Ø stole money.' (Ø: Yenghi, *Chelswu)

(86) kyengchal-un kwuncwung-i deymo-hanun-kes-ul  
police TOP crowd SM demonstrating OM
makassta. waynyahamyen,  
prevented because
'Police prevented the crowd from demonstrating, because,'

(87)a Ø phoktong-ul twulyewehayss-ki ttaymwun-iessta.  
riot OM feared because was
'Because Ø feared a riot.' (Ø: kyengchal 'police', *kwuncwung 'crowd')

b Ø phoktong-ul ilukhi-ki ttaymwun-iessta.  
riot OM rise-in because was
'Because Ø tended to rise in riot.' (Ø: kwuncwung
What is interesting and crucial in the use of incomplete sentences like (85b) and (87a-b) is why only one of the two choices given in parentheses is acceptable for the interpretation of the zero anaphor involved in each. For example, why must the zero pro-form of (87a) be interpreted as coreferential with 'police', whereas that of (87b) is coreferential with 'crowd'.

Such interesting questions cannot arise in the syntactic deletion analysis, because the grammar, with its powerful lexical insertion rule, automatically generates for (87a-b) the following underlying structures,

(88)a Kyengchal-i phoktong-ul twulyewehayss-ki
police SM riot OM feared

Ttaymwun-iessta.
because was
'Because the police were afraid of a riot.'

B Kwuncwung-i phoktong-ul ilukhi-ki ttaymwun-iessta.
crowd SM riot OM rise-in because was
'Because the crowd tended to rise in riot.'

in which the zero anaphors of (87a-b) are assigned kyengchal and kwuncwung, respectively, by some mystical device. Therefore, the real questions involved in the use of incomplete sentences and discourse anaphora are effectively eliminated on a principled basis under the deletion analysis. It seems that the syntactic deletion analysis severely underestimates the role played by context and by semantico-pragmatic knowledge about the world in our understanding of linguistic forms.
Worse yet, the questions posed by the syntactic deletion analysis may be the wrong ones.

3.1.4.3 A similar problem is observed in Step 1 of the afore-mentioned derivation, the stage of generating the underlying phrase marker. The problem arises because incomplete sentences are vague and indeterminate not only semantically, as in (81b), but also syntactically. Consider:

(89) yePO, nay cikap Ø.
    darling my wallet
    'Darling, my wallet.'

In this case, the deleted material can have many interpretations: eti-ev isso 'where is it'; chaca-cwusio 'find it for me'; poasso 'did you see it', and so on. The speaker of (89) must have meant only one out of these multiple interpretations. In order for the grammar to generate the underlying structure of (89), it must decide which interpretation the speaker intended in this particular context. Depending upon which was intended, the node VP will be expanded differently. For example, eti-ev isso should have a VP structure \([PP V]_{VP}\); poasso will be given a structure \([NP V]_{VP}\), etc. Again there is no way for the grammar to determine the correct structure intended in (89), unless it is given the power to find out the intention of the speaker in this context.

Finally, the identity condition, which is crucial to the operation of the Ellipsis rule (56), is very difficult to define (Wasow 1975; Chap 2: 3.3.4). Unless this identity
condition is not satisfactorily defined, the deletion analysis remains problematic.

3.1.5 Before we leave this section, we will review two alternative proposals to the syntactic deletion analysis, which are also framed in the transformational deletion approach.

3.1.5.1 Deletion-via-pronominalization analysis

Sanders and Tai (1969) and D.W.Yang (1975) propose an alternative analysis which might be called the "Deletion-via-pronominalization" (henceforth "Pronominalization") analysis. They propose that all deletions are derived via an intermediate stage of pronominalization, which they call the "analyticity condition." Consider the following sentences.

(90) kasi-l ttay, Ø pwuluseyyo.
    go when call
    'Call Ø when you leave. (Ø: na 'I', kay 'dog', John,)

(91)a Chelswu-nun com ettehsupnikka?
    TOP rather how
    'How is Chelswu?'

    b Ø kwaynchansupnita.
        fine
        'Ø is fine.' (Ø: Chelswu)

Note that (90) and (91b) involve exophoric and anaphoric ellipsis respectively.

In Pronominalization analysis, the zero anaphor with exophoric reference is represented as a base-generated pronoun in underlying structure. For example, (90) is derived from the underlying structure (92) by an optional pronoun
deletion rule.

\[(92) a \] kasi-l ttay (\text{na 'I' (uri 'we')} \{ \text{ku 'he'} \text{ ku veca 'she'} \text{ kukes 'that'} \text{ kukestul 'those things'} \text{ kutul 'they'} \}) -ul pwuluseyyo.

\[ b \] kasi-l ttay $\emptyset$ pwuluseyyo. (by an optional pronoun drop rule)

Note that the pronouns in (92a) are base-generated ones. Zero anaphors with anaphoric reference like (91b), on the other hand, are explained in the same manner as the syntactic deletion analysis proposes, except for an intermediate pronominalization stage. For example, (91b) is derived as follows:

\[(93) a \] Chelswu-nun kwaynchansupnita. (Deep structure)

\[ b \] ku-nun kwaynchansupnita. (Pronominalization)

\[ c \] $\emptyset$ kwaynchansupnita. (Optional pronoun drop)

What is new and advantageous in this analysis is its proposal about the derivation of exophoric ellipsis, such as that in (92). Note that the zero anaphor is represented as a set of pronouns in the underlying structure. This proposal is justifiable in an interesting sense. As we will show later (Chap 2: 3.3.1), the referent of an exophoric anaphor is always definite because it is some object uniquely established for both of the discourse participants. Likewise, deictic pronouns such as 'I', 'he', and 'they' are also definite. Therefore, the use of deictic pronouns for exophoric anaphora
in underlying structure seems to be compatible with the definite nature of the referent, and so is justifiable.

This proposal can also avoid the two problems we mentioned with regard to the syntactic deletion analysis in 3.1.2. One is that unconditional deletions—those without linguistic antecedents—of pronouns are easier to defend than unconditional deletions of full noun phrases. The other problem is that the possible underlying structures for (90) can be reduced to no more than nine because the set of deictic pronouns is not open-ended but closed: ne 'I', uri 'you', ne 'you', nehitul 'you', ku 'he', ku yeca 'she', kukes 'that', kukestul 'those things', and kutul 'they'.

However, this seeming advantage is spurious. Note that when (90) is uttered, the speaker uses a zero anaphor for a definite object. There is no way, however, to tell which one of the nine pronouns given in the presumed underlying structure (92a) is the correct one intended by the speaker. The only difference between the syntactic deletion analysis and the pronominalization analysis is that the latter reduces the number of possible referents, and thereby the number of possible underlying structures, from infinity to nine. But the pronominalization analysis cannot propose any principled way to choose from among these nine candidates. Thus, the basic problem facing the syntactic deletion analysis remains unsolved. Although the pronominalization analysis appears to be an improvement over the syntactic deletion analysis, it is still manifestly inadequate.
There is a still more serious problem with the pronominalization analysis, however, which indicates that it is actually less preferable than the syntactic deletion analysis. It has nothing to say about verbal ellipsis. An analysis which can provide a unified explanation for nominal and verbal ellipses should be more highly evaluated than one which cannot or which can handle only one of them. The syntactic deletion analysis can accommodate both types of ellipsis since it assumes that the elliptical elements are full NP's or VP's in underlying structure. The pronominalization analysis is problematic because no pro-verbal substitution is possible for verbal zero anaphora. For example, kulehata 'do so' is a pro-verb, but no pro-verbal substitution is possible for both exophoric and anaphoric zero anaphora.

(94)a  ca ili -lo ℹ.
please here to
'ℹ this way, please.' (ℹ: osio 'come', sesio 'stand',)

b *ca ili-lo kulehasipsiyo.

(95)  ecey eti -ey kassessni?
yesterday where to went
'Where did you go yesterday?'

(96)a  hakkyo -ey ℹ.
school to
'ℹ to school.' (ℹ: kassessta 'went')

b *hakkyo -ey kulehayssessta.
did-so

(94b) and (96b) are unacceptable because they contain pro-verbal substitutions instead of the zero anaphors in (94a) and (96a), respectively.
Other objections to the syntactic deletion analysis discussed in 3.1.3 and 3.1.4 apply equally to the pronominalization analysis.

3.1.5.2 Mixed Theory

Any transformational deletion analysis is necessarily doomed to failure in exophoric ellipsis, because it is not possible in syntactic terms to identify uniquely the full NP's or VP's the zero anaphors were derived from in underlying structure. Recognition of this impossibility came early in work on pronominal anaphora. It was noted very early that exophoric pronouns cannot be derived from fully specified NP's by the transformational rule of pronominalization. Consider

(97) # I think that she will die.
(98) # That's it.

Like exophoric zero anaphors, the underlined pronouns in (97-98) cannot be transformationally derived from full NP's because we do not know what they were in underlying structure. That is, it is impossible for the grammar to specify all and only the unique referents of these exophoric pronominal anaphors.

Again, just as in anaphoric ellipsis, it may be possible to derive anaphoric pronouns from fully specified NP's which are identical to some linguistic antecedents in the text. Thus, we could say that (99a) came from (99b):
(99)a The boy\(_i\) knew that he\(_i\) would win.
b The boy\(_i\) knew that the\(\text{boy}_i\) would win.

This fact led Chomsky (1965:145) to propose a mixed analysis, in which anaphoric and exophoric pronouns are to be derived from different deep structure sources. According to his suggestion, exophoric pronouns are inserted into the deep structure directly by the lexical insertion rule. Anaphoric pronouns, on the other hand, are inserted into surface structure transformationally by a pronominalization rule.

Hankamer and Sag (1976) and Sag (1977) have applied the same analysis to zero anaphora, and proposed a "Mixed analysis", in which all exophoric ellipses, which they call "pragmatically controlled deep anaphora," are directly generated in underlying structure by PS rules; and all anaphoric ellipses, which they call "linguistically controlled surface anaphora," appear first as fully specified lexical formatives in underlying structure, later to be deleted by an identity deletion rule. The cases of exophoric anaphora are supposed to undergo semantic interpretation rules later (though none are proposed) which assign to the zero anaphor some appropriate referent in the situational context. The Mixed analysis can handle exophoric ellipses adequately because the transformational deletion account has been given up. It can also handle many cases of anaphoric ellipsis as long as a linguistic antecedent is available in the surrounding text. However,
we maintain that the interpretive approach is to be preferred over this partly-interpretive and partly-transformational analysis.

First, it is not clear whether the situational and the discoursal contexts are so distinct as to warrant this kind of a bifurcated analysis. Consider the following texts:

(100) Ø nayil kayaci. tomorrow go-will 'Ø will go tomorrow.' (Ø: na 'I')

(101) Ø nayil kayaci -lako Chelswu-ka malhayssta. tomorrow go-will Quotative SM said 'Chelswu said that Ø will go tomorrow.' (Ø: Chelswu )

The zero anaphor of (100) has exophoric reference to the speaker, but (101) has anaphoric reference to the matrix subject Chelswu. There is no doubt that the zero anaphors of (100-101) can be accounted for by a single generalization insofar as they both refer to the speaker in some speech act situation (see Chap 2: 3.2.1.1). However, the Mixed theory claims that the zero anaphor of (100) is generated as such in underlying structure and is interpreted later by a semantic interpretation rule, whereas that of (101) appears as Chelswu in underlying structure, later to be deleted transformationally under identity with the matrix subject. This is surely a loss of generalization.

In Chapter 2, we will propose a unified concept of contexts based on the notion of "universe-of-discourse." This notion can eliminate the distinction between the situational
and discoursal contexts. The justification for the Mixed analysis is thus significantly weakened and a unitary explanation for both exophoric and anaphoric anaphora (as in (100-101)) is made possible. Other things being equal, a unitary analysis should be more highly evaluated than a bifurcated analysis which depends on two different explanatory mechanisms. Since the deletion approach cannot handle exophoric ellipsis, it cannot offer a unitary explanation. An interpretive approach would seem to offer the best hope of a unitary explanation. The Mixed analysis also cannot escape the criticisms mentioned in 3.1.3 and 3.1.4.

To conclude this section, we have reviewed three analyses within the deletion approach: the syntactic deletion analysis, the pronominalization analysis, and the mixed analysis. We have argued that none of these deletion analyses are adequate either empirically or theoretically.

3.2. Interpretive Analysis: Shopen (1972a)

In interpretive studies of pronominal anaphora (Dougherty 1969; Jackendoff 1972; Lasnik 1976) pronouns are directly generated in deep structure by the PS rules, and then they are matched, through a set of semantic interpretive rules, to other fully specified noun phrases in the text for interpretation. For example,

(102)a na-nun John-ul mannassnentey
I TOP OM met and 'I met John.'
The underlying structures of (102b-103b) contain the same elements as their surface structures, with the following constituent structures:

(104) \[
\begin{align*}
&\quad \text{[\{ku-nun\} NP \{minam-iessta\} VP]} \\
&\quad \text{S}
\end{align*}
\]

(105) \[
\begin{align*}
&\quad \text{[\{na-nun\} \{ku nye-lul\} \{hakkyo-eyse mannassta\} VP]} \\
&\quad \text{S}
\end{align*}
\]

Later, semantic interpretation rules will assign the underlined pronouns to 'John' and 'Yenghi', respectively.

It should be possible to extend the same line of analysis to zero anaphora and incomplete sentences. That is, incomplete sentences are generated from underlying structures containing exactly the same elements that turn up on the surface. Later semantic interpretation rules will assign appropriate readings to unrealized ellipted elements. Such an analysis was proposed by Shopen (1972a), who proposed that incomplete sentences are generated directly by the PS rules of a language. Shopen says:

I propose that the PS rules of the grammar generate incomplete sentences directly, i.e., that there will be rewrite rules of the form \(U \rightarrow X\), where \(U\) stands for utterance and \(X\) is a string of one or more categories. Thus we will have \(U \rightarrow PP PP\) (into the dungeon with
him), U $\rightarrow$ NP (a cup of coffee), U $\rightarrow$ AP (medium rare) etc. (Shopen 1972a:24)

In order to illustrate what Shopen proposed, let us consider the following sentences, which are identical to (102-103) except that zero anaphora are used instead of pronominal anaphora.

(106)a na-nun John-ul mannassnuntey
'I met John.'

b minam -iessta.
handsome was
'(He) was handsome.'

(107)a ne-nun eti-se Yenghi-lul mannassni?
'Where did you meet Yenghi?'

b na-nun hakkyo-eyse mannassta.
'I TOP school at met
'I met (her) at school.'

According to Shopen's proposal, (106b-107b) have the following underlying structures (108-109), respectively:

(108) \[[[minam-iessta]_{VP}]_U

(109) \[[[na-nun]_{NP}[hakkyo-eyse mannassta]}_{VP}]_U

(108-109) will undergo semantic interpretation rules, which will assign readings to them. Below we will argue that this kind of interpretive analysis is not only inadequate but also unrevealing.

First, Shopen's analysis leads to a system of totally unconstrained base rules which are redundant as well as unwieldy. For example, Shopen's base rules should look like (110):
Notice that the left-hand symbol $X'$ can be expanded almost indefinitely because almost any constituent of a sentence can be ellipted (if appropriate contexts are given), and because every incomplete sentence should be generated by the base rules. Furthermore, many of the rules which rewrite $X'$ will be redundant, i.e. they will replicate the rules of full sentences like (110c).  

In addition, the left-hand symbol $X'$ need not be a constituent at all. (110f), for example, will generate sentences like (111b):

(a) na-nun Yenghi-hako nollayyo.
   I TOP with play-will
   'I will play with Yenghi.'

(b) na-nun Chelswu-hako-yo.
   I TOP with
   'For me, with Chelswu.'

Ross (1969:266) has also pointed out that this kind of analysis will bring a "mal-odorous" base rule like (112) into English:

(112) $NP \rightarrow NP^P$

in order to generate sentences such as (113)
(113) She was dancing, but I don't know who with.

This sort of a problem does not arise in pronominal anaphora. Let us compare (104-105) with (108-109). In the former, the pronominal anaphora involved do not destroy the surface grammatical structure of the sentence because pronominal traces are left behind. Sentences such as (104-105) are thus easily generated by the PS rules. In contrast, the zero anaphora involved in incomplete sentences destroys the grammatical structures of the sentences at the surface, so that any attempt to generate them directly by the PS rules cannot avoid constructing arbitrary PS rules such as (110a-g). Thus, applying the interpretive mechanism worked out for the pronominal anaphora directly to the zero anaphora does not seem adequate.

Second, Shopen argues that "people succeed in communicating in incomplete sentences without passing through the intermediary of complete sentences" (p.10). This is tantamount to claiming that syntactic properties of complete sentences play no role in well-formedness and interpretation of incomplete sentences (cf. Morgan 1973b), and that the surface forms and semantic representations are mapped directly by interpretive rules without going through any intervening stage involving the syntactic representation of a complete sentence.

However, there is some syntactic evidence showing that this claim is difficult to maintain. First, if the generation and interpretation of incomplete sentences have nothing
to do with corresponding complete sentences, it would be very 
difficult to explain the occurrences of the particles and 
grammatical morphemes appearing in incomplete sentences. 
For instance, consider the following:

(114)a ne nwukwu-lul mannasstakwu?
you who OM met
'Who did you meet?'

b Yenghi-lul yo. 
OM
'Yenghi.'

(115)a nwu-ka ikes-ul kulyessni?
who SM this OM painted,
'Who drew this picture?'

b Chelswu-ka yo. 
SM
'Chelswu.'

In Shopen's analysis, it would be difficult to explain why 
the subject and object particles ka and lul appear in these 
incomplete sentences. Shopen's proposal amounts to claiming 
that their occurrences in (114b-115b) are just accidental, 
which is not true. On the contrary, the occurrences of 
these particles should be understood as indicating that the 
NP's in these sentences are not isolated occurrences but 
parts of larger grammatical structures.

It might be argued that the occurrences of these par­
ticles in (114b-115b) can be explained in relation to their 
preceding sentences. However, this escape is not available 
in the case of exophoric ellipses such as (116-117):

(116)a ani sacangnim.
'Oh Mr. President'

b Chelswu.
(116)b ani Miss Lee-nun ettehkey yeki-l?
   oh TOP how-come here OM
   'Oh, Miss Lee, how are you here?'

(117) ani wuli ay-ka? (=41)
   my dear our son SM
   'My dear. Is my son ?'

It would be difficult to explain the occurrences of the case particles *lul* and *ka* in (116-117) without postulating intermediate sentential structures for these incomplete sentences.

There are also several grammatical morphemes whose occurrences in incomplete sentences cannot be accounted for in Shopen's analysis.

(118) nayil poypkeyssupnita.
   tomorrow see-will
   '(I) will see (you) tomorrow.'

(119) ese o-sip-siyo.
   quickly come HON
   'Please come in.'

If (118-119) were generated directly by a PS rule like $S \rightarrow VP$, it would be difficult to explain why the mood marker *keyss* 'will' and the honorific marker *sip* appear in (118-119), because both of them co-occur only with a grammatical subject.

Finally, Shopen (1972a:287-299) proposes that all inflectional forms of a word be listed as separate lexical items in the lexicon. This proposal is unavoidable in his analysis since it needs to generate sentences like (120) directly by PS rules:

(120) Needs an oil change, doesn't it?
Thus, all of the inflected forms, such as see, saw, seen, seeing, must be separate lexical items in the lexicon. It is easy to imagine how complicated the lexicon would be if this proposal is accepted.

Note that all of the above-mentioned problems are artifacts of an analysis which generates incomplete sentences directly by the base rules. Furthermore, it is not clear whether the semantic interpretation rules should be given the ability to conjure up the missing nodes and reconstruct the missing syntactic structures of incomplete sentences. For these reasons, we believe that grammatical facts involved in incomplete sentences and discourse anaphora must be accounted for in relation to full sentences.
Notes to Chapter I

1. 'Tongmu-wul-noli' by Tong-Yul Ko (1975). (1-4) are uttered by a single speaker. The example sentences in Chapter 1 and Chapter 2 are numbered throughout, because the two chapters are about the same subject. In romanizing Korean sentences the Yale system is used.

2. A similar definition was also proposed by Shopen (1972a: 65): "There is ellipsis when propositions are not fully realized in the grammatical form of utterances." Halliday and Hasan (1976:144) also give a similar definition.

3. 'Discourse ellipsis' should be distinguished from what we may call "sentential deletions", which include such rules as Equi-NP Deletion and Relativized Head NP Deletion. Consider:

(i) Yenghi-ka(∅ pap -ul mekki)-lul kecelhayssta.
    SM    rice OM eat-ing OM rejected
    'Yenghi refused to eat rice.' (∅: Yenghi)

(ii) (Chelswu-ka ∅ salanghanun)Yenghi-nun cwukessta.
    SM    love-ing TOP died
    Yenghi, whom Chelswu loves, died. (∅: Yenghi)

Deletions in (i-ii) are mainly governed by structural factors. For example, the zero anaphor of Equi-NP Deletion, as in (i), is always coreferential with the matrix subject Yenghi; that of Relativized Head NP Deletion in (ii) with the following head NP Yenghi.

Note that sentential deletions usually occur in an embedded sentence and are interpreted structurally with regard to a constituent in the matrix sentence. Discourse anaphora usually occur in simplex sentences or across sentence boundaries, and its phoric relation is not determined structurally but by semantico-pragmatic factors. The structurally determined deletions are sensitive to constituency, usually applying to a single constituent, whereas discourse ellipsis is not sensitive to constituency and any element of a sentence can be ellipted. There are other differences between the two which will be mentioned again in fn. 11.

4. For some recent works on discourse anaphora, see Morgan (1973b), Hinds (1976; 1978) and Halliday and Hasan (1976).

5. The PS rules take care of the possible and admissible arrangements of symbols in the language. The transformations, such as movement rules, are responsible for the possible rearrangements of symbols allowed in the system. A number of arguments have been advanced against this kind of conception of grammar (cf. Chafe 1970; Halliday 1967; Winograd 1972; Hudson 1976).
From this point of view, it would not be a mere coincidence that an interpretive analysis has been proposed for pronominalization and ellipsis rather than for other areas of syntax.

According to Chomsky (1957; 1965), the grammar of a language should distinguish acceptable from unacceptable sentences, and it seems that in order to do so, the grammar must take the context into consideration. Even Chomsky admits this: "It is not clear at all that it is possible to distinguish sharply between the contribution of grammar to the determination of meaning, and the contribution of so-called 'pragmatic considerations', questions of fact and belief and context of utterance." (Chomsky 1972:111)

A number of arguments have been advanced for this position (Winograd: 1972; Matthews 1979; Garcia 1975; Thrasher 1974). We will quote a passage from Thrasher (p.125):

"We have to stop playing what George Lakoff (1973) has quite properly characterized as 'the performance/competence game.' As long as performance is a dirty word we will never be able to understand what happens when we use language in the job of communicating."

This is why Choi explains (15-16) as the "avoidance of repetition" and (9-14) as "simplicity and emphasis" (Choi 1975:798-799).

Cataphoric reference is sometimes dictated structurally, as in a relative clause:

(i) (nay-ka ə manman) salam-un minam-iessta.
I SM met man TOP handsome was 'The man whom I met was handsome.' (ə: ku salam 'the man')

The zero anaphor of (i) refers to the relativized head NP. Except for these kinds of structurally determined cases, cataphoric reference is quite rare in discourse ellipsis. We will give one example:

(ii)a ə cham nollap-kwun.
very surprising 'ə is very surprising.'

b Chelswu-ka cwuk-tani.
SM died 'Chelswu died.'
(ə: Chelswu-ka cwuk-un kes 'that Chelswu died')

This rule has several characteristics. First, it is optionally applicable. Consider
(i) Chelswu-nun Yenghi-lul mannassta. kulayse
   TOP OM met       and-so
   'Chelswu met Yenghi, and so'

(ii)a Chelswu-nun Yenghi-wa kathi hakkyo-ey kassta.
   TOP         with together school to went
   'Chelswu went to school with Yenghi.'

   b ø₁ ø₂ kathi hakkyo-ey kassta.
      together school to went
   'ø₁ went to school ø₂ together.'
   (ø₁: Chelswu ; ø₂: Yenghi-wa 'with Yenghi')

In contrast, structurally determined deletion rules like Equi-NP Deletion and Relativized Head NP Deletion are all obligatory.

Second, the Ellipsis rule must be a variable rule. That is, what can be ellided is not specifiable and it can be a noun, a verb, a PP, etc. Moreover, Index 3 of the Ellipsis rule (56) can contain anything, including more than one sentence boundary. This is again a good contrast with other deletion rules, which are mostly sentence-bound and which can be formulated without variables, as in Equi-NP Deletion, Relativized Head NP Deletion, that-Deletion, by-Agent Deletion, etc.

Third, the Ellipsis rule is the last rule of the sentence grammar, even after the Scrambling rule. But deletion rules are not; e.g., Equi-NP Deletion is a cyclic rule. That the Ellipsis rule applies after Scrambling is shown in the following elliptical coordinate sentences (see Chap 2: 4.1).

(iii)a Chelswu-nun ocen-ey Yenghi-hako ø, Yengswu-nun
      TOP morning at      with      TOP
      ohwu-ey Swunhi-hako hakkyo-ey kassta.
      afternoon with school to went
      'Chelswu ø with Yenghi in the morning, and
      Yengswu went to school with Swunhi in the after-
      noon.' (ø: hakkyo-ey ka-ko 'went to school and')

   b *Ocen-ey Chelswu-nun Yenghi-hako ø, Swunhi-hako
      morning at       with       with
      ohwu-ey Yengswu-nun hakkyo-ey kassta.
      afternoon     TOP school to went
      'In the morning Chelswu ø with Yenghi, and with
      Swunhi in the afternoon Yengswu went to school.'
      (ø: hakkyo-ey ka-ko 'went to school and')

The Ellipsis rule operating in coordinate structures requires that the order of constituents in each conjunct of (iii) be
identical. This is why (iiib) is unacceptable. If the Scrambling rule, which itself is a post-cyclic rule, applies after the Ellipsis rule, the Scrambling rule must be constrained in a number of ways to prevent the generation of unacceptable sentences in coordinate structure, such as (iiib), which under the reverse ordering (i.e., Scrambling before Ellipsis) would be automatically prevented (cf. Williams 1977a; Bresnan 1976).

12(58) is acceptable as a discourse-initial sentence in appropriate contexts, e.g., when uttered in a restaurant. Note that one may say (58) even before the waitress has said a single word. This is so because the very presence of the waitress with an order memo implies that the customer is going to eat or order something.

13Problems involved in referential identity and identity in logical form will be discussed in 3.3.4, Chap 2. As for the structural identity, see Chap 2: 4.1.

14It is not clear whether definiteness is transformationally introduced or not (see Chap 2: 3.3.1).

15'Hakwi nonmwun' by Jin-Sung Suh.

16See Chapter 3.

17For example, Perlmutter has proposed an optical pronoun drop rule (Perlmutter 1972). However, there is some evidence that deletion of pronouns is not optional (cf. Gundel 1977; Chafe 1976).

18Also consider (6a-c) in 1.3.
CHAPTER II
DISCOURSE ELLIPSIS: A PROPOSED ANALYSIS

1. Introduction

1.1 The purpose of this chapter is to account for how an incomplete sentence, deficient in form, achieves its communicative function of conveying a full proposition. For this purpose, we will propose a theoretical model of a sort of language understanding system which takes an incomplete sentence as input and produces the semantic representation of its corresponding full sentence as output. This model, which is a model of the language user, is assumed to be part of our linguistic competence. While developing the machinery of such a system, we will touch on various syntactic and semantic properties of incomplete sentences.

The chapter is composed of four sections. Section 1 sketches the outline of the analysis and defines the notion of "incompleteness". Section 2 discusses the syntactic generation of incomplete sentences under this analysis. Section 3 describes the semantic interpretation processes of zero anaphora in discourse ellipsis. These will be divided into data-based interpretation rules and inference-based interpretation rules. The former is the main topic of 3.2, in which the importance of verb morphology to the interpretation process will be stressed. Section 3.3 will discuss the latter kind of interpretation rules, showing the significance of non-syntactic logical inferences for anaphora resolution.
Section 4 will take up the problem of ellipsis in coordinate structures, a somewhat different phenomenon from other kinds of discourse ellipsis.

1.2 In order to illustrate how the surface form of an incomplete sentence is associated with its semantic representation in the proposed analysis, let us take the following sentence, which I heard myself saying to one of my friends who had finished his program at the University of Hawaii and was going to return to Korea.

(121) #encey ka-si-pnikka?
     when go HON Ques
     'When go go?'

The same sentence might have been used in a discourse like (122):

(122)a hankwuk-i sal-ki cohciyo.
Korea SM live Nom good
'Korea is a good place to live in.'

b encey ka-si-pnikka? (=121)
when go HON Ques
'When go go?'

(121) is an incomplete sentence with an exophoric reference, and (122b) contains both exophoric and anaphoric ellipses.

Both (121) and (122b) mean (123):

(123) tangsin-un encey hankwuk-ey ka-si-pnikka?
     you TOP when Korea to go HON Ques
     'When do you go to Korea?'

We propose that the mapping of (121) or (122b) onto (123) is
accomplished in two steps, or more correctly, by the interaction of two linguistic processes which we will call "generation" and "interpretation."

When (121) and (122b) are uttered in these contexts, they are recognized as being incomplete by the addressee. This recognition comes in various ways. Language is a complicated system in which a number of sub-systems are closely interrelated to each other. These sub-systems appear in so-called syntagmatic relations, so that the presence of a certain element is often a necessary condition for the presence of another element in the sentence and vice versa. For instance, the presence of the subject honorification marker _sî_ in (121) and (122b) implies the presence of a grammatical subject. Furthermore, our semantic knowledge of the verb 'go' tells us that this verb needs at least an agent and a goal argument to complete its semantic make-up. At the same time, our grammatical knowledge of the same verb tells us that the agent argument normally functions as the grammatical subject and that the goal argument normally appears as a PP with the particle _-ey 'to_'.

Based on this sentence-internal analysis, we can reconstruct the underlying structure of (121) and (122b) as something like (124):

\[
(124) \left[ \left[ \Delta \right]_{NP} \left[ \text{encey} \right]_{NP} \left[ \Delta \right]_{PP} \left[ \text{kasipnikka} \right]_{VP} \right]_{VP} \right]_{S}
\]

This is the process of generating the underlying structure of
an incomplete sentence. What is important is the fact that the phrase marker (124) is all that our grammatical knowledge alone can contribute to our understanding of (121) and (121b). On this level we absolutely do not know what the dummy NP and PP stand for. All we know about (121) and (122b) are the sentence-internal relationships represented in (124).

In order to find out what these NP's refer to, we have to go beyond (121) and (122b) and look for some relevant sentence-external information in the context in which they are used. That context may be either situational, as in (121), or discoursal, as in (122b). This is the process of semantic or discourse interpretation of the anaphora, in which other cognitive faculties of human beings interact with the faculty of language to determine the value of the anaphora. By discourse interpretation rules, to be discussed in 3.2, the NP will be interpreted as tängsin 'you' and PP as hankwuk 'Korea'. These interpretations will produce (125):

\[
(125)[[\text{tängsin}_\text{NP}][[\text{ence}_\text{NP}\text{hankwuk-e}_\text{PP}[\text{kasipnikke}_\text{V}\text{VP}]]_S
\]

This is the correct semantic structure of (121) and (122b). Thus, by going through the two processes, syntactic generation and semantic interpretation, the surface forms (121) and (122b) are mapped onto the meaning in (125).

If this characterization is correct, then the analysis proposed here may be outlined schematically as follows:
(126) The conceptual model of the proposed analysis

- surface form
- generation process
- underlying structure
- interpretation process
- semantic structure

This model is somewhat oversimplified, as we point out in 2.2, but this is the basic thrust of our analysis. Note that this model takes the position of the hearer rather than that of the speaker.²

At the risk of some repetition, we will clarify how the proposed analysis differs from the two theories discussed in the preceding chapter.

First, since incomplete sentences are semantically indeterminate and vague to the intentions of the speaker, and since determining their semantic values is not a grammatical process, this semantic vagueness is represented as such in underlying structure in our analysis. By restricting the role of grammar to syntactic analysis and by reserving the task of finding out meaning to interpretation rules, our analysis reduces the power of grammar drastically. It also offers a conceptually more plausible account than the syntactic deletion approach.

Second, the semantic interpretation rules in our
analysis are not restricted to interpreting overtly manifested linguistic forms. As we will see later, they reveal how the human language faculty interacts with other cognitive faculties such as memory (see 3.2.1.1) and intelligence (see 3.3.3). The significant interaction between linguistic and non-linguistic factors involved in the use of incomplete sentences and discourse anaphora is simply lost in the deletion approach. In fact, the very question of whether such interactions exist or not does not arise.

Third, we disagree with Shopen (1972a) in two respects. We assume that incomplete sentences involving zero anaphora must be explained differently from cases of pronominal anaphora because, even though their phoric relations may be the same, their influences on surface syntax are different. Shopen needs semantic interpretation rules only, but our analysis needs not only semantic interpretation rules but also a device which restores or reconstructs the missing grammatical structure.

Finally, we maintain that incomplete sentences are always associated with corresponding full sentences, both in production and understanding. Thus, they are governed by the same mechanism as are full sentences, and are generated by the same base rules that generate full sentences.

1.3. Before we leave this introductory section, we will discuss what the lexicon will look like in our grammar. We assume that, in the lexicon, the lexical entry of a verbal
predicate (a verb or an adjective) carries at least four kinds of information which are crucial to the well-formedness of the propositional and syntactic structures of the sentence.

First, it specifies the logical argument structure of the verbal predicate, i.e. the number and kinds of arguments which are intrinsically needed for the make-up of the proposition. Such information may be given in terms of Fillmorean deep semantic case relations with the definitions given thereof (Fillmore 1968, 1977; I.S.Yang 1972:6-7).

Incomplete sentences are semantically well-formed, containing a well-formed propositional structure. Since the semantic well-formedness of a proposition is determined by the logical argument structure of its verbal predicate (Parisi and Antinucci 1976), the verbal predicate in the surface form of an incomplete sentence and its logical argument structure specified in the lexicon should give us a clue to the propositional structure of the underlying sentence. This clue can in turn be utilized to reconstruct the full syntactic structure. In this sense, the verbal predicate and information on its argument structure are crucial to the generation and interpretation processes (see 2.1 and 3.2).

Second, when the proposition is realized as a sentence, some of the arguments are assigned nuclear syntactic functions of the sentence. Information like this is also specified in the lexicon. The nuclear syntactic functions of a sentence are the subject, the object, and the topic in Korean, which we assume has the following base rules:
Other semantic arguments that are not assigned a nuclear syntactic function are realized as PP's in the VP. Their semantic and syntactic functions are expressed by postpositions. Incidentally, there is no fixed order among the PP's in the VP.

The other two kinds of information are the specification of obligatoriness or optionality of arguments, and the semantic selectional restrictions between the predicate and the other arguments.

The following is a sample lexicon in our grammar. (Optional arguments are put in parenthesis and selectional restrictions are omitted.)

\[(128)\]

\(\text{(128)a} \quad \text{veyppute 'pretty':} \quad \{\text{+Pat}\} \quad \{\text{+Exp}\} \quad \{\text{+Loc}\} \quad \{\text{+Tm}\} \quad \{\text{+Sub}\} \quad \{\text{+Top}\} \quad \)

\(\text{b} \quad \text{ttaylita 'hit':} \quad \{\text{+Agt}\} \quad \{\text{+Pat}\} \quad \{\text{+Ins}\} \quad \{\text{+Com}\} \quad \{\text{+Tm}\} \quad \{\text{+Sub}\} \quad \{\text{+Obj}\} \quad \{\text{+Loc}\} \quad \)

\(\text{c} \quad \text{hwumchita 'steal':} \quad \{\text{+Agt}\} \quad \{\text{+Pat}\} \quad \{\text{+Src}\} \quad \{\text{+Com}\} \quad \{\text{+Sub}\} \quad \{\text{+Obj}\} \quad \{\text{+Ins}\} \quad \{\text{+Tm}\} \quad \{\text{+Loc}\} \quad \)

\(\text{d} \quad \text{nehta 'put':} \quad \{\text{+Agt}\} \quad \{\text{+Pat}\} \quad \{\text{+Loc}\} \quad \{\text{+Src}\} \quad \{\text{+Com}\} \quad \{\text{+Sub}\} \quad \{\text{+Obj}\} \quad \{\text{+Ins}\} \quad \{\text{+Tm}\} \quad \)

Information such as that in (128) provides us with a basis to define which sentences are to be considered
"incomplete". So far we have assumed that if semantic material in proposition is not realized in surface structure, the sentence is incomplete. But consider the following sentences:

(129) Chelswu-nun Yenghi-lul ttaylyessta.
    TOP OM hit
    'Chelswu hit Yenghi.'

(130) Chelswu-nun ton -ul hwumchyessta.
    TOP money OM stole
    'Chelswu stole money.'

Are sentences (129-130) incomplete or not? Logically speaking, they are all incomplete because we are not told how Chelswu hit Yenghi, where he stole money, and so on. But the speaker of (129-130) may want to assert just that Chelswu hit Yenghi and that he stole money, without bothering how and where. So we maintain that these are full sentences. Note that the ellipted elements in these sentences are all optional. Thus, omitting an optional argument may or may not lead to intuitions of incompleteness, and in normal situations--e.g., where there is no specific intention of describing the instrument of hitting and the source of stealing in (129-130)--those sentences would be considered "full."

Compare (129-130) with (131-132):

(131) Chelswu-nun chayk-ulo ttaylyessta.
    TOP book with hit
    'Chelswu hit with a book.'

(132) Chelswu-nun Yenghi-eykeyse hwumchyessta.
    TOP from stole
    'Chelswu stole from Yenghi.'
Whatever the communicative intentions of the speaker, (131-132) are all felt to be incomplete. Note that they contain an optional argument in place of an obligatory argument. Thus, the correct generalization seems to be that the lack of an obligatory argument in a proposition always leads to an intuitive feeling of incompleteness, whereas omission of an optional argument need not. Compare the (a) sentences with the (b) sentences in (133-134). The former lack optional arguments, while the latter lack obligatory arguments.

(133)a na-nun onul  phyenci-lul patassta.
   I TOP today letter  OM received
   'I received a letter today.'

   b na-nun onul  Yenghi-eykeyse patassta.
   I TOP today from received
   'I received from Yenghi today.'

(134)a Chelswu-nun Yenghi-lul mannassta.
   TOP OM met
   'Chelswu met Yenghi.'

   b Chelswu-nun hakkyo-eyse mannassta.
   TOP school at met
   'Chelswu met at school.'

The (b) sentences are felt to be incomplete, but the (a) sentences are not, even though neither (133a) nor (133b) specifies when and where the speaker received the letter, and neither (134a) nor (134b) gives when and with whom Chelswu met Yenghi. Thus, incomplete sentences might be defined as follows:

(135) A definition of 'incompleteness'

A sentence is incomplete when it lacks one or more of the obligatory arguments of its predicate.
However, it soon becomes clear that this definition does not always fit our intuitions about incompleteness. Consider the following anaphoric ellipses:

\[(136)\] a. Chelswu-nun Yenghi-eykeyse mwues-ul hwumchyesnii?
  TTOP from what OM stole
  'What did Chelswu steal from Yenghi?'

b. ton-ul hwumchyessta.
  money OM stole
  '(He) stole money.'

\[(137)\] a. ne-nun Yenghi-eykeyse mwues-ul patassni?
  you TOP from what OM received
  'What did you receive from Yenghi?'

b. phyenci-lul patassta.
  letter OM received
  '(I) received a letter.'

\[(136b-137b)\] are definitely incomplete sentences. Definition (135) predicts that if all the obligatory arguments are restored to (136b-137b), they will become full sentences. But this prediction is not borne out and the resulting sentences (138-139) are still felt to be incomplete.

\[(138)\] Chelswu-nun ton-ul hwumchyessta. (=\(130\))
  'Chelswu stole money.'

\[(139)\] na-nun phyenci-lul patassta. (=\(133a\))
  'I received a letter.'

The same sentence is felt to be complete, as in (130), and incomplete, as in (138); the same is true for (133a) and (139).

(136-137) show that incompleteness of a sentence cannot be determined in absolute terms, i.e., on the basis of the incomplete sentence alone, but only in relation to the discourse context in which it appears. The correct
generalization seems to be that if an optional argument has already appeared in the preceding text, e.g. the 'source' argument in (136a-137a), it must also be included in the following sentence; otherwise, the sentence becomes incomplete. Thus, we define an incomplete sentence as follows:

(140) **Definition of incompleteness**

A sentence is incomplete if it lacks an obligatory argument of its predicate and/or if it lacks one of the optional arguments already appeared in the antecedent sentence.

Another theoretical question we can entertain in this connection is whether or not incomplete sentences are fully legitimate sentences. This question is related to another issue we raised in Chap 1: 1.3—the grammaticality and un-grammaticality of incomplete sentences like (7-8). We repeat the sentences here for ease of reference.

(7) Chelswu-nun Yenghi-lul chacawassta. (kuliko)
    TOP OM visited and 'Chelswu visited Yenghi, and'

(8)a Chelswu-nun Yenghi-eykey semmwul-ul cwuessta.
    TOP to present OM gave 'Chelswu gave Yenghi a present.'
    b Chelswu-nun semmwul-ul cwuessta.
        'Chelswu gave a present.'
    c Yenghi-eykey semmwul-ul cwuessta.
        '(He) gave a present to Yenghi.'
    d semmwul-ul cwuessta.
        '(He) gave a present.'
    e *cwuessta.
        '(He) gave.'

(8b-e) are all incomplete sentences on both superficial and formal grounds. But (8b-d) are acceptable and (8e) is not.
The only difference between the two types is that the former sentences can get the interpretation of a full proposition since the missing arguments can be supplied by the context, whereas the latter cannot. Thus, we maintain that as long as an incomplete sentence can get the interpretation of a full proposition, as in (8b-d), it should be considered a fully legitimate sentence, whatever its superficial shape. However, if an incomplete sentence cannot get such an interpretation, as in (8e), it is not a sentence.

This also raises a question as to the adequacy of defining the sentence on formal grounds. For example, Bloomfield's (1933:170) definition based on grammatical non-substitutability and "distributional independence" predicts that an incomplete sentence is not itself a sentence (see Lyons 1969:175).

2. The Generation Process

The process of reconstructing the underlying structure of an incomplete sentence, based on various syntactic and semantic cues available from the surface structure will be called the "generation process." Most important among them are the base rules of the language, various sentence-internal co-occurrence restrictions and the argument structure of the verbal predicate. We will discuss nominal ellipsis first; verbal ellipsis will be taken up later. Nodes in the underlying structure thus generated which are not matched by surface constituents will be represented by deltas.
2.1 **Nominal Ellipsis**

When an incomplete sentence with an exophoric nominal ellipsis is given, we assume that the first operation to take place is "Parsing", which assigns labels to each of the surface constituents from a set of grammatical category symbols such as N, V, Adv., and so on. We will take the following sentences to illustrate.

(141)a *tto poypkeyssupnita.*  
again see will HON  
'See again.'

(142)a *ka-si-1 ttay pwulseeyyo.*  
go HON time call HON  
'When go, call.'

Parsing will yield the following parsed structures for (141a-142a):

(141)b \([tto]_{Adv} [poypkeyssupnita]_{V}\)

(142)b \([ka-si-1]_{V} ttay [pwulseeyyo]_{V}\)

The next operation posited is "'Catalysis' by Base Rules", a term borrowed from Siertsema (1965). Catalysis by Base Rules matches parsed structures against the base rules of the language. For example, (141b-142b) are checked against the PS rules (127). This will "encatalyze" the former to (141d-142d) through (141c-142c):

(141)c \([tto]_{Adv} [poypkeyssupnita]_{V} \) \(v_{VP}\) (by (127c))

\(d \left[ tto \right]_{NP} \left[ tto \right]_{Adv} [poypkeyssupnita]_{V} \) \(v_{VP}\) \(s\)' (by (127b))
The rule (127a) --$S \rightarrow (NP)S'$-- does not apply here because the topic NP is optional and the rule itself does not tell us whether or not it should apply to (141d-142d).

The third operation we hypothesize for the generation process is "Catalysis by the Argument Structure of the Verbal Predicate" (henceforth "Catalysis by Predicate"). This operation checks the encatalyzed structures (141d-142d) against the following information given in the lexical entries for pota 'see', kata 'go', and pwuluta 'call':

(143)a pota 'see': [+Agt] [+Pat] [+Ins] [+Com] [+Loc] [+Tm]

(143)b kata 'go': [+Agt] [+G1] [+Com] [+Ins] [+Loc] [+Tm]

(143)c pwuluta 'call': [+Agt] [-Subj] [+Pat] [+Ins] [+Loc] [+Tm]

Based on this information, (141d-142d) will be further encatalyzed to (141e-142e) respectively:

(141)e $[\Delta]_{NP}[\Delta]_{NP}[tto]_{Adv}[$poykeyssupnita$]_{VP}S,$

(142)e $[\Delta]_{NP}[[kasi]_{VP}]S,ttag[[\Delta]_{NP}[[\Delta]_{NP}[$pwuluseyyg$]_{VP}]_{VP}]S,$

(141e) and (142e) are the correct underlying structures for (141a) and (142a), respectively. Note that (141e-142e) are all that the grammar knows on the basis of the intra-
sentential criteria of (141a-142a). The particular referents of these cases of zero anaphora must later be resolved by discourse interpretation rules.

Anaphoric nominal ellipses are generated in the same way as the exophoric ones, except that the optional arguments already included in the antecedent sentence must be taken into consideration. Let us take the following sentence for illustration.

(144) ohwu-eyn nwukwu-lul pwululkkayo?
    afternoon who OM call-will
    'Who shall we call in the afternoon?'

(145a) Chelswu-lul pwulupsita.
    OM call-let's
    'Let's call Chelswu.'

(145a) will be generated as follows:

(145)b Parsing: [Chelswu-lul]_NP[pwulupsita]_V

   c Catalysis by Base Rules:
   [\[\Delta\]_NP[[Chelswu-lul]_NP[pwulupsita]_V]_VP]_S, 9
   (by (127b-c))

(145c) is certainly an acceptable underlying structure for (145a), if it is a discourse-initial sentence. However, note that (145a) is a discourse-bound sentence. As the definition (140) suggests, this sentence may contain optional zero anaphora. Actually, the underlying structure of (145a) must contain the 'time' argument.

The question is whether the generation process should
look for the discourse context to find out what optional argument have already been specified in antecedent sentences, e.g. (144). It seems that this is not necessary. All we need is a convention whereby the syntactic slots of optional arguments, those which are specified by the lexical entries of verbals in the lexicon, are to be generated in the underlying structure, even though they may not be realized on the surface. This convention applies both to exophoric and anaphoric ellipses. Later, semantic interpretation rules will scan the situational or discoursal context and any of the optional argument slots not interpreted by the semantic interpretation rules will remain unrealized.

The above convention will further encatalyze (145c) to (145d), based on the PS rule (127c) and the lexical information of the verb *pwuluta* 'call' given in (143c):

(145)d Catalysis by Predicate

\[
[[\Delta]_\text{NP}[[\text{Chelswi}]_\text{NP}[\Delta]_\text{PP}[\Delta]_\text{PP}[[\text{pwulupsita}]_\text{YP}]]_\text{VP}],
\]

The three optional PP nodes in (145d) are to accommodate any or all of the arguments of 'time', 'instrument' and 'locative' available in the context. So, (145d) is the correct underlying structure of (145a) and the first NP and one of the PP's will later be associated with *wuli* 'we' and *ohwu-ey* 'in the afternoon' respectively by discourse interpretation rules. (see 3.2.1).

What is clear from the above discussion is that the
logical argument structure of the predicate is crucial to the generation process, at least for those involving nominal ellipsis. Actually there are many cases in which Catalysis by Predicate is the only cue available. Consider the following sentences of the so-called "X-nun Y-ta" pattern (Kuno 1973:33-34).

(146)b #na-nun nayngmyen-ipnita.  
I TOP cold noodle be  
'I am cold noodles.' (or 'I am for cold noodles.')

(147)a i pwuntan-un kyosil -ul matass-ko  
second row TOP classroom OM took-care-of  
'The second row took care of the classroom, and'

b wuli-nun yulichang-iessta.  
we TOP window was  
'We were the windows.' (or 'We were assigned to the windows.')

(148)a Yenghi-nun cip -ulo ka-la.  
TOP house to go  
'Yenghi, go home.'

b Chelswu-nun pyengwen-ita.  
TOP hospital be  
'Chelswu is the hospital.' (or 'Chelswu is for the hospital. ')

It might be argued that (146b-148b) are not elliptical at all. To take (146b) for an example, it may be argued that na 'I' is the subject and nayngmyen 'cold noodle' is the nominative complement. Thus, (146b) may be analyzed as having the following configuration, both in the underlying and derived structures:

(149)[[na]_NP[[nayngmyen]_NP[ita]_V]_VP]_S,'
However, it is clear that (149) is not the correct underlying structure, if (146b) is compared with (150):

(150) na-nun sensayng-ipnita.
     I TOP teacher be
     'I am a teacher.'

As for (147b-148b), it might be argued that (148b), for example, is derived from the structure underlying (151):

(151) Chelswu-nun pyengwen-ulo ka-la.
     TOP hospital to go
     'Chelswu, go to the hospital.'

from which the verb kata 'go' and the particle ulo 'to' are deleted under identity, and to which the verb ita 'be' is added. However, this deletion analysis does not seem plausible. First, the hypothesized deletion transformation must be allowed to build structure, but it is highly doubtful whether transformations should be allowed to do so (Chomsky 1977). For example, the alleged underlying structure is

(152)a[[Chelswu]\_NP[[pyengwen-ulo]\_PP[kala]\_\_VP]\_S,

while its derived structure (152b) differs drastically:

(152)b[[Chelswu]\_NP[[pyengwen]\_NP[ita]\_\_VP]\_S,

Second, the derived structure (152b) does not correctly reflect the meaning of (148b), because Chelswu cannot be equated with a hospital. The derived structure is thus not semantically correct. Third, this deletion analysis cannot offer a unified explanation for both (146b) and (147b-148b).
Finally, the ita-insertion is rather ad hoc in the sense that we cannot explain why ita 'be', and not some other form such as the pro-verb hata 'do-so', should be used. (Note that what is deleted is an action verb 'go'.)

We argue that the correct underlying structures for (146b-148b) must be something like (153), if (148b) is taken as an example:

\[
\begin{align*}
\text{(153) } & \quad [\text{Chelswu}_\text{NP} \left[ \Delta_{\text{NP}} \left[ \text{pyengwen}_\text{NP} \left[ \text{ita}_\text{V} \right] \text{VP} \right] \right] S, S]
\end{align*}
\]

But, how have we arrived at this kind of underlying structure?

It is well-known that the predicate of a proposition need not be verbal, but can also be a noun. When a sentence contains the copula ita, we might say that the predicate is not the copula but the whole of 'NP+ita' ('NP+be' in English). For example, the following (154-155)

\[
\begin{align*}
(154) \quad & \text{John is a teacher.} \\
(155) \quad & \text{Mary is tired.}
\end{align*}
\]

must be represented logically as (156-157), respectively:

\[
\begin{align*}
(156) \quad & \text{TEACHER (John)} \\
(157) \quad & \text{TIRED (Mary)}
\end{align*}
\]

The representations (156-157) mean that each of these predicates is an attribute of the argument -- each attributes a property to its argument.
If (149) and (152b) are taken for the derived constituent structures for (146b) and (148b) respectively, then they must have the following logical representations:

(158) NAYNGMYEN ( na )
cold noodle I

(159) PYENGWEN (Chelswu)
hospital

However, (158-159) are not well-formed semantically, because the predicates cannot be attributes to their respective arguments. If the speaker of these sentences is cooperating in the conversation, there must be another nominal argument in (146b) and (148b), to which these predicates are attributed in a semantically well-formed way. This knowledge will encatalyze (146b) and (148b) to (160a-161a):

(160)a [[Δ]_NP [[nayngmyen]_NP[ita]_V ]_VP ]_S
(161)a [[Δ]_NP [[pyengwen]_NP[ita]_V ]_VP ]_S

where the whole VP is a predicate. These encatalyzed structures are returned to the Catalysis by Base Rules and further analyzed to (160b-161b) by the PS rule (127a):

(160)b [[na]_NP [[Δ]_NP[[nayngmyen]_NP[ita]_V ]_VP ]_S, ]_S
(161)b [[Chelswu]_NP[[Δ]_NP[[pyengwen]_NP[ita]_V ]_VP ]_S, ]_S

These are the correct underlying structures for (146b) and (148b).
2.2 **Verbal Ellipsis**

Verbal ellipsis is somewhat different from nominal ellipsis in structural generation. In the latter, the logical argument structure of the verbal predicate played an important role; in the former, however, the verbal predicate itself is missing on the surface. Thus, this useful semantic and syntactic cue is not available. So, the base rules and other syntagmatic co-occurrences of the language must be used instead.

Let us take the following exophoric and anaphoric verbal ellipses to illustrate how verbal ellipsis sentences are generated.

(162)a ani, yeki-l ettehkey?
my-dear here OM how
'My dear, how come you are here?'

(163)a ca, ili-lo.
well here to
'This way, please.'

(164) Speaker A: emeni-nun eti-ey kasyessni?
mother TOP where went-HON
'Where did your mother go?'

(165)a Speaker B: hakkyo-ey yo.
school to
'To school.'

As in the cases of nominal ellipsis, the first operation to take place is Parsing, which analyzes (162a-165a) to (162b-165b) respectively:

(162)b [yeki-l]_{NP} [ettehkey]_{Adv}

(163)b [ili-lo]_{PP}
Here too, the next operation to take place is Catalysis by Base Rules, which will check the parsed structures (162b-165b) against the base rules (127). Rule (127c) indicates that the verbal predicate—an obligatory constituent—is missing in (162b-165b). Thus, they are encatalyzed respectively to (162c-165c):

(162)c [yeki-]NP [ettehkey]Adv[Δ]V (by (127c))
(163)c [ili-lo]pp [Δ]V (by (127c))
(165)c [hakkyo-ey]pp [Δ]V (by (127c))

The base rules are not the only syntactic cues available. The occurrences of case markers such as lulu for the object, lo for the goal, ey for the locative, indicate that these NP's cannot be the subjects of these sentences. This knowledge, together with the PS rules (127b-c), further encatalyze (162c-165c) to (162d-165d):


(162d-165d) are the correct underlying structures for (162a-165a).

(162a-165a) are the simplest cases of verbal ellipsis,
in which case markers unmistakably show that the nominal constituents on the surface are not the grammatical subjects in their respective sentences. However, there are many cases in which structural catalysis is blocked by the lack of case markers. These cases show that case markers are very important syntactic cues operating in the generation process.

Consider the following texts:

(166)a i os -un pissayo.
    this clothes TOP expensive
    'This garment is expensive.'

    b ce os -to yo.
    that clothes too
    'That one is, too.'

(167)a ese tu-sey-yo.
    hurry eat HON
    'Help yourself.'

    b sensayngnim-to yo.
    you too
    'You, too.'

(168)a apeci -ka eti-ey key-si-ni?
    father SM where at be HON
    'Where is your father?'

    b secay yo.
    study
    'The study.'

The (b) sentences of (166-168) will be parsed to (169a-171a) respectively:

(169)a [ce os-to] NP
(170)a [sensayngnim-to] NP
(171)a [secay] NP
The Catalysis by Base Rules will encatalyze (169a-171a) to (169b-171b):

\[(169)b \, [\text{ce os-to}]_{\text{NP}} [\Delta]_V \, (\text{by (127c)})\]
\[(170)b \, [\text{sensayngnim-to}]_{\text{NP}}[\Delta]_V \, (\text{by (127c)})\]
\[(171)b \, [\text{secay}]_{\text{NP}} [\Delta]_V \, (\text{by (127c)})\]

However, we cannot tell whether the NP's in (169b-171b) are subjects or constituents of the VP, such as objects. This is because, unlike (162a-165a), there are no case markers in (166b-168b). Thus, the syntactic catalysis based on the sentence-internal cues cannot proceed further. This means that, given surface forms such as (166b-168b), the grammar cannot reconstruct the syntactic structures underlying these sentences, and that the grammar must consult the discourse context for more clues.

However, consulting the discourse context for information is no longer generative; so, we propose that our language understanding system proceeds to the next stage of the 'interpretation process.' As we will discuss later in 3.2.2, one of the interpretive rules involved in verbal ellipsis is the rule of "Verb Copying", which copies the verbal predicate of the preceding sentence into the unmaterialized V slot of the underlying structure reconstructed by the generation process. The same rule has been proposed for English by Williams (1977a:105). Note that this rule is not generative, but rather interpretive by nature.
will copy the verbal predicates of (166a-168a) into the reconstructed $V$ nodes of (169b-171b) to produce the following:

$$(169)c\ [\text{ce os-to}]_\text{NP}\ [\text{pissayo}]_V$$

$$(170)c\ [\text{sensayngnim-to}]_\text{NP}\ [\text{tuseyyo}]_V$$

$$(171)c\ [\text{secay}]_\text{NP}\ [\text{keyseyyo}]_V$$

$(169c-171c)$, once they have received verbal interpretations, will be returned again to the process of syntactic generation. The verb \text{pissata} 'expensive' is an one-place predicate, whereas the verbs \text{tulta} 'eat' and \text{issta} 'be' are all two-place predicates. $(169c-171c)$ will be encatalyzed to $(169d-171d)$ on the basis of the argument structures of the verbal predicates and the occurrences of the honorific marker \text{si}, which cooccurs only with the subject.

$$(169)d[[\text{ce os-to}]_\text{NP}[[\text{pissayo}]_V}_\text{VP}]_S$$

$$(170)d[[\text{sensayngnim-to}]_\text{NP}[[\Delta]_\text{NP}[[\text{tuseyyo}]_V}_\text{VP}]_S$$

$$(171)d[[\Delta]_\text{NP}[[\text{secay-ey}]_\text{PP}[[\text{keyseyyo}]_V}_\text{VP}]_S$$

These are the final underlying structures for $(166b-168b)$, and they will be subjected to the second cycle of the interpretation process to resolve the remaining anaphora.

The above discussion suggests two things. First, when an incomplete sentence contains both nominal and verbal ellipses, the latter must be resolved prior to the former. Second, the generation and interpretation processes are actually not
independent, but rather interdependent, operations. The latter conclusion is in agreement with what Winograd (1972) and Schank (1972) have found in artificial intelligence research:

Syntactic processing is by no means done in isolation from the rest of the processes. It would be non-sense to consider that syntactic processing is completed before conceptual processing ([similar to our interpretation process]) begins. The two processors 'talk together' while they work. (Schank 1972:559)

Bresnan (1976:18-19) has come to similar conclusions:

In the realization of this model of grammar the active processing components corresponding to the syntax and the semantics can be assumed to operate simultaneously.

Let us take one further example, which also shows why verbal ellipsis requires a cyclic application of the generation and interpretation processes. Consider

(172)a nwu-ka ilen cis -ul hayssni?
who SM this thing OM did
'Who did this kind of thing?'

b cey-ka yo.
I SM
'I did.'

The structure and meaning of (172b) are resolved in the following way:

(173)a Parsing: [cey-ka]NP


c Interpretation (Verb Copying):


Now, the derived structure (173c) should again be subject to catalysis, because it is semantically ill-formed; the
predicate  

\[ \text{hata 'do'} \] requires at least an agent and a patient.

\[(173)d\text{ Catalysis by Predicate:}\]
\[
[ce^\text{NP}[[\Delta^\text{NP}[\text{hayssyeyo}^\text{V}]]^\text{VP}]_S
\]

\[e\text{ Interpretation:}\]
\[
[cey^\text{NP}[[[\text{kulen cis-ul}^\text{NP}[\text{hayssyeyo}^\text{V}]]^\text{VP}]_S
\]

On the basis of the above discussion, the conceptual model of the analysis proposed in 1.1 must be revised as follows:

\[(174)\text{Revised conceptual model of the proposed analysis}\]

\[
\text{surface form} \xrightarrow{\text{generation process}} \text{underlying structure} \xrightarrow{\text{interpretation process}} \text{semantic structure}
\]

We have so far tried to show how the underlying structures of incomplete sentences involving nominal and/or verbal ellipses are to be identified or recovered by various intrasentential syntactic and semantic cues. However, this syntactic generation process is not always successful, and there are a few sentence types which defy all our attempts to resolve their structure and which remain vague. These are mostly forms with exophoric verbal ellipses. Consider
(175a) yepe, nay cikap.
darling my wallet
'Darling, my wallet.'

(175a) will be analyzed to (175b):

(175b) [nay cikap]_NP [Δ]_V

But this is all we can do about (175a). We do not know whether the NP is the subject or the object in the underlying structure. Since the zero anaphor has exophoric reference, the interpretive rule of Verb Copying cannot apply to (175b). Therefore, the syntactic structure of (175a) remains unsolved. The zero anaphor can be interpreted in various ways: _ati-ey issci_ 'where is it'; _owusio_ 'give me', etc. This fact is neither surprising nor problematic to the proposed analysis, because it does not claim that every sentence can be interpreted. But sentences like (175a) are problems for the deletion analyses.

3. The Interpretation Process

3.1 Universe of Discourse: A Unified Concept of Context

The interpretation process is the semantic operation of identifying the correct referent of a zero anaphor out of a set of possible candidates provided by the situational or discoursal context. In the following two sections, some of the most important and productive interpretation rules will be discussed.
But first, let us consider the notion "context." We argue that the distinction between situational and discoursal contexts is real, but that its theoretical correlate—the distinction between exophoric and anaphoric references—is spurious. In other words, the fact that there are two kinds of contexts does not mean that the distinction must be reflected in a theory seeking to explain phoric relations. This means that a unified theory of phoric relations or contexts is possible.

Let us consider the relationship between referents and anaphora (zero or pronominal). A predominant view in linguistics is that an anaphor refers to its linguistic antecedent, which in turn refers to a real world object, the referent. In this view, an anaphor does not refer directly to its referent but refers only indirectly through its linguistic antecedent. Let us take the following sentence for illustration.

(176)a ecey Reagan-i i kos-ey wassessci. (=S₁)
yesterday SM this place came
'Yesterday Reagan came here.'

b Ø cal sayngkyesstekwuman. (=S₂)
well looking
'Handsome.' (Ø: Reagan)

It has been assumed that the zero anaphor of (176b) refers to the linguistic antecedent Reagan of (176a), which in turn refers to the real world object, Mr. Reagan, the President of the United States. When the linguistic antecedent is associated with the real world object, it will bring in a set of
properties pertinent to the object, Mr. Reagan. This view might be shown schematically as follows:

(17')

President

\[\text{Reagan} \leftarrow \emptyset \]

\[\text{(S}_1\text{)} \quad \text{(S}_2\text{)}\]

\text{a Westerner}

\text{handsome}

\text{a Republican}

real world

text

Now, let us consider a slightly different text:

(178)a  
\text{ecey Reagan-i i kos-ey wasseisci. (=S}_1\text{)}

\text{yesterday SM this place came}

'Yesterday Mr. Reagan came here.'

b  
\text{ku sepwu-nyesek-i mopsi ttetultekwuman. (=S}_2\text{)}

\text{that westerner SM very noisy}

'That westerner was quite noisy.'

in which the full NP anaphor (an epithet in this case) of (178b) refers to Mr. Reagan. Since there is no linguistic identity between \text{ku sepwu-nyesek} 'that westerner' of (178b) and \text{Reagan} of (178a), both of them refer to the real world object, Mr. Reagan, directly and independently. This relation may be shown as follows:

(179)

\text{Reagan}\leftarrow \text{ku sepwu-nyesek}

\[\text{(S}_1\text{)} \quad \text{(S}_2\text{)}\]

\text{a Westerner}

\text{President}

\text{handsome}

\text{a Republican}

real world

text

According to this view of anaphoric reference, then, the
same coreference phenomenon may be explained in two different ways, depending upon the presence or absence of an identical linguistic antecedent in the text.

Much the same is observed when an anaphor has exophoric reference as in the following sentence:

(180) (Mr. Reagan has just left a town after his campaign. One citizen says to another:)

# O cal sayngkyesstekwuman. (=S₁)
    well looking
    'Handsome.' (Ø: Reagan)

Again, since there is no linguistic antecedent available in the text, the relationship between the zero anaphor and its referent is a direct one:

This view of the phoric relationship between an anaphor and its referent is undesirable for several reasons. First, the phoric relation is treated in two different ways: it is considered indirect if an identical linguistic antecedent is available in the text; direct if not. Thus, the same reference is handled differently in (177) and (181). One undesirable outcome of this treatment has already been pointed out in reviewing Hankamer and Sag's mixed theory. Note that the there is no a priori reason why reference should go through
the intermediate stage of a linguistic antecedent, just because such an antecedent happens to be available in the text. I believe that this view stems from too close an adherence to linguistic forms in generative grammar (Chomsky 1965).

A second problem has also been mentioned. A number of problems about the status of "identity" are difficult to handle (3.3.4 and 4.1).

Third, tying an anaphor to a linguistic antecedent under strict linguistic and coreferential identity binds approaches such as the syntactic deletion analysis, which is based on this view, into a strait jacket. The deletion analysis ends up being inadequate even from a purely descriptive point of view. Problems such as inference-related anaphora and sloppy identity are most typical. The former problem has been discussed in Chap 1: 3.1.3. The latter will be dealt with in 3.3.4.1 of this chapter.

We adopt a different assumption here. We hold that, in order to explain the phenomenon of zero reference in an adequate way, we must shed the strait jacket of linguistic identity. We adopt this view because many cases of zero anaphora cannot be described in terms of formal linguistic identity (see 3.3.3). For this reason, we hold a different view of the relationship between an anaphor and its referent, a view which has been well accepted among philosophers (cf. Kripke 1972) and which has also been familiar in linguistics (Karttunen 1976; Kuno 1973). This is the notion of "universe-of-discourse", a notion which I believe is crucial to the
understanding of anaphoric reference.

Let us assume that, when a discourse starts, the discourse participants enter a hypothetical world which we will call the "universe-of-discourse" (henceforth, UOD), a shared conceptualization of the world in which the discourse takes place. The UOD consists roughly of the information accumulated by the discourse. Let us take (17a-b) for illustration. When (17a) is uttered, the linguistic form Reagan will refer to the real world object, Mr. Reagan. Then, a mental image of Mr. Reagan will be registered in the UOD as a semantic concept. This semantic concept of Mr. Reagan will consist of all the relevant properties pertaining to Mr. Reagan: as President of the United States, as a man from the West, as tall and handsome, as a Republican, as a man whose wife's name is Nancy, and so on. (This is the so-called "cluster concept" theory (Kripke 1972) of reference in philosophy.)

Our hypothesis is that what the zero anaphor of (17b) refers to is not the real world object, Mr. Reagan, but the semantic concept of Mr. Reagan registered in the UOD. This semantic concept, which we will call a "discourse entity", is in turn associated with the real world object. This view might be shown schematically as follows:
Compare (182) with (177). Note that the zero anaphor in $S_2$ is related to the linguistic antecedent in $S_1$ indirectly, and that it refers directly to the discourse entity REAGAN (which we will capitalize to distinguish it from linguistic form) existing in the UOD. Also note that a discourse entity is not a real world object, but represents a sum of subjective intensional correlates of the real world object, which are shared by the discourse subjects. It can include even personal evaluations such as 'fool', 'an actor', etc.

The full NP anaphora of (178b) can be explained in the same fashion.

The full NP anaphor in $S_2$ refers to the discourse entity REAGAN. In this case, the main feature in focus is that he is a Westerner, which has been brought into the UOD by the mentioning of Mr. Reagan in $S_1$. 
The exophoric zero anaphor of (180) is subject to the same explanation. As we will argue later, a UOD exists even at the beginning of a discourse. In this particular case, the semantic concept of Mr. Reagan has already been listed in the UOD, and what the zero anaphor refers to in (180) is this discourse entity in the UOD, which is in turn associated with the real world referent, Mr. Reagan. The phoric relation in (180) may be represented as follows:

\[
\begin{array}{c}
\text{Mr. Reagan} \\
\text{(UOD)} \\
\left\{
\text{President, a Westerner, handsome, a Republican,}\right. \\
\text{real world}
\end{array}
\]

This view of the relationship between an anaphor and its referent has several advantages. First, the distinction between exophoric and anaphoric references (which rests on the absence or presence of a linguistic antecedent) is meaningless, and a unified concept of context, defined as the UOD, is possible. Thus, the semantic interpretation of an anaphor can be stated in terms of a single univocal notion, i.e. reference to a discourse entity in the UOD. Note that the relationship between an anaphor and its referent is an indirect one, mediated by the conceptualized discourse entity in the UOD.

Second, since an anaphor does not refer directly to
the linguistic antecedent, all the difficult problems involving identity condition disappear.

Third, and perhaps best, a number of problems facing the syntactic deletion analysis, particularly problems related to inference phenomena and sloppy identity, can be explained in a systematic and principled way (see 3.3). Thus, unlike the syntactic deletion analysis, this view offers a descriptively adequate analysis of anaphora.

As for the UOD, it must be distinguished from general long-term knowledge which has nothing to do with a given discourse. Though it is not yet clear what kinds of information are stored in the UOD in what way, it seems that the UOD at a given point in a discourse may be defined as a repository of the discourse entities introduced up to that point. Let us call these entities explicitly mentioned in the discourse "data-based discourse entities." It seems that the data-based discourse entities are interrelated among themselves in the UOD, depending upon the semantic and pragmatic ties between them. Furthermore, human inference processes bring in other discourse entities in the UOD, which are derived from the data-based entities. For example, the discourse entity REAGAN will create related entities such as MRS. REAGAN, GOVERNOR OF CALIFORNIA, CALIFORNIA, etc. We will call them "inferentially derived discourse entities." Thus, we will divide discourse interpretation rules into two types: data-based interpretation rules and inference-based interpretation rules. These will be discussed separately.
3.2 Data-based Interpretation Rules

Data-based interpretations represent straightforward cases in which the referent is found directly in the context. In other words, the discourse entity in the UOD which an anaphor refers to has its referent available in the surrounding text or in the situational context. We will discuss nominal ellipsis first. Verbal ellipsis will be taken up later.

3.2.1 Nominal Ellipsis

The most important role in resolving nominal ellipses is played by verbal structure and related concepts. The verbal structure consists of a stem, modal operators, and a sentence-final ending. We will start with sentence-final endings. The modal keyss and honorific si, verbal stems, and one syntactic factor, 'main clause', will be discussed later.

3.2.1.1 Sentence-final Endings

Performative analyses of Korean, such as that of H.B. Lee (1970), have shown that sentence-final endings function mainly as markers of the illocutionary force of a sentence. When the illocutionary force of a sentence is performative proper (ordering, promising, proposing, etc.), the subject of the sentence is allowed to refer only to the speech roles. Its referent is unambiguously interpretable because the doer and the recipient of promising, ordering, and proposing are predetermined by the nature of the illocutionary act in question. This fact is mainly responsible for the extensive use...
of subject nominal ellipsis in these performative sentences.

Consider the following promises:

(185) Ø nayil ka-ma.
      tomorrow go promissory
      'Ø will go tomorrow.'

(186) i ton -un Ø kkok kaphu-lita.
      this money TOP by-any-means pay promissory
      'Ø will pay this money back by any means.'
      (Ø: I, *you, *Chelswu,...)

The illocutionary force of promising is expressed by such endings as ma or lita. Since promises can be given only by the speaker, the elliptical subject which is to perform the promised act in (185-186) must be the speaker.

(187) Ø nayil il -ul ha-la/key(na).
      tomorrow work OM do imperative
      'Work tomorrow.'

(188) Ø annyenghi ka-si-psiyo.
      in peace go imperative
      'Goodbye.'
      (Ø: you, *I, *Chelswu,...)

(187-188) have the illocutionary force of ordering, which is expressed by such endings as la, key(na), or psiyo. Since the ordered activity is normally supposed to be performed by the addressee, the elliptical subject in (187-188) refers to the addressee.

(189) Ø ese uysa-lul pwulu-ca/sev.
      hurry doctor OM call propose
      'Let's call the doctor quickly.'

(190) Ø pap -ul meku-psita.
      rice OM eat propose
      'Let's eat rice.'
are all proposals, whose illocutionary force is expressed by such endings as ca, sey, psita. The referent of the elliptical subject in (189-190) normally refers to the set of the speaker and the addressee.

The elliptical subjects in (185-190) can be assigned their referents by the following interpretation rule:

(191) When a sentence has the illocutionary force of a promise (marked e.g. by ma), an order (marked e.g. by la) or a proposal (marked e.g. by ca), its elliptical subject refers, respectively, to the speaker, to the addressee, and both.

However, when the ellipted element is not the subject, but a non-subject term such as an object or a locative, its referent is not limited to the speech roles and become multi-ambiguous. Consider

(192) nayil ø ka-lita.
    tomorrow go promise
    'I will go ø tomorrow.'
    (ø: hakkyo-ev 'to school', Seoul-ev 'to Seoul,)

(193) ese ø sicakha-key.
    hurry begin Imperative
    'Start ø quickly.'
    (ø: il 'work', swukey 'homework',...)

(194) ca, ppalli ø kkuthnay-ca.
    well quickly finish propose
    'Well, let's finish ø quickly.'
    (ø: siksa 'dinner', noli 'play',...)

Sentences (192-194) are respectively a promise, an order, and a proposal, but the zero anaphora in them can mean a number of things. Since the problem we face in (192-194) is identical to that occurring in the case of subject ellipsis in
statements or questions, which are illocutionarily neutral, let us examine the latter first.

Consider the following statements:

(195) nayil-un Ø hakkyo-eyse kongpwuha-ney.
    tomorrow TOP school at study
    'Ø study at school tomorrow.'

(196) Ø piano-lul chi-l cwul a-o.
    piano OM play how-to know
    'Ø know how to play piano.'
    (Ø: I, Chelswu, Yenghi,...*you)

A statement is marked by such endings as (ni)ta, ney, o. Unlike the elliptical subjects in sentences of promising, ordering, and proposing, those of statements (195-196) can refer to a number of things (except for 'you'), even though we might say that the least marked interpretation tends to be 'I'. This multiple ambiguity arises when sentences like (195-196) are used either as discourse-initial or isolated sentences.

If sentences (195-196) occur in some specific discourse contexts, however, the zero anaphora involved have an unambiguous interpretation. Compare (197-198) with (195-196).

(197)a nayil Chelswu-ka mwues-ul ha-o?
    tomorrow SM what OM do
    'What is Chelswu doing tomorrow?'

b nayil-un Ø hakkyo-eyse kongpwuha-ney. (=195)

(198)a Chelswu-ka akki -ul tha-l cwu-l apnikka?
    SM musical instrument play know
    'Does Chelswu know how to play a musical instrument?'

b Ø piano-lul chi-l cwul a-o. (=196)

In (197b-198b) the zero anaphors refer to Chelswu only.

The exactly same phenomenon is observed in questions.

(199) nayil -un Ø hakkyo-eyse kongpwuha-ni?
tomorrow TOP school at study
'Do(es) Ø study at school tomorrow?'

(200) Ø akki -lul tha-l cwu-l a-pnikka?
musical instrument play how-to know
'Do(es) Ø know how to play a musical instrument?'
(Ø: you, Chelswu, Yenghi,...*I)

The illocutionary force of asking is marked by such endings as kka, ni, or (nu)nya, and again the elliptical subjects in (199-200) can refer to a number of things except for the speaker 'I'.

As in statements, the elliptical subjects of these questions have a definite interpretation when they occur in a specific discourse:

(201)a Chelswu-ka pappukwuna.
SM busy
'Chelswu, you are busy.'

b nayil-un Ø hakkyo-eyse kongpwuha-ni? (=199))

(202)a Chelswu-nun yenglhapnita.
TOP smart
'Chelswu is smart.'

b Ø akki-lul tha-l cwu-l a-pnikka? (=200))

The zero anaphors in (201b-202b) all refer only to Chelswu.

The generalization we observe from (195-202) is this:
when a statement or a question is used as a discourse-initial
or isolated sentence, its elliptical subject can refer to a number of things, even though in the least marked interpretation it refers to the speaker (in a statement) or the addressee (in a question). However, when it occurs in a specific discourse context, the elliptical subject normally refers to the entity mentioned in the immediately preceding sentence. How can we explain this general observation? Is there any explanation which can account for this pattern of anaphora resolution? It should be clear that a syntactic explanation is out of the question. What we need is an explanation with a semantic or pragmatic basis.

When we start a conversation, we already have a number of things in our long-term memory, e.g., Father, Chelswu (who is a friend to Yenghi and me), a flower, happiness, the university, and so on. In normal situations, however, none of these objects can be referred by zero, because there is no guarantee that the addressee can identify them successfully. For example, suppose that I meet Yenghi on the street. Can I say any of the utterances (195-196) and (199-200) in this context, hoping that Yenghi would understand that the referent of the zero is Chelswu? Definitely not. In contrast, let us suppose that Chelswu happens to pass us at this moment and so both my attention and Yenghi's are drawn to him. If Chelswu becomes our common focus, reference to him by zero, as in (195-196) and (199-200), will normally be felicitous. This means that the speaker can refer to an object with zero if and only if he is sure that the referent is obvious enough to
be identifiable without being specifically mentioned.

When a specific discourse is underway, zero reference to an entity in the immediate discourse is normally successful, as in (197b-198b) and (201b-202b), because the entity has just been mentioned and the speaker can be reasonably sure that the entity is identifiable to the addressee without further being mentioned.

Thus, regardless of whether a zero anaphor occurs in a specific context or in an isolated sentence (whether it is anaphoric or exophoric), it can be interpreted successfully if and only if its referent is obvious enough to be identifiable to the addressee. The question, then, is: What kinds of things are identifiable without being mentioned?

What we propose in relation to this question is the semantic notion of "foregroundedness," originally discussed by Chafe (1972, 1976) and based on the relationship between discourse and limited human consciousness (or short-term memory). To use a metaphor, this relation is similar to that of an unfolding drama and the stage on which it is played. As the story goes on, each event will bring its own characters on stage. Since the stage has a limited space and the audience cannot concentrate on more than one or two activities displayed at once, the characters which are already on stage and have finished their roles yield the center of the stage to newly incoming characters and move gradually toward the wings of the stage, finally to get off of it.

A similar model can be envisaged for the relationship
between discourse and human consciousness. Our consciousness, like the stage, has a certain limited space. As the discourse goes on, each sentence in the discourse produces its own discourse entities and enters them onto the stage of consciousness. Since our consciousness is limited, it cannot accommodate all of the continuously incoming entities (unless it is provided, like a computer, with a transcript of all the incoming entities properly indexed and arranged in their temporal order) and so old entities yield the center of the stage to newly-incoming entities and retreat into the background of consciousness.

Then, at any point of a discourse, there will be certain discourse entities which are more salient than others; they are so to speak "in sharp focus" at that point. These are the entities in the foreground of the consciousness of the discourse participants at that moment. These discourse entities will be called "foregrounded." Informationally, they are what Gundel (1978) calls "activated givens." Now, we claim that a zero anaphor will normally be interpreted as referring to the most foregrounded discourse entity at the moment when the sentence with the zero anaphor is uttered.

At the beginning of a discourse the most foregrounded elements are of two kinds: one is the entity which happens to be under common focus of the speaker and the addressee at the moment. This may be an object that is present in the immediate field of perception, such as Chelswu in the aforementioned example, or an entity uniquely identifiable from
the context, such as 'Korea' in (121). On the other hand, when there are no specific discourse entities under common focus at the beginning of a discourse, the most foregrounded entities are the participants in the speech act. As Chafe (1972:54) correctly points out, the speech roles are more "foregrounded" than others because "when someone says something that action necessarily puts him on stage." Thus, any reference with zero in this situation is interpreted either as the speaker (in a statement) or the addressee (in a question). This is why the elliptical subject in (195-196) and (199-200) can refer to a number of things, but the least marked interpretation is one or the other of the speech roles.

Within a discourse, what contributes most to foregrounding is the recency of mention, though this is sometimes overridden by other factors. In other words, the entity that has been mentioned most recently, most commonly in the immediately preceding sentence, will be most foregrounded. This explains why the zero anaphor in (197b-198b) and (201b-202b) is interpreted as Chelswu only. 22

Based on this assumption, an interpretive rule for zero anaphora in statements and questions is formulated as follows:

(203) In statements and questions, a zero anaphor refers to the most foregrounded entity at the time of utterance (either 'under common focus' or 'recently mentioned'). If no particular foregrounded entities are available, the zero anaphor refers to the speech roles.

It may be added that there are other sentence-final endings which bear on the interpretation process. For
example, the ending nola can normally co-occur with the speaker subject, whereas tela only with a third person that is most foregrounded.

3.2.1.2 Modal 'keyss' and Honorific 'si'

The modal operator keyss contributes to the resolution of an elliptical subject in statements and questions. Consider the following sentences:

(204) ø nayil Pusan-ey ka-keyss-ta.
   tomorrow to go will
   'ø will go to Pusan tomorrow.'
   (ø: I, *you, *Chelswu,...)

(205) ø Yenghi-lul manna-po-keyss-ni?
   OM meet try will
   'Will ø try to see Yenghi?'
   (ø: you, *I, *Chelswu,...)

When the modal operator keyss co-occurs with the infinitive form of a verbal, it always expresses volition. Since the speaker can assert his own volition, but not that of others, and since he can directly ask about the addressee's volition, but not that of others (including himself), the elliptical subjects in a statement and a question involving keyss are always interpreted as referring to the speaker and the addressee, respectively.

On the other hand, when keyss is used with non-infinitive forms of verbals, it carries the meaning of 'guess'. And the elliptical subject of the sentence in which it appears refers to the entity that is most foregrounded at the time of utterance, regardless of whether the sentence is a statement
or a question. Consider

(206)a Chelswu-ka ettehkey toyessulkka?
   SM how     have-become
   'How is Chelswu doing?'

b Ø cikum-ccum hakkyo-ey ka-{iss -keyss-ta.
   now      by    school to go{past  must
               ko-iss}{-ing
   'Ø must{be at    }school by now.'
      {be going to}

Note that these two usages of keyss are in complementary distribution. This fact leads to the following interpretive rule for subject ellipsis involving the modal operator keyss.

(207) The elliptical subject of a sentence in which keyss is used with the infinitive form of the verbal refers to one or the other of the speech roles; when keyss is used with a non-infinitive form of the verb, the subject refers to the most foregrounded discourse entity.

Various features of grammatical or lexical honorification contribute to the interpretation process, either directly by identifying the intended referent, or indirectly by narrowing down the range of possible referents. For example, the morpheme si conveys the speaker's deference to the grammatical subject, and this knowledge is used in resolving anaphora in the following text.

(208)a Chelswu-nun sensayngnim-i tuleo-si-nun kes-ul
   TOP teacher       SM enter HON     that OM
   poassta. kulayse,
   saw. and so
   'Chelswu saw the teacher coming in. And so'
When (208a) is uttered, there are two discourse entities available for the two zero anaphora of (208b): Chelswu and sensayngnim. What determines the correct interpretation in this case is the occurrence of si in the embedded sentence. The honorific marker si indicates that $\varnothing_2$, not $\varnothing_1$, refers to sensayngnim.

3.2.1.3 Verbal Predicate

Certain semantic properties of verbals contribute to determining the intended referent of an elliptical nominal. Here we will discuss three of them: subjective feeling and non-intentional perception, implicit causality, and selectional restrictions.

Subjective feeling and non-intentional perception

There is a group of verbs which represent one's subjective internal feelings or non-intentional perception. Among the verbs of subjective feeling are -ko siphta 'like to', silhta 'hate', cohta 'like', mwusepta 'fear', sulphuta 'sad', salangsulepta 'lovely', tepta 'hot', chwupta 'cold', and cilwuhata 'boring'. Among the verbs of non-intentional perception are poita 'see', tullita 'hear', and masissta 'tasty'.

When one of these verbs occurs in a sentence, the elliptical subject (or sometimes topic) of the sentence refers
either to the speaker (in a statement) or the addressee (in a question) (H.B.Lee 1970), unless the speaker makes an authoritative conjecture of others' feelings (I.S.Yang 1972:163).

(209)a Ø ku yenghwa-lul po-ko siphta.
   'I like to see the movies.' (Ø: I, *Chelswu)

b Ø ku yenghwa-lul po-ko siphunya?
   'Do I like to see the movies?' (Ø: you, *Chelswu)

(210)a Ø cengmallo sulphuta.
   'I really sad.' (Ø: I, *Chelswu)

b Ø cengmallo sulphunya?
   'I really sad?' (Ø: you, *Chelswu)

(211)a Ø san -i pointa.
   'I see the mountain.' (Ø: I, *Chelswu)

b Ø san-i poini?
   'I see the mountain?' (Ø: you, *Chelswu)

This is natural because we cannot make accurate judgments about others' feelings but only about our own. Similarly, the speaker can ask about the internal feelings of the addressee, but not about his own or those of a third person.

However, when these verbs co-occur with the pro-verb hata 'do' or with other epistemic modal auxiliary verbs such as -nka pota 'may', -n kes-i thullim epsta 'must', -n kes kathta 'seem', -ul kes-ita 'probable', etc., the reference of the elliptical subject is not restricted to the speech roles.

(212) Ø cengmallo sulphe-ha-ko issta.
   'I really sad.' (Ø: Chelswu, Yenghi, ... *I)
The pro-verb *hata* seems to imply that the non-observable internal feelings of a person are revealed in his outward behavior. Thus, his internal feelings can be read from his observable behavior and this is why the elliptical subjects of (212-213b) can refer to a third person *Chelswu*. Compare (212-213b) with (210a), in which the elliptical subject cannot refer to other than the speaker 'I'.

Epistemic modal auxiliaries put a sentence into an opaque context, i.e., into the speaker's belief world. This enables the speaker to guess the internal feelings of another person.

Even though (214b) is a statement, the elliptical subject can refer to a third person *Chelswu*. This is again a different behavior from what we observed in (209a).

Thus, we propose the following interpretation rule:

(215) The missing subject (or topic) of a sentence in which the verb represents internal feeling or non-intentional perception refers to the speaker (in a
statement) or to the addressee (in a question); if the verb co-occurs with the pro-verb hata or an epistemic modal, the elliptical subject refers to the most foregrounded entity at the time of utterance.

Implicit causality

Another semantic property of the verbals which contributes to resolving the anaphora of an elliptical subject is what Garvey and Caramazza (1976) have called "implicit causality." Consider the following sentences.

(216)a Chelswu-nun Yenghi-lul namwulassta.
       TOP OM scolded 'Chelswu scolded Yenghi.'

    b ø tongsayng-ul ttaylyess-ki ttaymwun-iessta.
       brother OM hit because was 'Because ø hit the brother.' (Ø: Yenghi)

(217)a ø sihem-ey mwusahi hapkyekhayessta-nun sosik-ul
       exam at easily passed that news OM tut-ko,
       heard and 'After hearing that ø passed the exam easily,'

    b Chelswu-nun Yenghi-lul chingchanhayssta.
       TOP OM praised 'Chelswu praised Yenghi.' (Ø: Yenghi)

(218)a ø swuyengcang-ey ka-keyss-ta-ko ha-ca,
       swimming pool go will said when 'When ø said that ø would go to the swimming pool,'

    b Chelswu-nun Yenghi-eykey cuksi helakhayessta.
       TOP to immediately allowed 'Chelswu said 'yes' immediately.' (Ø: Yenghi)

(219)a Chelswu-nun Yenghi-eykey myenglyenghayessta.
       TOP to ordered 'Chelswu gave an order to Yenghi.'

    b ø myech-il swi-si-psiyo.
       a couple of days rest HON 'Have a rest for a couple of days.' (Ø: Yenghi)
The elliptical subjects in (216-219) all refer to Yenghi, which functions as the object (direct or indirect) in the surrounding antecedent sentences, but not to the subject Chelswu.

Compare (216-219) with another set of verbs in (220-223):

(220)a  Ø yangsim-ul soki-l swu epse, conscience betray cannot
'Since Ø cannot betray his conscience,'

b Chelswu-nun Yenghi-eykey capaykhayssta. TOP to confessed
'Chelswu confessed to Yenghi.' (Ø: Chelswu)

(221)a ilpon-un mikwuk-eykey hangpokhayessta. Japan TOP U.S.A. to surrendered
'Japan surrendered to America.'

b Ø cenuy-ul sangsilhayess-ki ttaymwun-iessta. will-to-fight lost because was
'Because Ø lost the will to fight.' (Ø: ilpon 'Japan')

(222)a tasi-nun Ø nuc-ci anh-keyss-ta-ko, again late not will Quotative
'That Ø will not be late again'

b Chelswu-nun Yenghi-eykey mayngsehayessta. TOP to vowed
'Chelswu vowed to Yenghi.' (Ø: Chelswu)

(223)a Chelswu-nun Yenghi-eykey yaksohayessta. TOP to promised
'Chelswu made a promise to Yenghi.'

b i ton -un Ø kkok kap-ciyo. this money surely pay-back will
'Ø will pay back this money by any means.' (Ø: Chelswu)

The elliptical subjects in (220-223) refer to Chelswu or ilpon 'Japan', which functions as the subject in a neighboring sentence, but not to the object Yenghi.
It seems that verbs differ from each other in their propensity to attribute the cause of the action they denote to some factor. Some verbs tend to attribute it to the instigator of the action, others to the patient of, i.e., the entity affected by, the action. We will call this phenomenon the "implicit causality bias" of the verb. Verbs can be classified into three groups according to their causality bias: patient-biased, actor-biased, and neutral. 24

The verbs of the first group are patient-biased. These verbs attribute the cause of the action to the patient, which is normally realized as the grammatical object. For example, if Chelswu is engaged in the action of scolding or praising Yenghi, he normally does so because Yenghi, who is the patient of the action, has done something good or bad, but not because of Chelswu's own behavior. Therefore, the cause of the action denoted by these verbs is normally considered to be the patient. Thus, the elliptical subject of the sentence which either provides the reason for the action or specifies the content of the action (as in (219b)), is usually considered to be the patient (the grammatical object) of the action in question. This explains why all the elliptical subjects in (216-219) refer to Yenghi. Other verbs which are biased toward the grammatical object in terms of implicit causality feature are: conkyenghata 'respect', samohata 'love', kosohata 'accuse', pinanhata 'blame', kwenhata 'recommend', chwukhahata 'congratulate', etc.

The verbs in (220-223) are actor-based. The elliptical
subjects in (220-223) normally refer to the grammatical subject rather than to the object of the neighboring sentence. This behavior is again related to the causality bias of these verbs, which is toward the instigator of the action. For example, if one confesses something to another, one usually does so because one has done something wrong rather than the other person confessed to, i.e. the patient in the act of confessing. Therefore, the elliptical subject of the sentence which either provides the reason of the action (as in (220-221)) or specifies the content of the action (as in (222-223)) is usually associated with the actor, usually the grammatical subject. This is why the elliptical subjects in (220-223) refer to Chelswu. The verbs which belong to this group include: kopenkhata 'confide', pilta 'beg one's pardon', sakwahata 'apologize', calanghata 'boast', cainhata 'admit', and so on.

The causer of the action is also attributed to the grammatical subject in the following examples:

(224)a Yenghi-nun Chelswu-ka mwusewessta.
    TOP SM afraid-of
    'Yenghi was afraid of Chelswu.'

    b Ø maywu kechil-ki ttaymwun-iessta.
        very rough because was
        'Because Ø was very rough.' (Ø: Chelswu)

(225)a Swunhi-nun Yenghi-ka pwulewessta.
    TOP SM envious
    'Swunhi was envious of Yenghi.'

    b Ø maywu yeyppu-ki ttaymwun-iessta.
        very pretty because was
        'Because Ø was very pretty.' (Ø: Yenghi)
The zero anaphor of (224b) refers to Chelswu, which is the subject of (224a). Part of what we know about fearing includes the knowledge that the cause of fearing exists outside the fearer Yenghi in (224a). Since (224b) provides the reason for the fearing, the zero of (224b) is naturally understood as the causer of the fearing; so, Chelswu is associated with the elliptical subject of (224b).

Not all verbs can be classified either as subject-biased or object-biased in attributing causality. There are a number of verbs which are more or less neutral in this regard. The underlying cause of the actions these verbs denote can be either the subject or the object. Consider:

(226)a  Ø palam-ul phiwess-ki ttaymwun-ey,  
fooling-around because-of  
'Because Ø did fooling around,'

b Chelswu-nun anay -wa ihonhayessta.  
TOP wife with divorced  
'Chelswu divorced his wife.'  
(Ø: Chelswu or anay 'wife')

The reference of the elliptical subject of (226a) is ambiguous between Chelswu and Chelswu's wife. It will remain ambiguous unless some other semantic or pragmatic cues are found. This sort of ambiguity seems common among reciprocal verbs, in which the direction of the action involved is not unilateral.

Based on the above discussion, the following interpretive rule is proposed:

(227) If a verb is biased toward its patient or agent in attributing implicit causality, the missing
subject of the surrounding text which provides the reason for the action denoted by the verb refers to the object or subject, respectively. If the verb is neutral in causality bias, the missing subject remains ambiguous in reference.

**Selectional restrictions**

By imposing certain semantic restrictions on the nominal constituents of a sentence, the selectional restrictions of the verbals play an important role in resolving the anaphora of elliptical nominals. Consider

(228)a Yenghi-ka yak -ul mekessni?
   'Did Yenghi take medicine?'

   b ney. Œ₁ Œ₂ mekesseyo.
   'Yes, she did.'

(Œ₁: Yenghi-ka, Œ₂: yak-ul 'medicine')

A crucial cue in determining that the first and second NP's refer to Yenghi and medicine respectively, seems to be the selectional restriction requiring that the subject of eating be animate and that the object be edible.

Selectional restrictions are actually nothing but more than a set of easily definable frozen constraints abstracted from our more general pragmatic knowledge. Selectional restrictions and pragmatic knowledge may be considered to be on a continuum. Then, it is hardly surprising that our general pragmatic knowledge is a useful aid in interpretation, as demonstrated recently in artificial intelligence research (Winograd 1972; Wilks 1975) and by Li and Thompson (1979). We will not go into a detailed discussion as to how
pragmatic knowledge contributes to anaphora resolution, but there are a number of cases which cannot be explained by any means other than a pragmatic account. The following sentences are typical.

(229) \( \emptyset \) ikye.<br>'Victory.' \( (\emptyset: \text{wuli phyen} 'our team') \)

(230) \( \emptyset \) chye pwusyela.<br>hit destroy<br>'To hell with them.'<br>(\( \emptyset: \text{cek phyen 'the other team'})\)

3.2.1.4 Main Clause

The last factor we will discuss in nominal ellipsis is syntactic. Consider the following sentences:

(231)a Chelswu-ka Inchen-ulo kass-ki ttaymwun-ey <br>SM to went because<br>Yenghi-nun Taychen-ulo kassta.<br>TOP to went<br>'Since Chelswu went to Inchen, Yenghi went to Taychen.'

b kulentey, \( \emptyset \) kosok besu-lul tha-ko kassta.<br>and express bus OM ride went<br>'And \( \emptyset \) went by an express bus.'<br>(\( \emptyset: \text{Yenghi, *Chelswu})\)

(232)a Chelswu-ka swuhak sihem-ul chilu-l ttay,<br>SM math exam OM take time<br>Yenghi-nun yenge sihem-ul chilessta.<br>TOP English exam OM took<br>'When Chelswu had a Math exam, Yenghi took an English exam.'

b kulentey, \( \emptyset \) cal chi-ci mos hayssta.<br>and well take not<br>'And \( \emptyset \) did not do well.' \( (\emptyset: \text{Yenghi, *Chelswu})\)

Note that the zero anaphors of (231b-232b) refer to Yenghi.
not Chelswu. Chelswu and Yenghi do the same thing in each (a) sentence--going to somewhere in (231a) and taking exams in (232a). Thus, semantically, they stand an equal chance of being associated with the zero anaphors in the (b) sentences. However, the fact that only Yenghi is the preferred referent seems to be due to the structural fact that Chelswu and his actions are described in subordinate clauses, while Yenghi and her actions are described in main clauses. We argue that constituents in main clauses are more easily accessible to anaphora resolution than those in embedded (or subordinate) clauses.

It may be argued that resolution of the anaphora in (231-232) in favor of Yenghi may be due to the topic collaboration tendency discussed in Chap 3: 3.2 and in Keenan and Schiefflin (1976), where it is argued that the topic of the previous sentence tends to be chosen as the topic of the next sentence in dyadic situations. However, more than topic collaboration is involved. Consider the following sentences containing a relative clause:

(233)a (swul-ey chwihan)ku namca-nun kyothon-sako-lul liquor drunk the man TOP motor accident OM

nayssta.
made .
'The man who was drunken made a traffic accident.'

b ð ttaymwun-ey ku-nun kwusoktoyessta.
because-of he TOP was arrested
'Because of that, he was arrested.'
(ð: kyothon-sako-lul naym kes 'that he made a traffic accident')
As we will see in 3.3.1, zero anaphora can refer to propositions. The anaphors in (233b-234b) refer to the nominals in the main clause rather than to those in the embedded relative clauses. Note that the activities in both the main and subordinate clauses are equally eligible antecedents semantically. Also observe that topichood has nothing to do with the choice of the main clause activity over the activity in the subordinate clause. Similarly consider:

(235)a (defense-lul machin)Chelswu-nun job-ul etessta. defense OM completed TOP job OM got 'Chelswu who finished his defense got a job.'

b Ø acwu elywun il-intey, yonghakwuman. very difficult thing is lucky 'Ø is a very difficult thing; he is lucky.'
(Ø: job-ul etnun kes 'to find a job')

(236)a (job-ul et-un)Chelswu-nun defense-lul machyessta. job OM got TOP defense OM completed Chelswu who got a job finished his defense.

b Ø acwu elywun il-intey, yonghakwuman. (=235a) very difficult thing is lucky 'Ø is a very difficult thing; he is lucky.'
(Ø: defense-lul machinun kes 'to finish one's defense')

Sentences (235-236) clearly show that anaphora resolution is,
in part, syntactically conditioned.

Compare (232) with the following (237):

(237)a Chelswu-nun swuhak sihem-ul chilu-ko, 
TOP math exam OM took and 

Yenghi-nun yenge sihem-ul chilessta. 
TOP English exam OM took

'Chelswu had a Math exam and Yenghi had an
English exam.'

b kulentey, Ø cal chi-ci mos hayssta. 
and well take not

'And Ø did not do well.'

(Ø: Chelswu and Yenghi)

Unlike (232b), the anaphor of (237b) refers to both Chelswu
and Yenghi, or at least is ambiguous between the two. This
difference in anaphora resolution between (232b) and (237b)
seems to be related to the fact that the latter is a co-
ordinate structure with two independent main clauses, whereas
the former has only one main clause. Therefore, we conclude
that the main clause in a complex sentence tends to override
the embedded sentence in anaphora resolution, unless there
are other semantic or pragmatic cues in the text which favor
an opposite interpretation.

This conclusion is also supported semantically. The
main and subordinate clauses differ in the degree of assertedness. Subordinate and relative clauses are generally pre-
supposed, whereas the main clause is asserted. It is natural
that the main assertion, which is informationally more sa-
lient than the presupposed subordinate clause, is more easily
accessible to anaphora resolution. In a coordinate structure,
the two clauses are equally asserted, so that anaphora is not likely to be resolved in favor of one clause over the other.

3.2.2 Verbal Ellipsis

One characteristic of verbal ellipsis that is fundamentally different from nominal ellipsis, and thereby demands an explanation, is the fact that verbal ellipsis is very rare, particularly when it has exophoric reference. The rarity of exophoric verbal ellipsis stands in sharp contrast to the productive and often obligatory nature of nominal ellipsis. For example, it is easy to find exophoric nominal ellipses such as (21-27) but it takes time and certain amount of effort to uncover exophoric verbal ellipses such as (39-43).

The rarity of exophoric verbal ellipsis seems to be related to two facts. First, the most significant fact about exophoric verbal ellipsis is that there is no linguistic cue whatever that can be of help in its resolution. This is quite different from exophoric nominal ellipsis, where there are a variety of syntactic and semantic cues in the verb morphology, such as performative-related endings, modal operators, the logical argument structure of the verbal, etc. No such linguistic cues are available for verbal ellipsis, because the verb morphology itself is missing. Thus, the main repository of information utilized in interpreting nominal ellipsis is missing. Therefore, the elliptical verbal
must be interpreted from the non-linguistic situational context by some other human cognitive faculties, the nature of which is not clear yet.

Another reason for the rarity of exophoric verbal ellipsis may be the nature of that which an elliptical verbal denotes. The referent of an instance of exophoric nominal ellipsis is an entity which may exist in the immediate perceptual field or in common focus. But an elliptical verbal denotes a state or an action which is not only non-referential and amorphous but also potentially vague.

Even though there are no hard and fast interpretive rules for exophoric verbal ellipsis, there are some syntagmatic cues. For example, the range of possible interpretations can be narrowed down by case markers co-occurring with a nominal. Consider (39-40), which we repeat here.

(39) ca ili -lo ø.
   well this-way to
   'ø this way, please.'
   (ø: osio 'come', sesio 'stand')

(40) ani ettehkey yeki-1 ø.
    well how here øM
    'How come you are here?' (ø: wassci 'came')

The object particle lul in (40) and the directional particle lo in (39) indicate that the missing verbal must be a motion verb such as kata 'go' or ota 'come'.

Sometimes elliptical verbals get their interpretations from the conventions of language use, as in idiomatic expressions like (44-45). Elliptical nominals in such sentences
as (30-32) are also interpreted by such conventions.

Anaphoric verbal ellipsis is not as rare as its exophoric counterpart, because its meaning can be interpreted from the linguistic context. Consider (238):

(238a) emeni -nun eti-ey kasyessni?
   mother TOP where went?
   Where did your mother go?'

b hakkyo-ey yo.
   school to
   'To school.'

(238b) is catalyzed by the generation process to (239a):


Next, the node V must be interpreted. The interpretive rule we propose for anaphoric verbal ellipsis is "Verb Copying", which copies the verbal of the immediately preceding sentence into the empty V node. The rule of Verb Copying will transform (239a) to (239b), the subject NP being interpreted as emeni 'mother':

(239b) emeni-nun hakkyo-ey kasyessseo.
   'Mother went to school.'

which is the correct semantic structure of (238b).

3.3 Inference-based Interpretation Rules

Discourse entities which are directly mentioned in the discourse were called "data-based" discourse entities. The data-based entities are often supplemented or enriched by our extra-linguistic knowledge about the world. These
entities are also linked together among themselves by various semantic or pragmatic relationships they participate in. This way, the data-based entities create related but new discourse entities in our mind. We will call them "inferentially derived discourse entities." These derived entities do not appear in the situational or textual context, but it is reasonable to assume that they are registered as fully legitimate discourse entities in the UOD.

The operations which create derived discourse entities from data-based ones are mostly reasoning or inference processes characteristic of human communication. It seems that they are governed basically by inference rules similar to those developed in propositional logic. The extent to which the logician's methods and apparatus in formal languages are useful to the study of natural languages has been a matter of controversy, but as Allwood et al. (1977:104-106) have suggested, some of the concepts and theories of logic "might be a useful tool for analyzing natural languages, [even though] very little has been done in this field." 26 What we will do in this sub-section is to argue that some of the basic deduction rules of propositional logic are actually utilized for resolving anaphora in natural languages. Specifically, we will discuss such inferential processes as iota operator introduction, conjunction, association, and abstraction.

A comment on notational conventions is in order first. There are many cases of inferentially derived discourse entities whose semantics is well-understood but whose exact
syntactic shapes are not clear and cannot be pinpointed. To overcome this problem, we will use several notations used in propositional logic, even though some of the inference rules can be represented equally well with the usual grammatical terms.

3.3.1 Definitization

The simplest and most basic inference process is definitization, by which an indefinite referent introduced into the discourse is registered as a definite entity in the UOD. Consider the following text:

(240a) Chelswu-nun Yenghi-lul salanghanta.
    TOP love
    'Chelswu loves Yenghi.'

b Ø yeyppu-ki ttaymwun-ita.
    pretty because be
    'Because Ø is pretty.' (Ø: Yenghi)

What the zero anaphor of (240b) refers to is a unique and definite discourse entity Yenghi introduced into the UOD by (240a). But how about the following (241-242)?

(241a) Chelswu-nun etten Seoul yeca-lul salanghanta.
    TOP certain girl OM love
    'Chelswu is in love with a Seoul girl.'

b kulena, ku-uy pwumo-nun acwu Ø silhehanta.
    but he of parents very dislike
    'But his parents do not like Ø at all.'
    (Ø: Chelswu-ka salanghanun ku Seoul yeca 'the Seoul girl whom Chelswu loves')

(242a) Yenghi-nun cip -ul han chay sass-nuntey,
    TOP house OM one unit bought and
    'Yenghi bought a house, and'
In the syntactic deletion analysis, we have to posit either (243) or (244) for the underlying structure of (241b), unless arbitrary coreference marking is used (see Chap 1: 3.1.4.1).

However, neither (243) nor (244) can be considered the correct underlying structure of (241b), because neither of them is identical to (241b) in meaning. We should say that what has been deleted in (241b) is the definite NP given in parenthesis in the gloss. Thus, the antecedent in (241a) is an indefinite NP, but the deleted material in (241b) is a definite NP.

In order to accommodate this kind of problem within the syntactic deletion analysis, Chomsky (1965:234) suggested that the definiteness of a noun is "left unspecified" in the underlying phrase marker and is transformationally introduced later, and that the deletion rule disregards the [± definite] feature for its operation. This proposal would make the deletion analysis at least descriptively adequate but it does
not seem to be acceptable.

First, definiteness cannot be considered a transformationally introduced feature. For instance, for the following sentences

(245) The boy came.
(246) A boy came.

Chomsky would presumably give a single deep structure (247):

(247) [ [boy]NP[[came]V ]VP _S

and [+definite] and [-definite] features would later be assigned to (247) to produce the surface forms (245-246). However, (245) and (246) differ from each other in cognitive meaning: e.g., the boy implies a boy, but not vice versa. So, either the transformational treatment is wrong or some surface semantic interpretation rule must be sought. The latter option is clearly undesirable.

Second, if the proposal really is correct, we should expect to find cases in which the antecedent is definite but the deleted NP is indefinite. But, contrary to our expectation, no such cases are found and this fact makes the adequacy of the proposal doubtful.

Third, Chomsky's proposal is hardly explanatory. It still remains a question why the deletion rule disregards the indefinite (the antecedent) and definite (the victim) distinction. Actually, Chomsky's claim is tantamount to
saying that deletion has nothing to do with definiteness, which is incorrect.

The disparity between the antecedent and the victim in terms of definiteness, together with the second and the third problems mentioned above, are automatically explained in the proposed analysis. What we need is a process called "Definitization" that is identical to the iota operator introduction in logic.

A referent may be definite or indefinite when it is first mentioned, but when the referent is registered as a discourse entity in the UOD it always becomes definite, at least to the discourse participants. Definiteness means that there is one and only one individual which is uniquely identifiable and this is true of all discourse entities.

For example, when (24la) has been uttered, several new discourse entities are registered in the UOD. Some of them are already definite: Chelswu and Seoul. Others such as etten Seoul yeça 'a certain Seoul girl' is not. But when this indefinite referent is registered in the UOD as a discourse entity, it is represented as a definite entity: 'the Seoul girl whom Chelswu loves and who was explicitly mentioned in sentence (241a)'. Therefore, this entity is uniquely identifiable. If we call this particular discourse entity $DE_{23}$, $DE_{23}$ might be represented as follows:

$$ (248) \ DE_{23}: \ \lambda x[(x: \text{a Seoul girl}) \cdot \text{LOVE} (\text{Chelswu, } x) \cdot \text{EVOKE } ((241a), x)] $$
which is read as 'the Seoul girl whom Chelswu loves and who was evoked by sentence (241a)'. Thus, what the zero anaphor of (241b) refers to is not the linguistic antecedent etten Seoul yeca, but the definitized discourse entity DE$_{23}$. This way, the disparity between an indefinite antecedent and its definite deleted counterpart is explained as a natural consequence of Definitization.

Definitization may be represented schematically as follows:

\[(249) \text{Definitization} \]
\[\exists x \ (P, x) \supset \exists^1 x \ (P, x)\]

which is read: when a referent $x$ which has a property $P$ exists, there exists in the UOD a definite discourse entity $x$ which has the property $P$. If we apply (249) to (241a), we get (248). Then, this newly derived iota-prefixed entity is what the ensuing zero pro-form refers to.

Definitization is not limited to objects such as 'a girl' and 'a house' in (241-242), but can also apply to propositions. For example, when (241a) has been uttered, the whole proposition can be definitized as an abstract concept: Chelswu-ka etten Seoul yeca-lul salanghanta-nun kes/sasil 'the fact that Chelswu loves a Seoul girl'. This propositional definitization may be represented schematically as follows:

\[(250) f (x) \supset \exists^1 y \ (f(x), y)\]
which is read: if there is a proposition \( f(x) \), there is also a definite discourse entity \( y \) which represents \( f(x) \). The right-side representation is syntactically manifested by -\textit{nun kes}, -\textit{um}, -\textit{ki} (all translatable to 'that'), -\textit{nun sasil} 'the fact that' or independent lexicalizations such as \textit{phan-may} 'sale' (for \textit{pha-nun kes} 'sell-ing'), \textit{samang} 'death' (for \textit{cwuk-nun kes} 'die-ing'), \textit{pihaying} 'flight' (for \textit{na-nun kes} 'fly-ing'), etc.

Propositional definitization is very productive, especially in conditional sentences. Consider

(251)a Chelswu-ka cwukesseyo.
SM died
"Chelswu died."

b \( \emptyset \) sulphun il-ipnita.
sad thing is
"\( \emptyset \) is a sad incident.\) (\( \emptyset : \text{Chelswu-ka cwuk-un kes} \) 'that Chelswu died', \( \text{Chelswu-uy samang/cwuk-um} \) 'Chelswu's death')

(252) cey-ka cenhwaha-myen, \( \emptyset \) an toynayo?
I SM telephone if not good
"If I call you up, isn't \( \emptyset \) O.K.?"
(\( \emptyset : \text{nay-ka cenhwaha-nun kes} \) 'that I call', \( \text{na-uy cenhwa} \) 'my call')

The referent of the zero anaphor of (251b) may be represented as (253):

(253) \( \forall x \) (DIED(Chelswu), \( x \))

which can be read in several ways as indicated in parenthesis in the gloss. Somewhat more complicated cases of propositional definitization will be discussed in 3.3.4.2.
3.3.2 **Conjunction**

Conjunction is another productive inference process. This is a set-formation process. Consider the following text.

**(254)**

a Speaker A: pwunmyenghi yeca-yessupnita. 
\textit{certainly woman was} 'It was surely a woman.'

b pays-sok-uy ai-nun namca-inци 
\textit{belly inside baby boy be whether} mollato 
\textit{do not know} 'Though the baby inside her may be a boy.'

c Speaker B: kulem hol-mom-i anikwunyo. 
\textit{then single not} 'Then, not a single person.'

d Speaker A: ø kathi cwuk-un-ke-cyo. 
\textit{together died so-to-speak} 'ø died together, so to speak.' 
(ø: \textit{ku yeca-wa pays-sok-uy ai} 'the woman and the baby in her')

**(255)**

a seymeyntu han photay-ey molay twu thong-ul 
\textit{cement one bag in sand two bucket} OM neh-ko, 
\textit{put and} 'Put two buckets of sand into a bag of cement.'

b ø₁ mwul-ul pwu-e 
\textit{water OM pour and} 'Pour water.' 
(ø₁: \textit{seymeyntu han photay-wa molay two thong} \textit{a bag of cement and two buckets of sand'})

c ø₂ cal sekusiyo. 
\textit{well mix} 'Mix well.' (ø: \textit{seymeyntu han photay-wa molay twu thong-kwa mwul} \textit{a bag of cement, two buckets of sand and water'})
What the zero anaphor of (255c) refers to is a set consisting of 'a bag of cement, two buckets of sand, and water'. This is the so-called "split antecedents" problem (Grinder and Postal 1971:279). As Dougherty (1969) has correctly pointed out, split antecedents are a problem for the syntactic deletion analysis, because no linguistic antecedent of the structural description \([\text{NP}] [\text{NP}] \text{and} [\text{NP}]\) is available in the surrounding text.

Conjunction is a set-formation process to account for reference of this kind. When (254a) and (254b) are uttered, they separately enter two discourse entities, \textit{ku yeca} 'the woman' and \textit{pays-sok-uy ai} 'the baby in her', into the UOD. Then, these two discourse entities are interrelated to form a combined set of the two, \textit{ku yeca-wa pays-sok-uy ai} 'the woman and the baby in her', which constitutes a new discourse entity. This is exactly what is referred to by the zero of (254d). The conjunction process may be represented schematically as follows:

\[\begin{align*}
\text{(256) a } & \text{ Chelswu-nun etten Seoul yeca-wa ihonhayess-ko} \\
& \text{ TOP certain girl with divorced and} \\
& \text{'Chelswu divorced a Seoul girl, and'} \\
\text{ b } & \text{ Yengho-nun etten Seoul yeca-wa kyelhonhayessta.} \\
& \text{ TOP certain girl with married} \\
& \text{'Yengho married a Seoul girl.'} \\
\text{ c } & \text{ kulentey, } \emptyset \text{ motwu mayak-cwungtok-ca yessta.} \\
& \text{ and all narcotic addict was} \\
& \text{'And } \emptyset \text{ were all narcotic addicts.'} \\
& \text{(} \emptyset : \text{Chelswu-ka ihonhan ku Seoul yeca-wa Yengho-ka kyelhonhan ku Seoul yeca} \text{ 'the Seoul girl Chelswu divorced and the Seoul girl Yengho married'})
\end{align*}\]
\[(257) \exists(x) \ldots \exists(y) \supset \forall z \{z \mid x \ldots y\}\]

where \(\{z \mid x \ldots y\}\) stands for the set of \(z\) which includes \(x \ldots y\) as its members. (257) may be read as "if there exists entities \(x, y, \ldots\), then a new discourse entity which is the set of \(x, y, \ldots\) also exists in the UOD." When (257) is applied to (254), we have the following representation:

\[(258) \exists x (x: \text{woman}) \cdot \exists y (y: \text{the baby in the belly}) \supset \forall z \{z \mid x (x: \text{woman}) \cdot y (y: \text{the baby in the belly})\}\]

and \(\forall z\) is the referent of the zero anaphor of (254d).

Example (256) seems to indicate that Conjunction must apply after Definitization. This is so because the two discourse entities combined in (256c) are already definitized.

### 3.3.3 Association

Discourse does not start in a vacuum. It starts in a specific context with specific discourse participants who already share an extensive amount of knowledge about the world. Thus, when a discourse entity is entered in the UOD, it is naturally associated with some portions of our prior world knowledge which are somehow relevant to the discourse entity. This association will add more extra characteristics to the discourse entity, thus deriving a related but new discourse entity. So, a new, derived entity, which possesses more than the properties explicitly mentioned in the discourse may often be the referent of a zero anaphor.
When (71) is uttered, a discourse entity sen will be listed in the UOD. Our world knowledge about this particular concept includes a lot of related information—-that there are a man and a woman at a sen, that the parents of both parties usually accompany them, that the meeting place is usually a restaurant, etc. Thus, the entity is associated with another related entity sen-ul pon yeca 'the girl at the sen', and this is referred to by zero in (71b).

The association process exhibits two patterns of logical deduction: modus ponens and hypothetical syllogism (or transitivity relation). Note that all of the examples (67-69) and (71-74) given in Chapter I follow the deduction schema of modus ponens:

\[(259)a \quad p \quad b \quad \quad p \supset q \quad c \quad q\]

That is, given the qualification p, and the premise \(p \supset q\), then the conclusion q is derived. In actual discourse, the qualification p is provided by the discourse, and the premise \(p \supset q\) is supplied by pragmatic knowledge. Then, going through the
reasoning process of *modus ponens*, a new discourse entity q is registered in the UOD. For example, consider (67) given in Chap 1: 3.1.3 and the following (260).

(260)a Yenghi-nun *haysan* -ul hayessta.  
TOP confinement OM did  
'Yenghi had her confinement.'

b *ilum-un* ø Chelswu-lako cie-cwuessta.  
name TOP Quot make-gave  
'As for the name, she called ø Chelswu.'  
(ø: *saylo thavena-n aki* 'the newly born baby')

When (67) and (260a) are uttered, they enter two discourse entities *cipwung* 'roof' and *haysan* 'confinement' in the UOD. At the same time, our world knowledge contains the following axioms:

\[(261) \forall x (\exists x (x:\text{roof})) \supset \exists y (y:\text{house})\]
\[(262) \forall x (\exists x (x:\text{confinement})) \supset \exists y (y:\text{baby})\]

These can be read: for all x, if there exists an x, which is a roof (in (261)) or a confinement (in (262)), there also exists y, which is a house (in (261)) or a baby (in (262)). Thus, (261-262) provide the premises. Since the qualifications 'roof' and 'confinement' have already been supplied by the discourse, we can derive new discourse entities 'house' and 'baby', and these derived entities are registered in the UOD, later to be referred to by zero, as in (67b) and (260b).

Though not as common as *modus ponens*, the deduction schema of *hypothetical syllogism* (or transitivity relation) is observed in the association process. The inference
pattern is represented schematically as follows:

\[(263)a \quad p \supset q \quad b \quad q \supset r \quad c \quad p \supset r\]

That is, given \(p \supset q\) as the premise and \(q \supset r\) as the qualification, then a new implication \(p \supset r\) is established. In actual discourse, \(p\) is supplied by the discourse, and the two implications \(p \supset q\) and \(q \supset r\) are supplied by our pragmatic knowledge. Thus, the deduction pattern (263) derives a new discourse entity \(r\).

Consider the following examples, one of which is taken from a drama.

\[(264)a \quad \text{onul -un Walker Hill-ey ka-se} \quad \text{today TOP to go and}'
\quad 'I went to Walker Hill today, and'

\(b \quad \text{sey sikan -ina tolyess-nuntey} \quad \text{three hour as-much-as turned but}'
\quad 'played as long as three hours, but'

\(c \quad \text{Ø ilh-ki-man haysse.} \quad \text{losing only did}'
\quad 'I only lost Ø.' \(Ø: \text{ton 'money.'}\)

\[(265)a \quad \text{ecey Kosam-ey ka-se} \quad \text{yesterday to go and}'
\quad 'I went to Kosam yesterday, and'

\(b \quad \text{Ø yel mali-na capassney.} \quad \text{ten as-many-as caught}'
\quad 'I caught as many as ten Ø.' \(Ø: \text{mwul-koki 'fish'}\)

In (265b), the verb \(\text{capta 'catch'}\) indicates that the elliptical grammatical object must be an animate being, but we do not know what the captured animals are. For the communication
to be successful in this situation, it should be understood among the discourse participants that *Kosam* is a place well-known for fishing. If this hidden pragmatic premise is understood, the referent of the zero anaphor of (265b) is established through the following deduction:

(266) a  Kosam $\supset$ a fishing place  
    b  $\forall x (\exists x (x:\text{fishing place}) ) \supset \exists y (y:\text{fish})$  
    c  Kosam $\supset$ fish

That is, *Kosam*, which is the initial discourse entity introduced by (265a) is associated with a new discourse entity *mwul-koki 'fish'* after going through the chain of reasonings in (266). This new entity is registered in the UOD, and is the referent of the zero in (265b). A similar explanation applies to (264c).

Except for conjunction, *modus ponens*, and the hypothetical syllogism, we have not found any other complex inference rules (e.g., *modus tollens*, *constructive dilemma*, etc.) operating in the interpretation of zero discourse anaphora. This seems to be due to the fact that if more complicated inference rules are allowed to operate, the necessary association between the anaphor and its referent will be hard to achieve. For pronominal anaphora, however, more complicated inference rules seem to operate, because the association with the referent can be facilitated by the pronominal trace.

3.3.4 Abstraction

3.3.4.1 In all of the examples we have examined so far,
the zero anaphor refers to a discourse entity in the UOD, an entity either directly introduced by the text or derived from another discourse entity by inference. However, the following sentences differ in that their zero anaphors do not refer to any discourse entities in the UOD. Consider:

(267)a Chelswu-nun cha₁-lul sassta.
   TOP car₁ OM bought
   'Chelswu bought a car₁.'

b na-to Ø sassta.
   I too bought
   'I bought Ø, too.' (Ø: chaⱽ 'car')

(268)a Chelswu-nun sakwaᵢ-lul mekessta.
   TOP appleᵢ OM ate
   'Chelswu ate an appleᵢ.'

b Yenghi-to Ø mekessta.
   too ate
   'Yenghi ate Ø, too.' (Ø: sakwaⱽ 'apple')

(269)a Chelswu-nun onul mikwuk yecaⱽ-wa kyelhonhanta.
   today American girl with marry
   'Chelswu marries an American girlᵢ today.'

b Yengswu-nun nayil Ø (kyelhon)hanta.
   TOP tomorrow marry do
   'Yengswu will do tomorrow.'
   (Ø: mikwuk yecaⱽ 'an American girl')

When (267a) is uttered, a discourse entity cha is entered into the UOD, but this is not what the elliptical element of (267b) refers to, because the car I bought must not be the same car that Chelswu purchased. The zero anaphor of (267b) refers to another cha 'car', for which no discourse entity exists in the UOD. The same thing is observed in (268-269). This is the phenomenon of so-called "sloppy identity" (Ross 1967) or "pronouns of laziness" (Partee 1975).
To accommodate sentences like (267-269) within the syntactic deletion analysis, it has been proposed (Hankamer 1979:334) that identity between the linguistic antecedent and the deleted material be defined only formally, but not referentially. That is, the deleted entity should be identical to the antecedent only linguistically; coreferentiality between them would not be necessarily required. This rather ad hoc patchwork approach may accommodate data like (267-269), but there is evidence that mere "linguistic or formal identity" as a condition of deletion is simply too strong and that coreference must be included as a condition for the deletion transformation to apply. Consider the following sentences. Two girls are looking at the pictures of two men:

(270)a na-nun i namca₁-ka coha.
‘I like this man₁.’

b *na-nun ø coha.
‘I like this man.’

In (270), the two girls like different men; so, the zero anaphor of (270b) is not coreferential with i namca₁ of (270a). But they are linguistically identical word-for-word. Note that (270b) is not acceptable in the intended sense. However, if the two girls like the same man as in (271-272):

(271)a na-nun i namca₁-ka coha.
‘I like this man₁.’

b na-nun ø silhe.
‘I do not like ø.’

(Ø: i namca₁ ‘this man₁’)

(Ø: i namca₁ ‘this man₁’)
(272)a  na-nun i namca₁-ka coha.
   'I like this man₁.'

   b  na-to Ø coha.
      'I too like
      'I like Ø, too.' (Ø: i namca₁ 'this man₁')

then, (271b-272b) are acceptable. (270-272) clearly show that coreferentiality is a necessary condition for deletion to take place.

Now, let us suppose that there are two people who happen both to be named Yenghi.

(273)a  Yenghi₁-nun Yenghi₁-ka 0-1 kes-ilako malhayssta.
      TOP SM come that said
      'Yenghi₁ said that Yenghi₁ would come.'

   b  *Yenghi₁-nun Ø 0-1 kes-ilako malhayssta.
      TOP come that said
      'Yenghi₁ said Ø would come.' (Ø: Yenghi₁)

(274)a  Chelswu-nun Yenghi₁-lul mannassta.
      TOP OM met
      'Chelswu met Yenghi₁.'

   b  *na-to Ø mannassta.
      'I too met
      'I met Ø, too.' (Ø: Yenghi₁)

In (273-274), the two Yenghi's are not coreferential, even though they are identical in linguistic form. Note that (273b-274b) are both unacceptable. Thus, coreferentiality is again a necessary condition for deletion. If not, we cannot explain sentences such as (270b) and (273b-274b). Hankamer's analysis thus faces a paradox: some anaphors (e.g. (267-269)) do not require coreference as a condition of deletion, and some (e.g. (273-274)) do.

We mentioned that the zero anaphors of (267b-269b) are
referentially distinct from their antecedents in (267a-269a). This need not always be so. Given appropriate contexts, they can be coreferential with the antecedents. For example, compare (268) with the following (275):

(275)a na-nun sakwa₁-lul mektaka, mek-ki silhecyese, I TOP apple⁰ OM eating eat-ing dislike 'I was eating an apple_i, but I didn't like it,'

b Chelswu-ka ø mekessta.
   SM ate 'Chelswu ate ø.' (ø: sakwa₁ 'apple_i')

Note that the zero of (275b) is coreferential with sakwa of (275a). The same phenomenon is observed in (267) and (269). Compare (267) with (276).

(276)a Chelswu-nun phalan cha₁-lul poassta.
   TOP blue car⁰ OM saw 'Chelswu saw a blue car_i.'

b na-to ø (kathi) poassta.
   I too together saw 'I saw ø, too.' (ø: phalan cha₁/j 'blue carᵢ/j')

The zero anaphor of (276b) can be either coreferential or non-coreferential with the possible antecedent phalan cha₁ of (276a). Therefore, regardless of whether deletion operates only under linguistic identity or under linguistic identity plus coreference, we are faced with another problem of explaining why and how the zero anaphors of (276b) can be interpreted as either coreferential or non-coreferential with regard to phalan cha of (276a).

To make matters worse, there is more evidence that even "linguistic identity" itself is not an entirely correct
condition. Consider the following sentences, which involve the so-called pronouns of bound variables (Partee 1972; Sag 1977: Chapter 2).

\[(277)\]

\[a\] Chelswu-nun ku-uy ayin-ul mannassta.
TOP he of lover OM met 'Chelswu met his girl friend.'

\[b\] \{Yengswu\}-to ø mannassta.
I too met 'Yengswu (or I) met ø, too.'

\[(278)\]

\[a\] Chelswu-nun coyonghi ku-uy nwun-ul kamassta.
TOP quietly he of eye OM closed 'Chelswu closed his eyes quietly.'

\[b\] emeni-to ø kamusyessta.
mother too closed 'Mother closed ø, too.'

The elliptical object in (277b) is ambiguous between Chelswu's girl friend and Yengswu's (or my) girl friend. Note that there is no linguistic identity between the antecedent and the victim in the latter reading. In (278), due to the pragmatic constraint that one can close only one's own eyes, the zero of (278b) is interpreted only as emeni-uy nwun 'Mother's eyes', which differs from the antecedent Chelswu-uy nwun 'Chelswu's eyes' linguistically. Therefore, the deletion analysis is faced with several difficult problems in characterizing 'identity' between the linguistic antecedent and the victim (particularly with regard to the relationship between formal and referential identities). Facts like
these led Bolinger (1976:289) to conclude that the key to zero anaphora and pronominalization:

is not to be found in syntax, perhaps even that key does not exist. Keenan ... says of open coreference that it has, generally, no structurally statable restrictions (Emphasis added).

Sentences (267-269) and (276-278) constitute a problem for our analysis, too. We will discuss below how these problems can be solved within our framework.

There are two different kinds of referring. One may be called "referring by identity" (or "identity referring"). This is the traditional notion of referring in philosophy (Strawson 1964; Searle 1969). What is important in this referring is the unique identifiability of one referent that fits the description. The use of a proper name as a referring expression is the most typical means of identity referring. For example, if we say (279),

(279) Chelswu-ka o-nta.
      SM come
'Shelswu is coming.'

the use of Chelswu in this sentence implies that there exists one and only one entity so labelled and that this entity is uniquely identifiable in the context when labelled Chelswu. In this case, (279) might be represented logically as (280):

(280) COME ( \( x \) ( Chelswu (x) )

Proper names such as Fido, Hawaii, Yenghi, etc. are usually used to fulfill this function.
The other kind of referring may be called "referring by property" (or "property referring"). What is emphasized in this kind of referring is a certain property of the referent which the referring expression denotes, but not necessarily the unique identifiability of the referent. Many common nouns are used for this function. For example, the noun cha 'car' in the following sentence

(281) cekes-un cha-ita.
that TOP car is
'That is a car.'

is not used to refer to a unique referent, but to represent the property of 'being a car.' In this case, (281) may be represented logically as (282)

(282) \( \exists x \ (x: \text{that}) \cdot \text{CAR} \ (x) \)

where the noun cha 'car' functions as a predicate. This differs from the iota-prefixed Chelswu of (280).

By this, however, we do not mean that common nouns cannot be used in identity referring. Demonstratives are usually used to make this function of common nouns clearer:

(283) ya! cha-ka kursahata.
car SM beautiful
'Look! The car is beautiful.'

(284) cikum ceki ka-nun ce cha-nun Chelswu-uy kes-ita.
now there going that car TOP of thing is
'That car which is going over there is Chelswu's.'

Here the same noun, cha, is used to pick up or identify a
certain specific entity. The difference between these two kinds of referrings is well exhibited in the following sentence.

(285) Seoul-un hankwuk-uy swuto-ita.
      Top Korea of capital is
      'Seoul is the capital of Korea.'

in which Seoul is used to refer to a unique entity and hankwuk-uy swuto 'the capital of Korea' is used to represent a property of that entity.

When a common noun is used as a property function, as represented by CAR(x) in (282), the property which the noun denotes is an abstraction. There is a set, the members of which possess a number of diverse properties. Of these diverse properties, a certain property that is shared by all members of the set is abstracted to represent the set as a whole. For example, there are a number of women in the United States; they are all different from each other in name, color, height, character, habit, the number of children they have, etc., etc. But they share a common property. That is, they all have American citizenship. Thus, the noun mikwuk yeça 'American woman', as in (269), is a property shared by the set of these women. The same function is served by the nouns cha 'car' and sakwa 'apple', when they are used in the contexts of (267-268).

In property referring, the entity referred to must have the specific property denoted by the expression. The unique identity of the individual members of the set does not
matter. Technically, each member of the set which is defined by the presence of a certain property is called "an alphabetical variant", and all of the alphabetical variants of the set are equal and can be freely substituted (cf. Reichenbach 1947:201-211; Church 1941; Allwood et al. 1977). The distinction between 'identity' and 'property' matches that between de re ('about the thing') and de dicto ('about what is said') interpretations respectively. In the former case, what is important is the thing in the actual world, but the important thing in the de dicto reading is what the description says about the object to which it refers (Allwood et al. 1977:114-116).

Now, let us consider (276a-b). When (276a) is uttered, a discourse entity phalan cha 'a blue car' is registered in the UOD. The speaker of (276b) can interpret this discourse entity in two ways: either as a unique entity (in its identity function) or as a representation of a property possessed by the entity (in its property function). We will adopt the usual practice of logic: the iota operator will be used to represent the former (de re) interpretation and the abstract lambda operator (Church 1941) will be used to represent the second (de dicto) interpretation.

The use of the iota operator is essentially identical to our Definitization rule (249), which will be repeated here:

(249) Definitization
\[ \exists(x) ( P(x) ) \supset \lambda x ( P(x) ) \]
Thus, (276a) will first enter a discourse entity *phalan cha*:

(286) \[ \exists x \text{(BLUE CAR} (x)) \text{)\] 

This will be converted to (287) by Definitization, thus introducing a new discourse entity, say \( \text{DE}_{15} \):

(287) \[ \text{DE}_{15} : \exists x \left[ \text{(BLUE CAR} (x)) \cdot \left( \text{SAW} \left( \text{Chelswu, x} \right) \cdot \left( \text{EVOKED} \left( (276a), x \right) \right) \right] \]

(287) is read as 'the blue car which Chelswu saw and which was mentioned in (276a). We argue that \( \text{DE}_{15} \) is exactly what is referred to by the zero anaphor of (276b) when it is interpreted as being coreferential to *phalan cha* of (276a).

When the same noun, *phalan cha* of (276a), is understood as property referring—an entity that has the property of being a blue car—it will be represented as (288):

(288) \[ \text{SAW} \left( \text{Chelswu, x} \right) \cdot \text{BLUE CAR} (x) \]

and the discourse entity \( x \) will further be represented by the abstract lambda operator as follows:

(289) \[ \lambda(y) \text{(BLUE CAR} (y)) , \text{ (a)\] 

which is read as 'a has the property of being a blue car' or 'a is such that it has the property of being a blue car'. In (289), \( a \) is a free variable and it can be substituted for by other alphabetical variants \( b, c, d \), and so on.

(290) \[ \lambda(y) \text{(BLUE CAR} (y)) , \text{ (b)\]
Let us call $b$ of (290) $DE_{16}$. We can identify $DE_{16}$ as the referent of the zero anaphor of (276b) in its de dicto reading. All of the zero anaphors of (267-269) can be explained in this way. They all refer to individual members of the set possessing the particular property at issue.

Sentences (277-278), which have been called 'pronouns of bound variables', can be explained by the abstraction process. Let us consider (277). When (277a) is uttered, it will be interpreted in two ways by the speaker of (277b).

One interpretation is that ku-uy ayin 'his girl friend' is interpreted as identity referring—as referring to a uniquely identified entity such as Yenghi. In this reading, (277a) will be represented as (291):

\[(291) \text{MEET (Chelswu, his sweetheart)}\]

The pronoun his will first be interpreted as Chelswu by a pronominal anaphoric interpretation rule. (Such a rule is beyond the scope of this study.) Then, a discourse entity, let us say $DE_{17}$, represented as $\exists x (\text{Chelswu's sweetheart}, x)$, will enter the UOD. $DE_{17}$ will be the referent of the zero anaphor of (277b) in its de re reading.

The other interpretation will analyze (277a) into (292):

\[(292) \text{MEET (x, y)· (x: Chelswu)· SWEETHEART (y, x)}\]

What is important in (292) is not the unique entity called 'Chelswu's sweetheart', who happens to be Yenghi, but an entity which has the property of being a sweetheart to
somebody. This entity might be represented as (293):

$$(293) \lambda(y)(x) \text{(SWEETHEART \(y\)(x))}, \ (a)$$

The variable $x$ in (293) is bound to the grammatical subject. (see Partee 1975:265-269, Sag 1977:69, Williams 1977a:115-121 for the binding of a variable to the grammatical subject.) In (293), $a$ is a free variable which can be substituted for by another alphabetical variant $b$:

$$(294) \lambda(y)(x) \text{(SWEETHEART \(y\)(x))}, \ (b)$$

which can be read as 'an entity that has the property of being a sweetheart to somebody'. Thus, (294) is exactly what is referred to by the zero anaphor of (277b) when it has a de dicto reading. The discourse entity (294) may be filled in the ellipted object node of (277b), producing the representation in (295):

$$(295) \text{MEET} (x:Yengswu, \lambda(y)(x) \text{(SWEETHEART \(y\)(x))}, \ (b))$$

This is read as 'Yengswu met an individual who has the property of being a sweetheart of Yengswu' (or simply 'Yengswu's sweetheart'). (295) is the correct semantic representation of (277b) in its de dicto reading.

Sentence (278) can be explained in the same fashion. However, the de re reading Chelswu-uy nwun 'Chelswu's eyes' is ruled out by the pragmatic condition that one cannot close others' eyes. Only the de dicto interpretation—an entity that has the property of being an eye to somebody—is
allowed in (278b).

3.3.4.2 It was mentioned in 3.3.1 that a proposition can, as a whole, be definitized to a discourse entity. However, it is not always a specific proposition but the logical form of a proposition that is definitized. The logical form of a proposition is called "a propositional function." It is constructed by substituting variables for all or some of the individual constants of a proposition. For example, from the propositions in (296)

(296)a John loves Mary.
   b Harry loves Jane.
   c Tom loves Sue.

we can abstract the following logical form or propositional function: \( \text{'}x \text{ love } y\text{' or '}\text{LOVE(x,y)}\text{'}\)

Sometimes the propositional function is first abstracted from a given proposition and then this abstracted propositional function is definitized to a discourse entity, which finally functions as the referent of a zero anaphor. Consider the following 'X-nun Y-ta' pattern sentences we discussed in 2.1.

(297)a Chelswu-nun nayil New York-ey kanta.
       TOP tomorrow to go
       'Chelswu goes to New York tomorrow.'

b na-nun molay-ta.
   I TOP the-day-after-tomorrow be
   'For me, it's the day after tomorrow.'

(298)a Chelswu-nun nayil New York-ey kanta. (= (297a))
We argued that (297b-299b) all have the following underlying structure (300), if (298b) is taken for illustration:

\[(300) \quad \left[ \left[ \text{na-nun} \right]_{\text{NP}} \quad \left[ \Delta \right]_{\text{NP}} \left[ \text{London-ita} \right]_{\text{VP}} \right]_{\text{S}}, \quad \right]_{\text{S}}\]

The problem is then how to interpret the subject NP.

When (297a) is uttered, it has the following proposition (301):

\[(301) \quad \text{GO (Chelswu, New York, tomorrow)}\]

By substituting variables for all or some of these arguments, we can arrive at several propositional functions. We will represent the argument replaced by a variable as a blank, which stands for an empty argument slot without semantic content. Thus, among others, we will have the following propositional functions:

\[(302) a \quad \text{GO (____, New York, ____)}\]
\[b \quad \text{GO (____, ____ , tomorrow)}\]
\[c \quad \text{GO (Chelswu, ____ , ____)}\]

These propositional functions are then definitized to produce such discourse entities as (303a-c):
We argue that these propositional functions (303a-c), which have been abstracted and definitized from the proposition (297a), are what the zero anaphors of (297b-299b) refer to.

A possible problem arises here. The referents of the zero anaphors of (297b-299b) actually mean more than what the logical representations of (303a-c) imply. For example, the zero anaphor of (297b) represents time and those of (298b-299b) refer to places. But the logical expressions (303a-c) do not show this kind of information. We do not yet have a solution for this problem (cf. Vendler 1967: Chapter 2).

3.3.4.3 Abstraction processes also seem to be involved in the anaphoric use of the verbs kata 'go' and ota 'come'.

(304)a  ne encey party-ey ol-lay?
you when party to come will 'When will you come to the party?'

b 10 si -ey ø
  o'clock at 'At ten o'clock.' (ø: kanta 'go')

(305)a nay 10 si -ey (ne-eykey) ka-ma.
I o'clock at you to go will 'I will go to your place at ten.'

b an tway. cikum ø.
  no now 'No. Right now.' (ø: ota 'come')
The antecedents of the elliptical verbs in (304-305) are respectively ota and kata, but the actual referents for the same elliptical verbs are respectively kata and ota.

A similar phenomenon is observed in the anaphoric use of deictic pronouns:

(306)a ikes-i mwues-inya?
this SM what be
'What is this?'

b Ø mannyenphil -ita.
fountain pen be
'Ø is a fountain pen.' (Ø: kukes 'that')

(307)a kukes-un elma -yo?
that TOP how much
'How much is that?'

b Ø 100-wen ipnita.
won be
'Ø is 100 won.' (Ø: ikes 'this')

In (306-307), the two deictic pronouns ikes 'this' and kukes 'that' refer to one and the same thing. However, strictly on the basis of linguistic form, the antecedent and the victim are distinct.

The reverse phenomenon to (306-307) is observed in the following sentences (cf. Partee 1978:75; I.S.Yang 1972:65):

(308)a na-nun tangsin₁-ul salanghapnita.
I TOP you OM love
'I love you.'

b na-to Ø salanghapnita.
I too love
'Me, too.' (Ø: tangsin₁ 'you')

In (308a-b), the same deictic pronoun tangsin 'you' refers to two different entities, but the antecedent and the victim
are formally identical.

Examples like (304-308) are problematic for the syntactic deletion analysis, because the verbs kata/ota and the pronouns ikes/kukes are not formally identical. One might suggest that these formatives are identical at some abstract level, e.g. in deep structure, and that the two expressions are transformationally related. For example, it might be proposed that only one of each pair of the expressions, say kata and ikes, appears in deep structure and that kata and ota differ only on the surface. This can be taken care of by a transformation which changes kata to ota in some pragmatic environments. The identity condition is thus met and the deletion takes place at deep structure.

However, this kind of a proposal could hardly be acceptable because of the drastic change involved in the lexical forms. This kind of practice is prohibited by the constraints of the Extended Standard theory (Chomsky 1972). In addition, there is considerable evidence that deletion must operate on surface forms or at a shallow level (Sag 1977: Chapter 2), even after the post-cyclic Scrambling rule (see 4.1; Bresnan 1976; Williams 1977a). Finally, the environment of the alleged transformation must be specified in pragmatic terms, as we will see. This too is unacceptable in a strictly formal theory. Furthermore, sentences like (306-308) present complications for the coreferentiality and identity conditions of the deletion analysis, as discussed in 3.3.4.1.

Whatever patch-ups are devised within the syntactic deletion
analysis, it is clear that these examples will require significant complication of the theory.

In our analysis, kata and ota are two separate lexical items. However, there is a logical equivalence or equivalence in logical form between the two lexical items. This logical equivalence is derived from the actual uses of these lexical items by the abstraction process. Because of this abstracted logical equivalence, the two expressions are interpreted as being interchangeable. This can explain the anaphoric uses involved in sentences (304-308). We will start with the verbs kata and ota.

The verbs kata and ota denote movement and resulting change of location. The correct use of these verbs in a dyadic situation is determined by the speaker's assumption of the "deictic center" (K.-D. Lee 1978), the goal of the movement involved. In a dyadic situation such as a question-answer pair, ota is used when the deictic center (or destination) of the motion is the speaker and the subject of the motion is either the addressee or a third person who is most foregrounded at the moment. Consider

(309) ne encey hakkyo-ey ol-lay?
   you when school to come will
   'When will you come to school?'

(310) Chelswu-ka hakkyo-ey wassta.
    SM school to came
    'Chelswu came to school.'

(309-310) imply that the speakers of these sentences are already at school, and that either the addressee (as in (309))
or a third person (as in (310)) is going to or has moved toward the speaker. In this specific situation—if the speaker is already at school—the verb kata cannot be used, as the unacceptability of (311-312) shows:

(311) *ne encey hakkyo-ey kal-­lay?
(312) *Chelswu-ka hakkyo-ey kassta.

If we use x to represent the speaker and y the addressee (or a foregrounded third person), then the verb ota might be represented logically as follows:

(313) ota: \( \wedge (x) (y) \text{ (MOVE } (y, x)) \), (a)

On the other hand, the verb kata is used when the deictic center of the movement is toward the addressee and the subject of the motion is either the speaker or a foregrounded third person. Consider (304) and the following sentences:

(314) cikum ney samwusil-lo ka-ma (*o-ma).
now your office to go will come
'I will go to your office now.'

(315) (A husband in his office telephones his wife at home, asking her to send his son Chelswu to him.)

a Chelswu-ka encey cip-­eyse ttenasso?
SM when house from left
'When did Chelswu leave home?'

b 10 si -ey kasseyo (*wasseyo).
o'clock went
'He went at ten.'

In (304) and (314-315), the verb kata implies that either the speaker or a foregrounded third person Chelswu is moving
toward the addressee, but not the other way around. Thus, 
kata might be represented as follows:

\[ (316) \text{kata}: \land (x)(y) (\text{MOVE}(x, y)) \]

The logical expressions (313) for ota and (316) for 
kata show that the two verbs are identical in logical form 
with an identical abstract predicate MOVE. In a question-
answer pair, the roles of speaker and addressee switch, and 
so \( x \) and \( y \) can be considered alphabetical variants. Thus, 
the two verbs have a logical equivalence. In the proposed 
analysis, we assume that the verb ota in (304a) is interpret-
ed as being equivalent to kata by the abstraction process, 
and so the latter verb can be evoked with a zero in (304b).

A similar treatment is possible for the deictic pro-
nouns ikes and kukes. Let us assume an abstract predicate 
PROXIMAL to show the distance between two entities, and three 
variables \( x, y \) and \( z \): \( x \) and \( y \) represent the speaker and the 
addressee respectively and \( z \) stands for an object. Then, 
the deictic pronouns ikes and kukes may be represented as 
follows:

\[ (317) \text{ikes}: \text{PROXIMAL}(z, x) \]
\[ (318) \text{kukes}: \text{PROXIMAL}(z, y) \]

(317-318) show that the two deictic pronouns are identical 
in logical form.

By the same token, tangsin in (308a-b) can be repre-
sented as having logical forms (319-320) respectively:
(319) **tangsin** in (308a): ADDRESSEE(x, y)

(320) **tangsin** in (308b): ADDRESSEE(y, x)

As the identical logical forms in (317-318) and (319-320) indicate, the anaphoric uses in (306-308) are all justified.

4. **Ellipsis in Coordinate Structure**

We have so far discussed discourse ellipses which occur either in a single sentence or across sentence boundaries, and whose anaphoric relations are not structurally but semantically-pragmatically determined. While doing so, we have covered all the types of sentences given by Choi (1975) in Chap 1: 2.1 with the exception of sentence (15), which will be repeated here:

elder-brother to younger-brother to went
'The elder brother utherford to Seoul, and the younger one went to Pusan.' (utherford: kass-ko 'went and')

We have not touched on sentences like (15), because they exhibit some characteristics which are not found in other discourse ellipses. Note that (15) is a coordinate structure with the verb in the left conjunct deleted.

Verbal ellipsis in a coordinate structure like that in (15) will be called "Gapping". 4.1 presents several constraints which apply only to Gapping. 4.2 defines what Gapping is. 4.3 tries to explain why constraints such as those mentioned in 4.1 operate in Gapping. 4.4 discusses the final
rule of "Conjoining" that is unique to ellipsis in coordinate structures.

Unlike other discourse ellipsis we have seen, Gapping is structurally governed in the sense that the gapped element is always the verbal in the left conjunct which is identical to the verbal of the right conjunct of a coordinate structure. Thus, Gapping and Coordinate ellipsis will be discussed within the framework of the deletion analysis, and in this section we will assume an ellipsis rule similar to the one given in Chap 1: 3.1. 30

4.1 Ellipses involved in coordinate structures are exhibited in the following types of sentences:

(321) Yenghi-nun hakkyo-ey ka-ko, Ø tapang-eyto kassta.
    TOP school to go and tea-room also went
    'Yenghi went to school and Ø to the tea-room.'
    (Ø: Yenghi)

(322) Yenghi-nun hakkyo-ey-to ka-ko, Ø tosekwan-eyse
    TOP school to also went and library at
    kongpwu-to hayssta.
    study also did
    'Yenghi went to school and Ø also studied in
     the library.' (Ø: Yenghi)

(323) Chelswu-nun ttek-ul mantul-ko, Yenghi-nun Ø
    TOP cake OM made and TOP
    mekessta.
    ate
    'Chelswu made cakes and Yenghi ate Ø.'
    (Ø: ttek 'cake')

(324) Chelswu-nun Yenghi-eykey Ø₁ Ø₂, Yengswu-nun
    TOP to TOP
    Swunhi-eykey chayk-ul cwuessta.
    to book OM gave
'Che1swu .getDocumentText()  Yenghi and Yengswu gave a book to Swunhi.
(∅₁: chayk 'book', ∅₂: cwi-ko 'give and')

(325) Chelswu-nun Seoul-lo ∅, Yenghi-nun Pusan-ulo TOP to TOP to

kassta.
went

'Che1swu ∅ to Seoul, and Yenghi went to Pusan.'
(∅: ka-ko 'go and')

In (321-322), the subject is ellipted. The verbs in the two conjuncts can be identical, as in (321), or different, as in (322). In (323-324), the object is ellipted; (324) also has the verbal ellipted. In (325), only the verbal predicate is ellipted.

Putting (324) aside for a while, let us compare sentences (321-323) with the gapped sentence (325). In the former, ellipsis takes place in the right conjunct, but in the latter, it occurs in the left. Except for this difference in site, the two types of sentences seem to be similar, but actually they are not. Gapped sentences like (325) are subject to several constraints which do not apply to other coordinate ellipses like (321-323). It is our hypothesis that coordinate ellipses like (321-323) are not basically different from other discourse ellipses we have seen in the preceding sections, but that Gapping in (325) is a special type of discourse ellipsis which is unique to coordinate structure. To justify this hypothesis, we will compare sentences like (321-323) with the gapped sentence like (325) in this subsection, and discuss how Gapping differs from other discourse
ellipses, including those in (321-323).

First, the conjunctors allowed in gapped sentences must be symmetrical ones such as kuliko 'and' (or its reduced form ko) and kulena 'but' (or its reduced form na).\textsuperscript{31} Consider

\begin{enumerate}
\item[(326)]
\begin{align*}
(326)a & \text{ Chelswu-nun Seoul-lo } \emptyset, \text{ kuliko Yenghi-nun} \\
& \quad \text{TOP to and TOP} \\
& \quad \text{Pusan-ulo kassta.} \\
& \quad \text{to went} \\
& \quad \text{Chelswu } \emptyset \text{ to Seoul, and Yenghi went to Pusan.}' \\
& b * \text{Chelswu-nun Seoul-lo } \emptyset, \text{ kulentey/kulayse} \\
& \quad \text{then so Yenghi-nun Pusan-ulo kassta.}
\end{align*}
\end{enumerate}

\begin{enumerate}
\item[(327)]
\begin{align*}
(327)a & \text{ Chelswu-ka pang-ey } \emptyset, \text{ kuliko Yenghi-ka pwuek-ey} \\
& \quad \text{SM room to and SM kitchen} \\
& \quad \text{tule-kassta.} \\
& \quad \text{entered} \\
& \quad \text{Chelswu } \emptyset \text{ into the room, and Yenghi went into the kitchen.}' \\
& b * \text{Chelswu-ka pang-ey } \emptyset, \text{ kuleca/kulayss-ul tta~} \\
& \quad \text{then at-that-time Yenghi-ka pwuek-ey tule-kassta.}
\end{align*}
\end{enumerate}

But there is no such constraint in other discourse ellipses, including (321-323). Compare (326b) and (327b) with (328) and (329) respectively:

\begin{enumerate}
\item[(328)]
\begin{align*}
(328) & \text{ Chelswu-ka ttek-ul mantulessta; kulentey/kulayse} \\
& \quad \text{SM cake OM made then so} \\
& \quad \text{Yenghi-ka } \emptyset \text{ mek-ci anhassta/mekessta.} \\
& \quad \text{SM eat not ate} \\
& \quad \text{Chelswu made cakes, but/so Yenghi did not eat/ate } \emptyset.'
\end{align*}
\end{enumerate}
(329) Chelswu-ka pang-ey tule-kassta; kuleca/kulayss-ul
SM room to entered then at-that-
ttay Yenghi-ka nawassta.
time SM came out
'Chelswu entered the room; then/at that time
Yenghi came out ø.'

Second, gapped sentences, unlike other discourse ellipses, must meet the structural identity condition. That is, the two conjuncts in gapped sentences must be strictly parallel in structure, a parallelism which includes identical node labels for constituents, semantic similarity, and constituents in identical order. Let us take the following sentences for example:

(330) *Chelswu-nun naccam-ul ø, Yenghi-nun pam-ey canta.
TOP nap OM night at sleep
'Chelswu ø a nap, and Yenghi sleeps at night.'
(ø: ca-ko 'sleep and')

The syntactic categories of the underlined constituents are not parallel in (330): one is an NP, the other a PP. Hence, (330) is unacceptable.

(331) *na-nun i coat-ka ø, Chelswu-nun kyewul-i chwupta.
I TOP this coat SM TOP winter SM cold
'This coat is ø to me, and winter is cold to
Chelswu.' (ø: chwup-ko 'cold and')

(332) *Chelswu-nun ø, kuliko i yenghwa-nun sulphuta.
TOP and this movie TOP sad
'Chelswu is ø and this movie is sad.'
(ø: sulphu-ko 'sad and')

In (331-332), the semantic properties of the underlined NP's are not the same. Chelswu is an experiencer and the 'movie'
is a patient in (332).

(333) *10 si-ey Chelswu-nun Yenghi-hako Ø, Swunhi-hako o'clock TOP with with

ohwu-ey Yengswu-nun hakko-ey kassta.
afternoon at TOP school to went

'At 10 Chelswu Ø with Yenghi, with Swunhi in the afternoon Yengswu went to school.'

In (333), the orders of constituents in the two conjuncts are different. Sentences (330-333) show that the two conjuncts in gapped sentences must be identical in syntactic structure and in semantic properties of the constituents. However, no such constraint is operative in other coordinate ellipses. Compare (330-333) with (334):

(334) Chelswu-nun sayngsen-ulo kwuk-ul kkuli-lyeko-ha-,
TOP fish with soup OM make is-going-to

-ko, Yenghi-nun pwul-ey Ø kwu-ulyeko-hanta.
and TOP fire on roast is-going-to

'Chelswu is going to make a soup with fish, and Yenghi is going to roast Ø on fire.'
(Ø: sayngsen-ulo 'fish')

Note that the two conjuncts in (334) differ in structure and in the semantic properties of the conjuncts, but that (334) is still acceptable.

Third, Gapping applies only to main, and not to subordinate, clauses. This constraint does not apply to other discourse ellipses, including the coordinate ellipses involved in (321-323). Consider the following sentences:
Gapping occurs in a subordinate clause in (335) and in a main clause in (336). Only the latter is acceptable. Compare (335) with (337):


'Chelswu forgiving his brother $\emptyset$ his father, and Yengswu forgiving his wife to please his mother.' ($\emptyset$: apeci-evkey 'to father')
gapped sentences cannot exceed three. There is no such constraint in other coordinate ellipses.

(338) Chelswu-nun Yenghi-eykey os-ul Ø, Yengswu-nun TOP to clothes TOP

Swunhi-eykey chayk-ul tollye-cwuessta.

to book OM returned

'Chelswu Ø clothes to Yenghi, and Yengswu returned a book to Swunhi.'
(Ø: tollye-cwu-ko 'return and')

(339) ?Chelswu-nun Yenghi-eykey 10 si-ey chayk-ul Ø,

TOP to o' clock book

Yengswu-nun Swunhi-eykey 1 si-ey kong-ul tollye-

TOP to o' clock ball OM

cwuessta.

returned

'Chelswu Ø a book to Yenghi at 10, and Yengswu returned a ball to Swunhi at 1.'
(Ø: tollye-cwu-ko 'return and')

(340) ?Chelswu-nun Yenghi-wa cip-eyse 10 si-kkaci TOP with home at o' clock till

chayk-ul Ø, Yengswu-nun Swunhi-wa hakkyo-eyse book TOP with school at

caceng-kkaci kulim-ul poassta.
midnight till picture read

'Chelswu Ø a book with Yenghi at home till 10, and Yengswu read a picture-book with Swunhi at school till midnight.' (Ø: po-ko 'read and')

Compare (340) with (341):

(341) Chelswu-nun cip-eyse Yenghi-wa 10 si-kkaci TOP home at with o' clock till

chayk-ul ilkess-ko, Yengswu-nun Ø Swunhi-wa ca-

book OM read and TOP with mid-
ceng-kkaci patwuk-ul twuessta.
night till go-game played
'Chelswu read books with Yenghi at home till 10 and Yengswu played the go-game with Swunhi $\emptyset$ till midnight.' ($\emptyset$: cip-eyse 'at home')

Whatever native speakers' intuitions about the acceptability of (340-341) may be, there is no doubt that (341) is much better than (340).

The fact that these four constraints apply only to gapped sentences such as (325), but not to other coordinate sentences with ellipsis, such as (321-323), leads to the conclusion that Gapping must be distinguished, as a special case, from other cases of ellipsis in coordinate structures. The latter do not differ from other discourse ellipses in general. But, why is Gapping, unlike discourse ellipsis in general, subject to these constraints?

4.2 In this sub-section we will first describe the pattern of ellipses in coordinate structures. This will define Gapping and set the stage for the explanation (in 4.3) of why Gapping is subject to the constraints mentioned above. Consider the following sentences:

(342)a Yenghi-ka hakkyo-eyse chayk-ul ilk-ko, $\emptyset$ kong-SM school at book OM read and park wen-eyse wuntong-ul haysts ta. at exercise OM did

'Yenghi read books at school, and $\emptyset$ took exercise in the park.' ($\emptyset$: Yenghi-ka)

b *$\emptyset$ hakkyo-eyse chayk-ul ilk-ko, Yenghi-ka kongwen-eyse wuntong-ul haysts ta. ($\emptyset$: Yenghi-ka)
(343) a Yenghi-nun hakkyo-ey ka-ko, tto Ø tosekwan-ey
TOP school to go and library to
tosekwan-ey kassta.
went
'Yenghi went to school, and Ø went to the
library.' (Ø: Yenghi-nun)
b *Ø hakkyo-ey ka-ko, tto Yenghi-nun tosekwan-ey
tosekwan-ey kassta. (Ø: Yenghi-nun)

(344) a Yenghi-nun hakkyo-ey Ø₁, kuliko Ø₂ tosekwan-ey
TOP school to ₁ and library to
tosekwan-ey kassta.
went
'Yenghi Ø₁ to school, and Ø₂ went to the li-
brary.' (Ø₁: ka-ko 'go and', Ø₂: Yenghi-nun)
b *Ø₂ hakkyo-ey Ø₁, kuliko Yenghi-nun tosekwan-ey
tosekwan-ey kassta. (Ø₁: ka-ko 'go and', Ø₂: Yenghi-nun)

(342-344) show that subject ellipsis in coordinate structures
always proceeds forward; no matter whether the verbs of the
two conjuncts are different, as in (342), or identical, as
in (343); or whether the verbs are not ellipted, as in (343),
or ellipted, as in (344).

On the other hand, ellipsis of a verbal predicate al-
ways proceeds backward, as in (345):

(345) a Chelswu-nun pap-ul Ø, Yenghi-nun ttek-ul mekessta.
TOP rice OM TOP oake OM ate
'Chelswu Ø rice, and Yenghi ate cakes.'
(Ø: mek-ko 'eat and')
b *Chelswu-nun pap-ul mek-ko, Yenghi-nun ttek-ul Ø.
'Chelswu ate rice and Yenghi Ø cakes.'
(Ø: mekessta 'ate')

In case both the subject and the verbal predicate are el-
lipted, the former is ellipted forward and the latter
backward, as in (346):

(346)a Chelswu-nun pap-ul \( \emptyset_1 \) kuliko \( \emptyset_2 \) ttek-ul mekessta.  
TOP rice OM and cake OM ate  
'Chelswu \( \emptyset_1 \) rice and ate cakes.'  
(\( \emptyset_1 \): mek-ko 'eat and', \( \emptyset_2 \): Chelswu)

b *\( \emptyset_2 \) pap-ul mek-ko, Chelswu-nun ttek-ul \( \emptyset_1 \).  
'\( \emptyset_2 \) ate rice, and Chelswu \( \emptyset_1 \) cakes.'  
(\( \emptyset_1 \): mekessta 'ate', \( \emptyset_2 \): Chelswu)

For the object, the direction of ellipsis is determined by the ellipsis of the verb. Consider:

(347)a Chelswu-nun Yenghi-lul salangha-ko, Yengswu-nun \( \emptyset \) miwehanta.  
TOP OM love and hate  
'Chelswu loves Yenghi, and Yengswu hates \( \emptyset \).'  
(\( \emptyset \): Yenghi-lul)

b *Chelswu-nun \( \emptyset \) salangha-ko, Yengswu-nun Yenghi-lul miwehanta.  
'Chelswu loves \( \emptyset \), and Yenghi hates Yenghi.'  
(\( \emptyset \): Yenghi-lul)

(348)a Chelswu-nun Yenghi-eykey chayk-ul cwuess-ko, Yengswu-nun Swunhi-eykey \( \emptyset \) cwuessta.  
TOP to book OM gave and gave  
'Chelswu gave a book to Yenghi, and Yengswu gave \( \emptyset \) to Swunhi.'  
(\( \emptyset \): chayk-ul 'book')

b *Chelswu-nun Yenghi-eykey \( \emptyset \) cwues-ko, Yengswu-nun Swunhi-eykey chayk-ul cwuessta.  
'Chelswu gave \( \emptyset \) to Yenghi, and Yengswu gave a book to Swunhi.'  
(\( \emptyset \): chayk-ul 'book')

(349)a Chelswu-nun Yenghi-eykey \( \emptyset_1 \) \( \emptyset_2 \), Yengswu-nun Swunhi-eykey chayk-ul cwuessta.  
TOP to TOP  
'Swunhi-eykey chayk-ul cwuessta.'  
'Chelswu \( \emptyset_1 \) \( \emptyset_2 \) to Yenghi, and Yengswu gave a
book to Swunhi.'  
(Ø₁: chayk-ul 'book', Ø₂: cwuess-ko 'gave and')

b *Chelswu-nun Yenghi-eykey chayk-ul Ø₂, Yengswu-nun Swunhi-eykey Ø₁ cwuessta.  
'Chelswu Ø₂ a book to Yenghi, and Yengswu gave Ø₁ to Swunhi.'  
(Ø₁: chayk-ul 'book', Ø₂: cwuess-ko 'gave and')

(347-348) show that the ellipsis of the object proceeds forward if the verb is not deleted. It does not matter whether the verbs of the two conjuncts are different, as in (347), or identical, as in (348). However, if the verbs of the two conjuncts are identical and the one in the left conjunct is ellipted, as in (349a), then object ellipsis always proceeds backward, as in (349a-b). Thus, the correct generalization is that the object can be ellipted either forward or backward. Backward ellipsis always co-occurs with a simultaneous verbal ellipsis which also applies backward.

This generalization holds true in the ellipses of other VP constituents such as locatives, datives, and temporal NP's. Consider the following sentences involving the ellipsis of a locative NP. Compare (350) with (351):

(350)a Chelswu-nun Yenghi-lul hakkyo-eyse manna-ko,  
TOP OM school at meet and

Yengswu-nun Swunhi-lul Ø mannassta,  
TOP OM met

'Chelswu met Yenghi at school, and Yengswu met Swunhi Ø.'  (Ø: hakkyo-eyse 'at school')

b *Chelswu-nun Yenghi-lul Ø manna-ko, Yengswu-nun Swunhi-lul hakkyo-eyse mannassta.
'Chelswu met Yenghi ø, and Yengswu met Swunhi at school.' (ø: hakkyo-eyse 'at school')

(350b) is not acceptable in its intended reading.

(351a) Chelswu-nun Yenghi-lul ø₁ ø₂, Yengswu-nun Swunhi-lul ø₁
       TOP OM TOP
       hi-lul hakkyo-eyse manassta.
       OM school at met

'Chelswu ø₂ Yenghi ø₁, and Yengswu met Swunhi at school.'
(ø₁: hakkyo-eyse 'at school', ø₂: manna-ko 'meet and')

b *Chelswu-nun Yenghi-lul hakkyo-eyse ø₂, Yengswu-nun Swunhi-lul ø₁ manassta.

'Chelswu ø, Yenghi at school, and Yengswu met Swunhi ø₁.' (ø₁: hakkyo-eyse 'at school',
       ø₂: manna-ko 'meet and')

The patterns of ellipses in coordinate structure may be summarized as follows: (a) the subject ellipsis is forward, (b) the verbal ellipsis is backward, and (c) the ellipsis of VP constituents (object, dative, time, and location NP's) is either forward or backward. Backward ellipsis co-occurs with the verbal ellipsis.

So far, coordinate ellipses have been described in terms of grammatical functions, but if they are analyzed in terms of grammatical categories we can obtain a better generalization. We propose that coordinate ellipses be again grouped into two types: nominal and verbal (see Chap 1: 2.1). We also maintain that the nominal VP constituents in a coordinate structure are ambivalent with regard to ellipsis: when they are ellipted along with the verbal predicate they
behave just like the verbal; but when their ellipses are not accompanied by the ellipsis of the verbal, they behave like nominals such as the subject. In other words, we suggest that ellipsis of the nominals in the VP be considered a part of the verbal ellipsis when the verbal is also ellipted, but that it be considered nominal ellipsis otherwise. Then, we have a neat pattern: all nominal ellipses proceed forward and all verbal ellipses apply backward. In order to maintain this generalization, we propose that Gapping, which we restricted to the ellipsis of the verbal predicate in coordinate structures be expanded in scope and be considered a VP ellipsis rule which has the verbal predicate as its obligatory target and other nominal constituents of the VP as its optional targets. Then, we can define Gapping as follows:

(352) **Gapping**

Gapping is the ellipsis of VP constituents in a coordinate structure. It always proceeds backwards.

It is not difficult to find the reasons why nominal ellipsis in coordinate structure proceeds forward and Gapping proceeds backward. The forward direction of nominal ellipsis seems to be subject to Langacker's precede-command condition on pronominalization. Langacker (1969) noted that backward pronominalization is not allowed if the victim and the antecedent are the elements of two separate coordinated sentences. The same condition applies to coordinate ellipsis. Ellipsis is not applicable backward if the victim and the antecedent appear in separate coordinate structures.
On the other hand, backward application of Gapping seems to be dictated by the structural requirement of Korean that every sentence must end in a verb (cf. Kuno 1973). Due to this 'verb-final' requirement, the verb in the right conjunct cannot be ellipted and the ellipsis must proceed backward. Furthermore, Hankamer (1979:69) has claimed that in no languages can the sentence-final constituent of the basic sentence pattern be ellipted in coordinate ellipsis. For example, the object, a sentence-final element in SVO languages like English, cannot be ellipted in this language. Consider:

\[(353)a \quad \text{John cooked } \emptyset, \text{ and Harry ate the eggplant.} \\
\quad b \quad *\text{John cooked the eggplant, and Harry ate } \emptyset.\]  
(\(\emptyset\): the eggplant)

If this typological finding is correct, then the verb, the sentence-final constituent in SOV languages like Korean, cannot be ellipted in sentence-final position, and so backward ellipsis is the only option available.

4.3 We are now ready to explain why Gapping is subject to the various constraints mentioned in 4.1. The first factor which seems to bear on this problem is the nature of the conjunctors used in Gapping. It was mentioned that the only conjunctors allowed in Gapping are symmetrical ones such as 'and' and 'or'. Because of this property of symmetry, the two sentences around them are also expected to be symmetrical in structure and meaning. Otherwise, there will be a
semantic conflict between the conjunctor and the two conjuncts around it. If the syntactic categories and semantic properties of the constituents of the two conjuncts are not parallel, they will be incompatible with the semantic feature of 'symmetry' implied by the conjunctor 'and' or 'or'.

The second factor is related to perceptual difficulty (cf. Sanders 1976) and to the communicative function of Gapping. If ellipsis takes place in the right conjunct--i.e. applies forward--all of the possible antecedents and their relations have already been understood and are available in prior memory when the site of ellipsis is first encountered in the right elliptical conjunct. In contrast, when ellipsis sites are encountered in the left conjunct, the missing elements and their relations are not available in memory at that time. So, the interpretation process must be suspended --and all previously obtained results held in memory--until the necessary cues are received in the right conjunct. Then, it is necessary to go back and complete the interpretation of the left conjunct. This process is understandably a lot more difficult to perform than the interpretation task involved in forward ellipsis, where the referents have already been supplied.

It seems that the strict parallelism in structure and semantics required for Gapping is a way to compensate for this extra difficulty in interpretation. For example, if the order of the constituents in the two conjuncts are different, the interpretation will be hindered much more than
when the order of the constituents is identical. This interpretational difficulty also seems to explain why the fewer the number of the remnants after Gapping is, the more acceptable the elliptical coordinate sentence becomes. For example, if only one constituent remains after Gapping, as in (15), it can easily be associated with the elliptical verbal, e.g., kassta 'went' in (15). However, if more than three constituents remain after Gapping, as in (339-340), it would be considerably more difficult to recover the relations among them because no verb immediately follows them.

Moreover, the communicative function of Gapping involves contrast between the two elliptical conjuncts. Contrast is most effectively achieved where there is only one focused entity in each of the conjuncts. If there are more than two remnants after Gapping, the contrast which Gapping tries to capture will be less clearly in focus. This seems to be another functional explanation for why the number of remnants is so limited in gapped sentences.

For the unavailability of subordinate clauses for Gapping, we propose another functional explanation based on perceptual difficulty. Let us repeat (335) for ease of reference.

(335) *Chelswu-nun [apeci-lul Ø] tongsayng-ul yongse-
TOP father OM brother OM forgave

hayss-ko, Yengswu-nun [emeni-lul kippukey-
and TOP mother OM please

halye-ko] anay-lul yongsehayssta.
to wife OM forgave
'Chelswu forgave his brother Ø his father, and Yengswu forgave his wife to please his mother.' (Ø: kippukey-halye-ko 'to please')

As is clear from (335), Gapping in a subordinate clause always leads to a structural ambiguity since it brings NP's or PP's in the subordinate clause and the matrix clause together, e.g., apeci-ului tongsayng-ul 'father brother.' Structural ambiguities of this kind will make the interpretation process confusing and difficult. For this reason, we propose a constraint similar to that proposed for English by Hankamer (1979):

(354) Non-ambiguity Condition
Any ellipsis which leads to the creation of structural ambiguity is not allowed.

Finally, let us consider sentences such as (349b) and (351b). We will repeat (349b):

(349)b Chelswu-nun Yenghi-eykey chayk-ul Ø₂, Yengswu-
TOP to book OM

nun Swunhi-eykey Ø₁ cwuessta.
TOP to 1 gave

'Chelswu Ø₁ a book to Yenghi, and Yengswu gave Ø₂ to Swunhi.'
(Ø₁: chayk-ul 'book', Ø₂: cwues-ko 'gave and')

The main reason for the unacceptability of this sentence is that we have to make two opposite references: one is a cataphoric reference to cwuessta 'gave', and the other is a backward reference to chayk 'book'. This kind of cross-referencing must be too confusing to be interpreted. Facts like this lead to another perceptual constraint on Gapping:
(355) Ban on Cross-referencing

Simultaneous forward and backward references are not allowed in Gapping.

Note that this constraint also provides a partial explanation of why Gapping must be a VP ellipsis rule, rather than a single verbal ellipsis. If Gapping ellipts only the verbal predicate, it would violate the above constraint because the simultaneous nominal ellipsis (e.g., chayk in (349b)) should proceed forward.

4.4 We have so far assumed that the rule of ellipsis in coordinate structures (except for Gapping) is identical to that for other discourse ellipses. However, there is evidence that a simple ellipsis rule is not sufficient to handle all of the elliptical coordinate sentences and that coordinate ellipsis requires one more rule, which we will call "Conjoining." Consider the following sentences.

    and or TOP school to went
    'Chelswu and (or) Yenghi went to school.'

(357) Yengswu-nun Chelswu-wa/na Yenghi-lul mannassta.
    TOP and or OM met
    'Yengswu met Chelswu and (or) Yenghi.'

(358) Chelswu-nun Yenghi-lul ttayli-ko/kena chassta.
    TOP OM hit and or kicked
    'Chelswu hit and (or) kicked Yenghi.'

We will assume that sentences like (356-358), particularly (356-357) that have the phrasal conjunctor wa 'and', are elliptical coordinate sentences which have two sentences combined with the sentential conjunctor kuliko (or its
Then, (356), for example, should have the following under-
lying structure

(359) Chelswu-nun hakkyo-ey kassta kuliko Yenghi-nun
TOP school to went
hakkyo-ey kassta.
"Chelswu went to school and Yenghi went to school."
with the following phrase marker

(360) S
   \_____________/    \_____________/    \_____________/  
   S_1 --kuliko-- S_2   
   \___________/   \___________/   \___________/  
   NP          VP    NP        VP
   |          |      |          |      |          |  
   Chelswu   PP   Yenghi PP   hakkyo-ey kassta
   hakkyo-ey kassta

Since PP and V are identical in S_1 and S_2, one of them will
be deleted to produce the structure (361):

(361) S
   \_____________/    \_____________/  
   S_1 --kuliko-- S_2    
   \___________/  \___________/  
   NP          VP
   |          |      |          |  
   Chelswu   PP   Yenghi hakkyo-ey kassta
   hakkyo-ey kassta

If we have only a single rule of ellipsis, (361) should be
the final output of the derivation, but it is clear that
(361) is not the correct surface constituent structure.

There are two pieces of evidence for this. First,
(361) predicts that any pause will occur between Chelswu and
the rest of the sentence, because a major constituent break falls in that place and Chelswu and kuliko Yenghi do not form a single constituent. However, if a pause occurs, it must be after Yenghi, and Chelswu and Yenghi should form a single constituent. Second, unlike English, which uses only one conjunct or and to combine any two constituents, Korean has two conjunctors (kuli)ko and wa both of which mean 'and'. The former is used to combine two sentences or verbs, and the latter wa to combine two lexical NP's. Thus, ko and wa might be called "sentential" and "phrasal" conjunctors, respectively (cf. W.C.Kim 1970), and they are in complementary distribution. It might be argued that the two conjunctors are contextual variants of a single conjunct or morpheme.

The fact that wa, not ko, appears in the surface forms of (356-357) is indicative of the fact that Chelswu and Yenghi must be two noun phrases conjoined under a single NP node, rather than under an S node, as is implied by the phrase marker (361).

Therefore, the surface structure of (356) must look like (362), not (361):

\[
\begin{array}{c}
S \\
| \\
NP \\
| \\
wa \\
| \\
NP \\
| \\
Chelswu \\
| \\
| \\
NP \\
| \\
PP \\
| \\
hakkyo-ey kassta
\end{array}
\]

Thus, the ellipsis rule is not sufficient to produce the surface structure (362), and we need one more rule which will be
called "Conjoining." This rule applies to the output of the coordinate ellipsis rule and is formulated as follows:

\[(363) \text{Conjoining}\]

If the ellipsis rule in coordinate structures (including Gapping) leaves only one constituent in one of the two conjuncts, the constituent and its counterpart in the other conjunct are restructured to a single conjoined constituent.

The rule of Conjoining will turn (361), the output of the ellipsis rule, into (362). The sentences (357-358) will be generated in the same way.

Incidentally, we would like to comment on the direction of ellipsis which should apply to derive sentences (356-358). It might be suggested that the direction of ellipsis of identical constituents in (360) does not matter. For instance, one might argue that either of the VP's in \(S_1\) or \(S_2\) can be deleted or that the VP in \(S_2\) is deleted under identity with the identical VP in \(S_1\). We assume that neither of these suggestions are correct and that the kind of ellipsis involved in deriving (362) from (360) must be Gapping. What is deleted in (360) is the VP of \(S_1\). This is a logical consequence of the proposed analysis--when a VP is ellipted, the direction of ellipsis must be backward. On the other hand, the direction of ellipsis involved in deriving (358) must be forward, and the deleted constituent from the underlying structure (364)

\[(364) \text{[Chelswu-nun Yenghi-lul ttaylyessta]} S_1 \text{ kuliko TOP OM hit } S_1 \text{ and}\]
must be the underlined constituents in $S_2$. Finally, sentence (357) is derived from (365)

(365) $[\text{Yengswu-nun} \text{ Chelswu-lul} \text{ mannaststa}]_{S_1}$ kuliko $\text{TOP OM met 1}$ and $\text{TOP OM met 2}$

by the usual ellipsis rule, which deletes $\text{Yengswu}$ from $S_2$, and by Gapping, which deletes $\text{mannastta}$ from $S_1$. After that, the rule of Conjoining will apply to produce the correct surface structure. 38

We have argued that Conjoining applies when only one constituent is left in one of the two conjuncts after the ellipsis rule has applied in a coordinate structure. Now, let us apply this proposal to (321-325) given in 4.1. According to our proposal, (321-322) will undergo Conjoining because the ellipsis rule will produce an intermediate structure like (366):

(366) $[[\text{Yenghi-nun}]_{NP} [\text{hakkyo-ey ka}]_{VP} -ko[\text{[]}]_{NP}$

$[\text{tapang-eyto kasstt}]_{VP} ]_{S}$

which meets the SD of the rule of Conjoining (A-over-A principle applies here). Then, (366) is restructured to (367):

(367) $[[\text{Yenghi-nun}]_{NP} [[[\text{hakkyo-ey ka}]_{VP}-ko [\text{tapang-eyto kasstta}]_{VP} ]_{VP} ]_{S}$
which is correct.

According to our proposal, however, (323-325) will not undergo Conjoining. Note that in (323-325) there are two constituents in the ellipted conjunct after the application of the ellipsis rule. The following (368-369) are the outputs of the ellipsis rule applied to (323) and (325) respectively.

\[
(368)[Chelswu\text{NP}[ttek-ul mantu]\text{VP}]_{S_1} \text{-ko}[[Yenghi]\text{NP} \\
[\emptyset \text{mekessta}]_{VP} ]_{S_2}
\]

\[
(369)[Chelswu\text{NP}[Seoul-lo \emptyset]\text{VP}]_{S_1} \text{kuliko} [[Yenghi]\text{NP} \\
[\text{Pusan-ulo kassta}]_{VP} ]_{S_2}
\]

It has been suggested, however, that sentences such as (324-325) which involve Gapping must also undergo some sort of Conjoining. For example, Ross (1971) proposed that (369) should be further related to (370) by a conjunction reduction rule similar to our Conjoining:

\[
(370)
\]

Ross has not formulated any rule to show just how (369) is converted to (370), but there is no doubt that it would not
be an easy task. To support his claim that the surface
structure of (324-325) must look like (370) rather than
(369), Ross gives the following argument. For some Japanese
speakers, a major constituent break falls immediately before
the final verb in the gapped structures of the Japanese coun­
terparts of (324-325)--i.e., [SO, SO // v], not [SO, // SOV].
That is, a pause is placed before the verb kata, not after
Seoul-lo in (325). Ross takes this as evidence supporting
(370) rather than (369) as the correct surface constituent
structure.

This intonation-break argument is very shaky. To me,
an intonation break seems to fall after Seoul-lo in (325).
Since this whole argument is unclear, we regard the surface
structure (370) as not motivated. On the contrary, there
are two pieces of evidence that (369), not (370), is the cor­
rect surface constituent structure. First note that the
phrase marker (370) is actually exceptional. The structural
configuration [S V] is allowed in Korean only when a sen­
tence is embedded as a nominalized NP constituent. However,
the node S3 of (370) cannot be considered a nominalized S.
Emonds (1970) argues that a transformation cannot create a
surface structure not accommodated by the base rules of the
language. 39 Phrase marker (370) is a new constituent
structure which is not found in the base rules of Korean.
Therefore, the structure (370) is difficult to justify.

Second, let us consider honorification phenomena. Con­
sider the following sentences.
(371) a na-wa sensayngnim-un Yenghi-lul{salanghanta. }
and teacher TOP OM love
'I and the teacher love Yenghi.'

b sensayngnim-kwa na-nun Yenghi-lul{salanghanta. }
'The teacher and I love Yenghi.'

(372) a na-nun Swunhi-lul Ø, sensayngnim-un Yenghi-lul
TOP OM teacher TOP OM
dsalamhanta }
salamhasinta.
'I Ø Swunhi, and the teacher loves Yenghi.'
(Ø: salangha-ko 'love and')

b sensayngnim-un Swunhi-lul Ø, na-nun Yenghi-lul
teacher TOP OM I TOP OM
dsalamhanta }
salamhasinta.
'The teacher Ø Swunhi, and I love Yenghi.'
(Ø: salangha-si-ko 'love(HON) and')

(371a-b) show that when the subject NP of a sentence consists of two nouns, one of which is honorified, the other of which is not, the sentence cannot undergo honorification and the subject is not honored. (371) is not a problem to both analyses, but (372) is certainly a problem for Ross's proposal. (372a-b) will be given the following surface structure in his analysis:

(373)
Since the subject of the verb salanghanta is both na 'I' and sensayngnim 'teacher' in (373), Ross's analysis predicts that honorification should not apply, just as in (371). However, this prediction is not borne out, as (372a) shows. Note that honorification is sensitive only to the subject of \( S_2 \) in (372a).

In our analysis, (371a-b) have undergone Conjoining and so honorification is blocked. However, (372a-b) do not undergo Conjoining because the output of Gapping in the left conjunct has more than one constituent. Thus, the surface structure of (372a-b) are as follows:

\[
\begin{array}{c}
\text{NP} \\
[\text{sensayngnim} \text{Swunhi na}] \\
\text{V} \\
[\text{sensayngnim}] \\
\text{NP} \\
[\text{salanghasinta}] \\
\text{VP} \\
\end{array}
\]

The honorification phenomena in (372a-b) are automatically explained in the surface phrase marker (374). Based on these facts, we conclude that Conjoining applies only when one constituent is left in one of the two conjuncts of a coordinate structure after ellipsis, and that the outputs of Gapping in (324-325) do not undergo Conjoining.

To summarize, we have argued that ellipsis in coordinate structures should be divided into Gapping and nominal ellipsis. The latter is essentially identical to other discourse ellipses. Gapping, however, is subject to various
constraints because of its backward application. When the output of the ellipsis in coordinate structure has a single constituent left in one of the two conjuncts, it undergoes one more rule, Conjoining. This final operation is unique to coordinate ellipsis.
Notes to Chapter II

1 One might wish to take issue with this use of "generative." For this criticism, see ft. 6 below. The deltas in underlying structures like (124) stand for forms that exist linguistically but are not realized phonologically.

2 If a grammar of incomplete sentences takes the speaker's position, it will end up in a deletion analysis, and the direction of the arrows will be upward starting from the semantic structure. We have found that this analysis is not plausible.

3 Jackendoff's pronoun interpretive rules (Jackendoff 1972: 174-177) and Dougherty's (1969) are limited to what is overtly expressed on the surface.

4 Pronominal and zero anaphora are the same in phoric relations. Both make use of certain pro-forms rather than fully specified lexical formatives in order to refer to some identifiable object in the context. The only difference between the two is that the former has a more restricted range of possible referents than the latter.

5 One might ask why 'topic' is a syntactic function, arguing that it is a purely semantic or pragmatic function that has nothing to do with the grammatical structure. This may be so for English, but we argue that topic in Korean is a syntactic category which also has semantico-pragmatic functions (see Chapter 3). Topic as a nuclear syntactic function for Korean is justified in Chapter 3. For the distinction between nuclear and non-nuclear syntactic functions, see Dik (1978) and Parisi and Antinucci (1976).

6 One might ask how this syntactic reconstruction process is 'generative.' The term 'generative' has been used in various ways (see Parret 1974: 27-28, 174, 249-250), but in Chomsky's original sense (Chomsky 1957, 1965: 9) it means that the grammar should provide a structural description for each permissible sentence in a language. The syntactic reconstruction process which we will discuss does just that, and is 'generative' in this sense.

7 There are cases in which it is difficult to judge whether an argument is obligatory or optional for a given predicate. kata 'go' is one such example. It is not clear whether the goal argument should be considered obligatory or optional. For example,

(1) Chelswu-nun ka-la.
    TOP go Imp.
    'Chelswu! go!'
If one simply wants Chelswu to leave without having a particular destination in mind, it seems that the goal argument is optional.

A similar difficulty is also observed in the verbs of psychological perception: yevpputa 'pretty', cohta 'good', mas-issta 'tasty', etc.

(ii) ku yenghwa-nun caymi-issta.
the movie TOP interest exist
'The movie is interesting.'

(iii) na-nun ku yenghwa-ka caymi-issta.
I TOP the film SM interest exist
'I find the movie interesting.'

It is not clear whether or not an experiencer argument like na 'I' in (iii) is obligatory with the verbal caymi-issta 'interesting'.

The modal operator keyss and the honorific marker si are also made use of in reconstructing the underlying structures such as (141e-142e).

Note that the object marker luI prevents the NP Chelswu from being analyzed as the subject NP. The significance of case markers for the generation process will be discussed later.

Looking for the context for information is not 'generative'. 'Generative' here means assigning structural descriptions to an incomplete sentence based on the sentence-internal information. Put in the strongest possible terms, the generation process need not consult the context.

As already mentioned in ft.12 of Chapter 1, (146b) is acceptable as a discourse-initial sentence in appropriate contexts. As an example of the same sort, consider the following sentence. Two boys are fishing at a pond. One of the boys comes to the other and says:

(i) # na-nun tases mali-ita.
you TOP five is
'I am five.'

Actually the occurrence of the copula be must be a surfacy and language-particular phenomenon of Korean: e.g. other languages such as Russian do not require the copula in sentences like (154-155) (cf. Parisi and Antinucci 1976).

Ellipses of the elements of the verbal predicate such as polarity, modality, voice, tense, and illocutionary forces are another set of complicated problems, which will not be considered in this study. For the operator ellipses involved
in VPD (verb phrase deletion) in English, see Sag (1977), and Halliday and Hasan (1976:174-196) give a detailed description of these phenomena.

14Delimiters such as man 'only' and to 'also' are of no use in the generation process.

15This is the view which is adopted by most of the deletion and interpretive analyses: Chomsky (1965); Hankamer (1979:334); Hankamer and Sag (1976); Jackendoff (1972); Shopen (1972a); Grinder (1976).

16The structure of the hypothetical world UOD is not clear. We will try to clarify this notion more concretely in a moment, but Winograd (1972) and Schank (1972) are good reference sources for the internal make-up of this theoretical concept.

17An exophoric pronominal anaphora as in (i) is explained in the same way:

(i) # ku-nun cal sayngkyess-kwuman.
   he TOP well looking
   'He is handsome.'

18Elliptical subjects in the so-called explicit performative sentences which have an overt performative verb on the surface, e.g.,

(i) Che1swu-nun Yenghi-lul mannassul-ttay,
   TOP OM met when
   'When Chelswu met Yenghi,'

(ii)a (Ø kkok party-ey kan-ta)-ko yaksokhayssta.
   certainly to go Quot promised
   'Ø promised that Ø would certainly go to the party.' (Ø: Chelswu, *Yenghi)

   b (Ø kkok party-ey ka-la)-ko myenglyenghayessta.
      Imp ordered
   'He ordered Ø to go to the party by any means.'
   (Ø: Yenghi, *Chelswu)

   c (Ø kkok party-ey ka-ca)-ko oevanhayssta.
      Propose proposed
   'He proposed that Ø go to the party by any means.'
   (Ø: Chelswu and Yenghi)

19can be interpreted in terms of the semantic properties of these verbs, which will be discussed in 3.2.1.3.

19For a nice example for this point, see Bates (1976:172-174).
20. Note that the notion of "foregroundedness" is a relative concept. Some entities are more foregrounded and some are less. However, at any point of a discourse, there are always most foregrounded entities. Two problems with the notion 'foregrounding' are (i) how to pinpoint precisely when a foregrounded entity ceases to be foregrounded and retreats into the background, and (ii) how to define relative degrees of foregroundedness among discourse entities in consciousness (cf. Chafe 1972).

21. Though she did not elaborate, this is why Gundel (1977) claims that "what is deleted is activated topics." We do not accept this claim that deletion is topic-conditioned, because we do not accept the "informational" definition of topic such as that of Gundel (1974). For details, see Chapter 3. Also consider the following sentences.

(i) (Chelswu is in a restaurant and a waitress comes up to him to take an order. Chelswu says:)

#nayngmyen.
'Cold noodles.'

(ii) a Speaker A: ne-nun nwukwu-wa hakkyo-ey kassessta?
you TOP who with school to went
'Who did you go to school with?'

b Speaker B: \( \varnothing_1 \) Chelswu-hako \( \varnothing_2 \) yo.
with
'With Chelswu.'

\( \varnothing_1: \text{na 'I', } \varnothing_2: \text{hakkyo-ey kassessta } \) ('went to school')

What is deleted above is \( \text{ne 'you', na 'I' and } \text{cwuta 'give'} \) in (i), and \( \text{na 'I' and hakkyo-ey kassessta 'went to school'} \) in (ii). Gundel would argue that the topics of (i) are \( \text{ne-ka na-eykey cwuta 'you give me'} \) and \( \text{nay-ka hakkyo-ey kassessta 'I went to school'} \), respectively. However, we argue that the topics of (i) and (ii) are simply \( \text{ne 'you'} \) and \( \text{na 'I'} \). If this is correct, it is clear that deletions involve more than 'topic', and that 'topic' cannot offer a comprehensive account for deletion phenomena.

22. Recency of mention is not enough. Often there is more than one entity in the immediately preceding sentence, and only one of them is associated with the zero anaphor. One such problem will be discussed in 3.2.1.4.

23. The zero anaphor of (214b) can be interpreted even as the speaker himself, if an appropriate discourse context is given. For example, in (i), the speaker is observing himself as an objective entity:
(i) a nay-ka way ilehkey kongpwu-ka an toy-ci?
I SM why like-this study SM not go
'Why can't I concentrate on my study?'

b ø yenghwa-lul po-ko siphun kes-i thullim-epsta.
(=(214b))
(ø: na 'I')

24 As we will see in a moment, verbs may be classified into three groups: actor-biased, patient-biased, or neutral. However, the causality bias of verbs is actually a gradient phenomenon, which may be graphically shown as follows:

Causality bias continuum

<table>
<thead>
<tr>
<th>actor-biased</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>patient-biased</th>
</tr>
</thead>
</table>

Some verbs are clearly actor-biased (point A), clearly patient-biased (point E), or neutral (point C). But there are many verbs which are normally actor-biased rather than patient-biased (point B) or vice versa (point D).

For example, the causality bias of the verb hangpokhata 'surrender' is not as clear as those of the other verbs given in this group. Consider

(i) ilpon-un mikwuk-eykey hangpokhayestta. (=221a)
'Japan surrendered to America.'

(ii) a ø cenuy-lul sangsilhayess-ki ttaymwun-iessta.
'Because ø lost the will to fight'
(ø: ilpon 'Japan')
(=(221b))

b ?ø nemwu kanghayss-ki ttaymwun-iessta.
too strong because was
'Because ø was too strong.'
(ø: mikwuk 'America')

(iiib) does not sound as natural as (iia). If the zero of (iiib) refers to mikwuk, the goal (or patient) of surrendering, it will normally be specified, not ellipted. In this sense, hangpokhata is normally actor-biased rather than patient-biased.

Similarly, verbs like pokconghata 'obey', chacakata 'go to', and pwuthakhata 'request' are usually actor-biased rather than patient-biased. Consider

(iii) Chelswu-nun Yenghi-lul chacakassta.
TOP OM went-to
'Chelswu went to Yenghi.'
(iv) a Ø ton -i philyohayss-ki ttaymwun-ey.
    money SM needed because
    'Because Ø needed money.' (Ø: Chelswu, ?? 'Yenghi)

b ?Ø ton -i manh-ki ttaymwun-ey.
    money SM much because
    'Because Ø has much money.' (Ø: Yenghi)

c Yenghi-ka ton-i manh-ki ttaymwun-ey.
    SM money much because
    'Because Yenghi has much money.'

(iv-a) and (iv-c) are fully acceptable. Note that (iv-c) is
more natural than (iv-b) in this context.

25 The object particle luli can co-occur with verbs of
motion such as kata 'go' and ota 'come' in Korean. luli kass-
ta implies that the act of going ended at a certain loca-
tion.

26 Those who argue for the incorporation of logical
theories into linguistics are Reichenbach (1947), Allwood et
al. (1977), Partee (1975) among others; those who oppose to
this include Chomsky (cf. N. Chomsky. 1955 'Logical Syntax
and Semantics: Their Linguistic Relevance' in Language 31:
36-45). Even among the scholars who oppose this view, it
seems to be well-accepted that logical theories might prove
more useful in dealing with semantics than formal and struc-
tural aspects of language. Note that the interpretation
process is basically semantic.

27 There is no way to predict which out of the three
possible syntactic manifestations of the semantic concept of
'Chelswu's death' in (251b) has actually been deleted. This
is a problem to the syntactic deletion analysis. This kind
of problem can be avoided by using a logical representation
like (253).

28 Note that the same kind of a paradox is also found
in the analysis which argues that coreferentiality is needed
for deletion, because sentences (267-269) are counterexamples
to this claim. But the difference between the two is that
Hankamer has no explanation for why most zero anaphors (e.g.
(270b)) require coreferentiality, whereas we can explain why
some anaphors (e.g. (267b)) do not require coreferentiality.

29 When the verbs kata and ota do not involve the speaker,
the addressee and the relation of movement between the
two, the problems we have in sentences like (304-305) do not
arise. Consider
(i)a Chelswu-nun nayil ku-uy samwusil-ey kanta.
   'Chelswu goes to his office tomorrow.'

b Yenghi-nun ø cip-ita.
   'For Yenghi, it is her home.' (ø: kata 'go')

(ii)a na-nun hakkyo-ey kanta.
   'I go to school.'

b ne-nun eti-ey ø ?
   'How about you?' (ø: ka-ni 'go')

(iii)a Chelswu-ka encey cip-ey ol-kka?
   'When will Chelswu come home?'

b 3 si ø yo.
   'At three o'clock.' (ø: ota 'come')

The speaker and the addressee are not involved in (i-ii) as the starting and ending points of movement. When both the speaker and the addressee are assumed to be in the same place, no complications arise, either.

30 The following discussion of Gapping and ellipsis in coordinate structures can be recast in the interpretive framework which this thesis adopted. Kac (1978: Chap 5) and Williams (1977a) have demonstrated that ellipsis in coordinate structures can be described in the interpretive framework without any loss of generality.

31 The conjunctor and can be asymmetrical in sentences like the following:

(i)a Harry stood up and objected to the proposal.
   b Harry objected to the proposal and stood up.

For details, see Schmerling (1973).

32 The sequence of the verb in the left conjunct and the following conjunct is usually contracted as follows: 'mantulessta; kulentey' → 'mantulessmuntey'; 'mantulessta; kulayse' → 'mantulesse'.

33 As already mentioned in Chap 1 (fn.11), this fact indicates that ellipsis rule must be a very late rule, applying even later than Scrambling.
The grammatical subject is excluded from this count, because the crucial elements in this ellipsis are the constituents of the VP, as we will argue in the next section.

This is also true in English, where Gapping is a kind of VP rule. Consider:

(i) Max wants to try to begin to write a novel and Alex (((to try) to begin) to write) a play.

It might be argued that Gapping can occur in subordinate clauses, as in the following sentences:

(i) Chelswu-nun Yenghi-ka ø, Yengswu-nun Swunhi-ka micyessta-ko malhayssta.
     crazy said
     'Chelswu ø Yenghi ø, and Yengswu said that Swunhi got crazy.'
     (ø: micyessta-ko malhayssta 'said that ø got crazy')

(ii) Chelswu-nun Yenghi-ka ø, Yengswu-nun Swunhi-ka tangsentoy-ki-lul palassta.
     be elected OM wanted
     'Chelswu ø Yenghi ø, and Yengswu wanted that Swunhi get elected.'
     (ø: tangsentoy-ki-lul palassta 'wanted that ø get elected')

However, note that the verbs of the embedded sentences are ellipted together with the verbs of the matrix sentences. It is not possible for the embedded verb alone to be ellipted.

     nun Swunhi-ka micyessta-ko malhayssta.
     TOP SM said and
     TOP SM crazy said
     'Chelswu said that Yenghi ø, and Yengswu said that Swunhi got crazy.'
     (ø: micyessta 'got crazy')

Note that Gapping in (iii) results in a structural ambiguity.

Cross-referencing is allowed if Gapping and the other usual ellipsis rule are both involved:
Conjoining' is a rule of Chomsky-adjoining, and is subject to A-over-A principle. For example, the actual derivational processes which produce (357) from (365) are as follows:

(i) \[\text{Yengswu-nun Chelswu-lul mannassta}_S \text{ kuliko } \text{Yengswu-nun Yenghi-lul mannassta}_S\]

(ii) \[\text{Yengswu-nun } \text{Chelswu-lul} \text{ } \text{mannissta}_S \text{ kuliko } [\text{Yenghi-lul mannassta}_S]_{VP}\]

After ellipsis and Gapping in (i), \(S_2\) in (ii) has a single VP constituent remaining (by A-over-A principle). Thus, Conjoining takes place to produce (iii):

(iii) \[\text{Yengswu-nun } [\text{Chelswu-lul}]_{VP} \text{ kuliko } [\text{Yenghi-lul mannassta}]_{VP}\]

(The node VP is Chomsky-adjointed.)

The VP in (iii) again has a coordinate structure with a single constituent remaining in the left conjunct. Thus, Conjoining re-applies to (iii) to produce (iv-v):

(iv) \[\text{Yengswu-nun } [\text{Chelswu-lul}]_{NP} \text{ kuliko } [\text{Yenghi-lul}]_{NP} \text{ mannassta}_V_{P}\]

(v) \[\text{Yengswu-nun } [\text{Chelswu-wa Yenghi-lul}]_{NP} \text{ mannassta}_V_{P}\]

Emonds also allows some rules that are not structure-preserving.
CHAPTER III

TOPIC

1. Introduction

In the traditional grammar of Korean (Choi 1975) topic was identified as the NP-nun in sentence-initial (henceforth S-initial) position and was treated as just a stylistic variant of the subject NP-ka. But it soon became clear that this was wrong. The topic need not be the NP in the role of subject; it may be the NP functioning as object, or even as locative, goal, or time.

Later, in transformational grammar, the topic was treated as a subject-predicate sentence constituent, which was fronted by means of a topicalization transformation (Oh 1971; C.M.Lee 1973). The topic role was thus considered a rather marginal syntactic phenomenon. Others, however, noting that the so-called topicalization operation is faced with derivational difficulties, have proposed that topic be generated directly by the FS rules of the language. Kuno (1973), Gundel (1974), and D.W.Yang (1975) have made such proposals. The former may be called the "transformational" analysis and the latter the "FS" analysis of topic. However, proponents of both theories seem primarily concerned with the syntactic derivation of the topic, either by a transformational or by a base rule, and have neglected such significant questions as why topic is used, and how it is generated in actual communicative situations.
Another problem with the notion of topic is definitional. While challenging the generally accepted assumption that the topic in Korean is the NP-nun, I.S. Yang (1973) argued that the notion not only includes the S-initial NP-nun but also other S-initial NP's such as case-marked NP's (e.g. NP-lul, NP-eykey 'to', etc.) and delimiter-attached NP's (e.g. NP-to 'also', NP-man 'only', etc.). I once entertained, and others such as Kuno (1979:8-9) have also suggested, a hypothesis that topic need not be the S-initial constituent and can be even a predicate. For example, in the following sentence

(1a) nwu-ka hakkyo-ey kass-ni?
    who SM school to went
    'Who went to school?'

b Chelswu-ka hakkyo-ey kasseyo.
   'Chelswu went to school.'

the topic is assumed to be the underlined predicate, because (1b) is about 'somebody went to school' from the informational viewpoint of the discourse. We will discuss this kind of characterization of topic shortly.

To make matters worse, what we call 'topic' here has also been known by various other terms: "theme" (in functional sentence perspective theories such as those of Mathesius (1928, 1939), Firbas (1964), and Kuno (1972) and also of Halliday (1967, 1970) in a different sense); "subject" (Strawson 1964; Searle 1969); "subject in the logic" (Kuroda 1965, 1972, 1976); "presupposition" (Chomsky 1972); "ground"
(Talmy 1979); and so on. Even the term 'topic' is understood in different ways by various scholars—see, for example, Gundel (1974) and Dahl (1969) as opposed to Magretta (1977). Some of these diverse terms refer more or less to the same concept, but others carry totally different meanings, though their extensions may partially overlap. It is important, then, to make explicit what we mean by 'topic.'

In this chapter we attempt first to clarify the notion of 'topic' and propose a reasonable definition of it for Korean, and then to provide an account of how topic is generated. In Section 2, we review four definitions or characterizations of topic proposed so far: informational, syntactic, functional, and morphological. We then propose that topic in Korean be defined either as the unstressed S-initial NP-\textit{nun} (which we will call "unmarked topic") or as a delimiter-attached definite NP in S-initial position (which we will call "focus topic"). Section 3 deals with the generation of topic. In 3.1 we examine the transformational operation of topicalization and argue that this analysis is not only undesirable theoretically but also inadequate on empirical grounds. In 3.2 we elaborate on the PS analysis of topic defended in this thesis, discussing when topics are or are not used. This is followed by discussion of a variety of factors involved in the process of topic choice. We then deal with the relationship between the topic and the comment. Finally, we present a hypothesis concerning the dual functions of the particles \textit{ka} and \textit{nun}. 
2. Definition

So far, four kinds of definitions or characterizations of 'topic' have been advanced. One is an informational characterization based on the discourse informational structure; another is a formal and structural definition based on the syntactic structure of a sentence; a third is a functional definition close to the formal definition; and a fourth is a morphological definition proposed specifically for languages like Korean and Japanese. The first three definitions, as we will see, appear to assume that topic is a universal notion. While a universal definition of topic is surely attractive, it is, as Li (1976) points out, extremely difficult, perhaps impossible, to formulate a universal definition of topic, at least at this stage of our understanding of the notion. We will not try to formulate a universally applicable characterization of topic. Our aim is much more modest. We will limit our concerns to Korean data, and try to define this concept language-internally. We will attempt to justify this definition and its usefulness for the description of Korean.

2.1 The Informational Definition

The informational definition, appearing in its strongest version in Gundel (1974) and Chomsky (1972), characterizes topic in terms of 'given' or 'old' information, regardless of the formal syntactic structure of the sentence. Lyons (1969:335) labels this approach to topic "topic as
contextual dispensability": 

the topic, or 'subject of discourse' is described as that element which is given in the general situation or in some explicit question to which the speaker is replying; and the comment as that part of the utterance which adds something new and thus communicates information to the hearer (Emphasis added).

According to this definition, one and the same sentence can have different topics depending upon the discourse contexts in which it appears. Jesperson (1924:146-147) gives a good example of this kind of characterization of topic. For example, the following sentence

(2) I am hungry.

can have at least three different topics. If (2) is used as an answer to the question in (3)

(3) Who is hungry?

then the topic of (2) will be 'somebody is hungry', which is the contextually given (or dispensable) information. If (2) is an answer to (4)

(4) Is there something amiss with you?

then the topic of (2) will be 'I'. If (2) follows (5)

(5) Are you really hungry?

the topic of (2) will be 'I ... hungry' and the comment 'am'. Thus, given an appropriate context, any constituent of (2) can be the topic of the sentence.
Chomsky (1972) has made a similar proposal. While equating the topic-comment structure of a sentence with its presupposition-focus bipartition in discourse contexts, Chomsky (1972:100) proposes that both are determined on the basis of the natural response and the range of permissible focus in the surface structure of a sentence. Thus, he says:

Each sentence is associated with a class of pairs \((F,P)\) where \(F\) is a focus and \(P\) a presupposition, each such pair corresponding to one possible (semantic) interpretation. ... The focus is a phrase containing the intonation center; the presupposition, an expression derived by replacing the focus by a variable (Emphasis added).

For example, the following sentence (6)

(6) Did the Red Sox play the Yankees?

... can have three topics (or presuppositions)

(7)a whom the Red Sox played
   b what the Red Sox did
   c what happened

... and three respective comments (or foci)

(8)a the Yankees
   b to play the Yankees
   c the Red Sox played the Yankees

... because (6) can have the following three natural responses:

(9)a No, the tigers.
   b No, they flew to Washington.
   c No, the game never took place.

Chomsky's claim that topic is anything minus the focus has been found untenable (cf. G. Lakoff 1971b: 261; Shopen 1972b:
basically because no distinction is made between normal stress (which typically falls on the S-final lexical item) and contrastive stress in English. Chomsky's proposal works only in the sentences containing the latter.  

Gundel (1974) defines the topic as the "presupposed (given or known) information" of the sentence. Though she does not make explicit what is meant by "presupposed", her presupposition appears to comprise at least two distinct notions: one "textual" and the other "existential." Textual presupposition is involved when the sentence contains contrastive stress; in this case her notion of 'topic' is exactly identical to that of Chomsky (1972)--anything minus the information focus is the topic. When the sentence has normal stress, Gundel argues that the topic is what is "existentially presupposed" or "given, either activated or unactivated" (Gundel 1977), a thesis similar to that of Strawson (1964).

We accept the Gundel-Strawson thesis that the topic of a sentence is existentially presupposed. We will refer to existential presupposition simply as "definiteness" (see 2.4.1.2). However, we will show that the Gundel-Chomsky claim that topic is anything minus the information focus when a textual presupposition is involved is wrong, at least for Korean.  

Gundel further argues that the topic-comment relation of a sentence need not correspond to its surface grammatical
structure. Since the topic is always presupposed or given, and the presupposed information is not always the S-initial element in surface structure, Gundel argues that there is no way to uniformly identify the topic at the surface structure by means of surface syntactic cues (e.g. the S-initial position in Chomsky (1972)). The only way to account in a uniform manner for topic-comment relations in various kinds of sentences is to separate the notion from surface syntax and represent it at a more abstract level, one she refers to as "the logical structure." Thus, she first opens the way for the topic-comment relation to lose its surface grammatical footing and to be represented more abstractly in deep structure. Later she devises various transformational rules to convert the topic-comment deep structure into a surface structure (cf. 1974: 93-96).

For example, given the following sentence (10)

(10) John took Bill to the zoo.

Gundel cannot tell what the topic of this sentence is; it can be John, Bill, or the zoo. But if (10) is an answer to (11), for example,

(11) What happened to Bill?

the topic of (10) is Bill. The deep structure of (10) occurring in this context will look like (12):

(12) [Bill:X]_{topic} [John took X to the zoo]_{comment}
In support of this deep structure she points out that (10) can be paraphrased as (13) without any serious meaning change.

\[(13) \text{As for Bill, John took him to the zoo.}\]

In (13), the expression 'as for' shows that Bill is the topic. Note that the topic need not be the S-initial element. Similarly for the following (14)

\[(14)\]
\[a\] What happened to John and Mary?
\[b\] John persuaded Mary to be examined by the doctor.

she argues that the topic of (14b) is John and Mary, because they are 'given' and (14b) can also be changed to (15):

\[(15) \text{As for John and Mary, he persuaded her to be examined by the doctor.}\]

Note that the topic can be represented in surface structure by discontinuous constituents.

Kuno (1979) also entertains a similar hypothesis. He argues that the topics of the (a) sentences of (16-19)

\[(16)\]
\[a\] John was born in 1960.
\[b\] John was born sometime.
\[(17)\]
\[a\] John was still a small boy in 1960.
\[b\] John was still something in 1960.
\[(18)\]
\[a\] Did you stay in a hotel in Paris?
\[b\] You stayed somewhere in Paris.
\[(19)\]
\[a\] Did you buy this perfume in Paris?
\[b\] You bought this perfume somewhere.

are the following (b) sentences respectively:
Note that his topics (16b-19b) are all informationally non-focal or old, i.e., less newsworthy than the comments.

What is common to these various versions of the informational view of topic is that topic is the "given" or "old" element of the sentence, the content whose informational status is largely determined by the discourse. The topic has nothing to do with the surface grammatical organization of the sentence. Below we will show that the equation of topic with the textually given or dispensable information is not adequate, at least for Korean, and that an abstract topic-comment analysis is untenable. Rather, we will argue later that while topic is clearly related to, and in part governed by, the informational structure of the discourse (3.2.3.2), topic is also tied to the surface grammatical structure of the sentence (2.4.1).

Unlike English, which does not usually mark the topic of the sentence grammatically, Korean marks topics morphologically. The problem is that the informational characterization of topic is incompatible with this surface-marked topic or, at least, cannot explain its behavior. For example, when sentences (16a-19a) are translated into Korean, the morphologically marked topics of these Korean sentences will be just John or you, but not (16b-19b), which are simply 'old information.' Also consider the following sentences:

(20) Chelswu-nun ecey Yenghi-hako mwues-ul hayssni?
TOP yesterday with what OM did
'What did Chelswu do with Yenghi yesterday?'
(21)a Chelswu-nun ecey Yenghi-hako hakkyo-ey kassesse. TOP yesterday with school to went 'Chelswu went to school with Yenghi yesterday.'

b Chelswu-wa Yenghi-nun ecey hakkyo-ey kassesse. TOP 'Chelswu and Yenghi went to school yesterday.'

Gundel (1974) and Kuno (1979) would predict that the topics of (21a-b) are both 'Chelswu did something with Yenghi yesterday' and the comments are both 'went to school'. However, (21a) and (21b) are clearly different in topichood (both formally and intuitively): (21a) is only about Chelswu, but (21b) is about Chelswu and Yenghi. Thus, (21a-b) show that the informational view of topic does not correctly characterize our intuitions about topic, and that the surface-coded topic in Korean does not always coincide with the 'old information' of the discourse.

The topic of a sentence is not something one can determine objectively on the basis of the discourse context in which the sentence appears. Topic is basically a speaker-centered subjective notion. Whatever the informational structure of the discourse, it is ultimately up to the speaker to decide what the topic of his sentence will be. In (21a-b), the speakers have chosen Chelswu and Chelswu-wa Yenghi as the topics of these sentences, whatever the given or old information is in the discourse.

Arguing that Chelswu-nun of (21a) is just a surface coding of the topic and that what (21a) is really about is 'Chelswu did something with Yenghi yesterday', one might
propose, with Gundel (1974), that the topic still be 'Chelswu did something with Yenghi yesterday' and that the comment be 'went to school', at least in deep structure. However, this kind of abstract topic-comment structure is faced with a number of technical problems in converting the alleged deep structure to the surface form. For example, consider the following sentences:

(22) Chelswu-nun swukcey-lul ha-ko issni?
    TOP homework OM do Progressive 'Is Chelswu doing homework?'

(23a) anya. Chelswu-nun (imi) swukcey-lul haysse.
    no TOP already homework did 'No, Chelswu already did homework.'

(23b) anya. Chelswu-nun (icey) swukcey-lul ha-l-kkeya.
    no TOP now homework OM do will 'No, Chelswu will do homework now.'

(23a-b) are appropriate responses to (22). What is presupposed or given in (23a-b) is the fact that Chelswu does homework at a certain time. The specification of the time is the focal point. Thus, if the abstract notion of topic-comment is adopted, the underlying structures of (23a-b) should look like the following (24a-b):

(24a) [Chelswu-ka swukcey-lul hata]topic [Past]comment

(24b) [Chelswu-ka swukcey-lui hata]topic [Future]comment

This analysis has several major flaws. First, the alleged topic does not reflect correctly the intuition that (23a-b) are only about Chelswu. Second, it is a matter of extreme
difficulty to set up a derivational process which converts (24a-b) to (23a-b).

Let us consider one more example. It is well-known that in many languages complementizers have semantic content (Bresnan 1970; I.S.Yang 1972). Consider

(25)a Chelswu-nun cip-ey o-ko issta.
   TOP house to come Progressive
   'Chelswu is coming home.'

b Chelswu-nun cip-ey o-a issta.
   Perfective
   'Chelswu has come home.'

The only difference between (25a) and (25b) is in the complementizers used: a indicates that the action has been completed and ko means that the action is still going on. Thus, if the informational abstract topic-comment underlying structure is adopted, the topics of (25a-b) are 'Chelswu comes home' and the foci are represented by a and ko, respectively. As in (24a-b), the alleged topic is not correct, and the problems of deriving (25a-b) from the abstract topic-comment underlying structure are insurmountable.

The conclusion to be drawn from sentences (20-25) is that topic cannot be equated with the contextually given or presupposed information in the discourse, either in deep or surface structure. The correct generalization would be that topic is part of the 'old' or 'given (textually)' information, and that an equation such as 'topic = given (textually)' holds true only in pseudo-cleft sentences, which is a type of topic-comment structure.
(26) a nusu-ka hakkoye ey kassni?
   who SM school to went
   'Who went to school?'

   b hakkoye ey ka-n kes-un Chelswu-ty.
   school to go that TOP be
   'He who went to school is Chelswu.'

where the whole textual presupposition 'somebody went to school' is taken as the topic. However, this does not mean that topic has nothing to do with the discourse informational structure at all. In 3.2.3.2, we will argue that topic choice is partly governed by the information structure of the discourse.

2.2 The Syntactic and Functional Definitions

Formal and structural definitions identify topic as the first constituent of a sentence in its surface form. A formal definition was first proposed by Travnicek, followed by Halliday (1967; 1970) and Chomsky (1965), and more recently by Magretta (1977).

Halliday, who is a strong proponent of this definition, argues that topic ("theme" in his term) is what the speaker is talking about now, regardless of what has been said in discourse. The topic of a sentence has nothing to do with the discourse information structure and can be defined as the surface S-initial constituent. For example, in the following (27-28):

(27) Did you see the movies?
(28) What did John find?"
Halliday argues that the topics are *did* and *what*. In addition, the topics of the following sentences

(29) *It* is raining.

(30) *There* goes a bus.

(31) *Nobody* came to see you.

are the underlined *it*, *there*, and *nobody* respectively. However, it is clear that semantically sentences like (29-30) are about none of the sentence constituents (cf. Kuroda 1972; Lyons 1977:177); the speaker is just describing certain ongoing events. In (31), the alleged topic *nobody* is a non-existing entity. Thus, the speaker is attributing to a non-existing entity the predication that it came to see the addressee, which is absurd (see 2.4.1.2) . The Korean equivalents of (27-28) are the following:

(32) ne-nun yenghwa-iul po-assni?
    you TOP movie OM see did
    'Did you see the movies?'

(33) John-un mwues-ul po-assni?
    TOP what OM see did
    'What did John find?'

Despite the differences in word-order between the two languages, it seems that both English (27-28) and Korean (32-33) are equally about *you* and *John*.

To overcome problems like these, Chomsky (1965:221) proposed that "the topic of the sentence [be defined as] the left-most NP that is immediately dominated by S in the surface structure and that is, furthermore, a major category"
(Emphasis added). Chomsky's proposal makes correct predictions in (27-28) and (29): the topics of (27-28) are you and John, and there is no topic in (29). But how about (30-31)? Chomsky would say that the topics of these sentences are a bus and nobody. But these indefinite NP's cannot be considered topics for semantic reasons (see 2.4.1.2).

A similar proposal was made for Korean by I.S.Yang (1973), who argued that topic can be identified as the element in S-initial position in the surface form of the sentence. Yang contends that topic is not necessarily marked by the particle nun only, but that any S-initial NP's are topics, whether they are case-marked or delimiter-attached. He gives the following examples to support his claim.

\[(34)\]
\[
\begin{align*}
\text{a} & \quad \text{nulkun seytay-ka celmun seytay-lul ihayhayeyahanta.} \\
& \quad \text{old generation young generation OM understand} \\
& \quad \text{The older generation must understand the younger generation.}' \\
\text{b} & \quad \text{celmun seytay-lul, nulkun seytay-ka ihayhayeyahanta.} \\
\text{c} & \quad \text{celmun seytay-[\underbrace{\text{nul\_kun\_seytay-ka}}_{\text{only}}, \underbrace{\text{\text{\text{man}}}}_{\text{\text{\text{to}}}}('\text{\text{\text{only}}'}, '\text{\text{\text{also}}'}), \text{\text{\text{ihayhayeyahanta}}}.}
\end{align*}
\]

\[(35)\]
\[
\begin{align*}
\text{a} & \quad \text{John-i Mary-lul i khal-lo ccillessta.} \\
& \quad \text{SM OM this knife with stabbed} \\
& \quad \text{John stabbed Mary with this knife.'} \\
\text{b} & \quad \text{i khal-lo, John-i Mary-lul ccillessta.} \\
\text{c} & \quad \text{i khal-lo-[\underbrace{\text{\text{\text{man}}}}_{\text{\text{\text{to}}}}], \text{John-i Mary-lul ccillessta.}}
\end{align*}
\]
Yang suggests that the underlined expressions in (34-36) are the topics of their respective sentences.

The syntactic definition is on the right track in that it views the topic not as a totally discourse-conditioned objective entity but as a speaker-centered subjective notion. That is, the speaker decides what he wants to talk about and opens the sentence with it. It has also been confirmed across languages (Givón 1978; Li and Thompson 1976) that S-initial position is a highly favored way to express topichood. However, this does not mean that any constituent, or any constituent belonging to a major lexical category, is the topic if it occurs in surface S-initial position.

In the first place, this kind of definition is linguistically almost insignificant. What kinds of linguistic facts can be explained by assuming that anything in S-initial position is the topic? We can do with or without this kind of topic.

Second, this definition predicts that the topics of the following sentences

\[(37) \text{pī-ka o-nunkwuna!} \quad \text{rain SM come} \quad \text{'Rain is pouring.'}\]
should be the underlined NF's. On semantic grounds, however, sentences like (37–38) might be said to be topic-less (cf. Kuroda 1972; Lyons 1977).

Third, if topic is defined as any element which is "preposed to the front of the sentence" (I.S. Yang 1973:87), even the subject must be considered the topic—e.g. nulkun seytay in (34a), John in (35a) and Mary in (36a)—because we can say that the 'preposing rule' has vacuously applied to the subject. In 2.4.2, we will show that the subject cannot be considered the topic, nor can other case-marked NP's such as the underlined ones in (34b-36b).

Fourth, sentences (34b-36b) must be derived from (34a-36a) by a 'preposing' operation which is part of the Scrambling rule. This claim that Scrambling creates the topic, as in (34–36), is tantamount to a claim that topic is a purely surface structure concept created transformationally and determined after the Scrambling rule. However, there is evidence that topic is a notion determined prior to the Scrambling rule. Consider the following sentences.

(39a) namwu-lo Chelswu-nun chayksang-ul mantulessta.
    tree with TOP desk OM made
    'With wood, Chelswu made a desk.'

(40a) Yenghi-eykey Chelswu-nun ecey ton-ul cwuessta.
    to TOP yesterday money gave
    'To Yenghi, Chelswu gave money yesterday.'
It seems that the topics of (39a-42a) are Chelswu, even though their topichood may not be as clear as the same Chelswu in (39b-42b) following, which are assumed to underlie (39a-42a):

(39)a Chelswu-nun namwu-lo chayksang-ul mantulessta.
(40)a Chelswu-nun ecey Yenghi-eykey ton-ul cwuessta.
(41)a Chelswu-nun ecey Yenghi-lul mannassta.
(42)a Chelswu-nun, Yenghi-ka wassul ttay, wul-ko issessta. 5

The fact that the topics of the (a) and (b) sentences of (39-42) are equally Chelswu indicates that the Scrambling rule does not seriously affect the topichood of these sentences, and that topic is a status already determined prior to the Scrambling operation. 6

There is further evidence that the topic is not determinable solely from the surface syntactic structure. 7 Consider the following sentences.

(43)a Chelswu-nun nwukwu-lul cohahani?
   TOP whom OM like
   'Who does Chelswu like?'

b Yenghi-lul cohahay.
   OM like
   'Likes Yenghi.'
There is no doubt that the topic of (44b) is Chelswu, not the object Yenghi which happens to be in the surface S-initial position. The topic Chelswu must have been deleted due to obviousness. Thus, in this case the topic refers to a constituent of a structure which existed prior to the deletion rule.

In sum, the syntactic definition, which identifies the topic as the S-initial element of the surface structure, is not adequate. As we will argue later, topic is not a surface concept. This definition also neglects the semantic (2.4.1.2 and 3.2.2) and discoursal (3.2.3.2) aspects of the topic.

Closely related to the syntactic definition is what Magretta (1977) has proposed as a "functional definition." His proposal is similar to that of Chomsky (1965), but his major lexical category includes not only NP's but also PP's, adjectives, and even verbs. Magretta (1977:9-10) says:

Taking these observations about the expansion of the notion of topic beyond the category of nominal elements ... I propose to define the topic as a specific, single function associated with the initial position in a sentence. This function appears most clearly when constituents other than the grammatical subject ... that is, adverbs, prepositional phrases, adjectives, or verbs ... appear initially, but it is in principle a characteristic of all sentence-initial elements. However, what is common to all sentence-initial elements is not that they identify what the sentence is about but rather that they announce a frame of conditions under which the sentence is to be understood. (Emphasis added).

Note that Magretta dismisses even the axiomatic assumptions that topic is what the speaker is talking about and that its main function is to identify the entity about which a comment is to be made. The functional definition is not only faced
with the problems of the syntactic definition, but is even more problematic, because, as we will argue in 2.4.3, non-nominal elements such as PP's, adjectives, and verbs cannot be topics.

2.3 The Morphological Definition

In languages where topic is marked by some specific morphological markers, attempts were made to define topic in terms of its coding properties. Thus, Ree (1969), Kuno (1972, 1973) and D.W.Yang (1975) tried to identify the topic in Korean and Japanese as the S-initial NP followed by the particle nun and wa, respectively.

One of the puzzling aspects in this definition is the relationship between the NP-nun in S-initial position and the other S-medial NP-nun which has been called "contrastive focus" (D.W.Yang 1975). Kuno (1972, 1973), dealing with the corresponding wa forms in Japanese, seems to assume that only the S-initial NP-wa is the topic, whereas Kuroda (1969) has argued that the S-initial NP-wa is a "limit case" of the contrastive focal wa, implying that all NP-wa are a kind of topic. D.W.Yang (1975) has called the former S-initial NP-nun a "special case" of the latter contrastive focal nun, assuming that only the former is the topic, but later in Yang (1974) he seems to assume, like Kuroda (1969), that all NP-nun are topics.

It is not clear when and why topichood came to be associated with the form nun in Korean, but this morphological
definition seems to be based on the native speaker's intui-
tion that, when an NP-nun appears in S-initial position,
this NP is somehow felt to be something that the following
sentence is predicated about. It seems that the morpho-
logical definition is basically correct, but it is legitimate
to ask why only the NP-nun should be considered the topic,
and not case-marked NP's or delimiter-attached NP's in S-
initial position. That is, why should the particle nun be
considered a unique marker of topic? What kinds of differ-
ences, formal and semantic, distinguish the NP-nun from
other case-marked and delimiter-attached NP's, so that only
the former can be the topic? Until we find a reasonable ans-
ter to this question, the intuition-based morphological de-
finite cannot be said to be adequately justified.

The claim that topic can only be marked by the particle
nun runs into difficulties in cases where topichood seems to
be manifested not by a particle but by a syntactic position.
Consider the following sentences.

(44) Chelswu-nun pay -ka kophuta.
    TOP stomach SM hungry
    'For Chelswu, his stomach is empty.' ('Chelswu is hungry.')

(45) Chelswu-fo ('also') \ pay-ka kophuta.
    man ('only') \ ka ('just')
    ' Chelswu too, \ his stomach is empty(he is hungry).' \ Chelswu only,
    \ Just Chelswu,'

The topics of these sentences are intuitively felt to be
Chelswu, regardless of the particles used. It seems that
this intuition is also shared by others. For example, Li and Thompson (1976:488) observe similar intuitive feelings among Japanese speakers. This feeling seems to be due to the fact that both Chelswu's in (44-45) share certain similarities: they identify the entity to talk about and set up the framework in which the following proposition is to be understood or judged to be true or false.

Sentences like (44-45) raise the question whether topichood is really expressed only by the particle nun. Why cannot delimiters express topichood? Thus, it is not convincing enough to argue, just on the basis of intuition, that only the S-initial NP-nun is the topic.

To summarize, we have so far examined four definitions of topic: informational, syntactic, functional, and morphological. The first three were found to be unacceptable, because they lead to conclusions that are either incompatible with the distribution of the particle nun or do not fit the native speakers' intuition about 'topic' in Korean. The last one is not fully justified as it stands.

2.4 A Proposed Definition

2.4.1 Topic Properties

What element of the sentence can be considered the topic in Korean? Rather than replying intuitively with "the NP-nun" or "any S-initial NP," we will start by asking what the topic is at all. We take it as axiomatic that topic is what the speaker is talking about. This axiomatic feature of
topic, aboutness, is reflected by two essential kinds of characteristics, one syntactic and the other semantic. We will consider the syntactic property first.

2.4.1.1 As we will argue later in 3.1 and 3.2.1, topic is a semantic concept not created transformationally at a later stage of derivation (also see sentences (39-43)), but one present in underlying structure as an inherent semantic and syntactic component of the sentence. This means that the topic should be generated directly by the base rules of the language. At the same time, if the topic is what the speaker talks about, one of its logical consequences is that in linearization, i.e. when a proposition or message is realized syntactically as a sentence, the topic must precede what is actually said about it. Then, the topic should be the first syntactic category generated by the base rules. Since the base rules constitute the basic syntactic structure of the language, we might say that the topic occupies the first syntactic slot in the basic syntactic structure of the language.

We first note that the topic function is normally carried by the subject in English, as Mathesius (1928:64) has pointed out. We maintain that this is so because in English, which is a strong subject-prominent language, the subject occupies the first syntactic position of its basic sentence structure SVO, and so the topic function must typically be filled by the noun phrase in the subject slot. This is why
the topic function so often combines with the subject function in English.

In contrast, the basic syntactic structure of Korean is not strict SOV, but something like \([(\text{NP}) \, [\text{SOV}])\]. That is, Korean has an extra syntactic NP slot before the subject. There is enough evidence for the basicness of this syntactic structure, which is as important as the SOV pattern.

First, consider the following sentences.

(46)a (Chelswu (Yenghi-ka cohta))
  SM good
  'For Chelswu, Yenghi is good.' ('Chelswu likes Yenghi. ')

(47)a (na (kyewul-i chwupta))
  I winter SM cold
  'For me, the winter is cold.'

As has been argued by many (Li and Thompson 1976:468; I.S. Yang 1972), the underlined NP's Chelswu and na cannot be derived from an SOV pattern like those in (46b-47b):

(46)b *Yenghi-ka Chelswu-eykey cohta.
(47)b *kyewul-i na-eykey chwupta.

Then, there is no other way but for these NP's to occur in the pre-subject NP node in underlying structure, according to the basic formation rules of the language.

Second, consider the following sentences.

(48) mas -i issta.
  taste SM exist
  'It is delicious.'
naymsay-ka epsta.
smell SM non-existing
'It has no smell.'

(48-49) are semantically incomplete as they stand, because
the possessor or location of the taste or smell is not known.
(48-49) are acceptable when the possessor or location argu-
ment either has exophoric reference to a certain food the
speaker is eating in the here-and-now context, or has ana-
phoric reference to another NP in a discourse like the fol-
lowing (50):

(50) ku umsik ettay?
the food how
'What do you think about the food?'

Note that possible subject-predicate sources like (51a-b)
are not acceptable:

(51)a *ku umsik-uy mas -i issta.
the food of taste SM exist
'The taste of the food is good.'

b *mas-i ku umsik-ev issta.

Thus, the only syntactic position in which the possessor or
locative noun umsik can appear is again the pre-subject NP
position: [ku umsik [mas-i issta]].

Third, consider the following sentences.

(52) caki-ka Yenghi-lul ttaylyess-ci?
selt SM OM hit
'You hit Yenghi, didn't you?'
Reflexivization in Korean is subject to two conditions: the "subject condition" and "command condition" (N. McCawley 1976; D.W. Yang 1975). The former is a language-particular condition which states that the antecedent is normally the subject of the sentence, though there are some exceptions. The command condition is a rather universal constraint (Langacker 1969) to the effect that the antecedent which triggers reflexivization must precede and command the reflexive pronoun in a simplex sentence. Note that the subject condition is not relevant to the reflexivization in (52-53). The reflexivization phenomena in (52-53) are exceptional in that no antecedents which precede (and command) the reflexive form caki are found in these sentences. If the pre-subject NP is assumed to exist in Korean, this exceptional behavior of the reflexives are automatically explained by the command condition. That is, there was a pre-subject NP in underlying structure, which has either exophoric reference to the addressee ne or anaphoric reference to Chelswu. This pre-subject NP was deleted in (52) and (53b) after having triggered reflexivization in underlying structures like the following.
Otherwise, the above reflexivization phenomena cannot be explained except by an ad hoc exceptional rule.

Given that the basic syntactic structure of Korean is 
\[(NP (SOV)_{S})_S\], then our hypothesis that topic is the first constituent in the linearization process entails that the topic function falls on the pre-subject NP node in Korean. We maintain that this pre-subject NP node is the optimal syntactic position for expressing topichood in Korean. This means that the topic function (or the semantic concept of 'topic') has not been given independent syntactic status in English, but it has in Korean.

Actually the claim that the pre-subject NP's like Chelswu and na 'I' in (46a-47a) are topics is in agreement with native speakers' intuition as to the identity of the topic in these sentences. This claim is also supported by the typological findings that the pre-subject NP's are the prototype of topic in what Li and Thompson (1976) refer to as "topic-prominent languages."

Furthermore, language-internally, there seems to be no other syntactic function but that of topic which the pre-subject NP can perform. There have been suggestions that the pre-subject NP be considered another subject, either a "sentential subject" (Park 1973) or a "macro-subject" (I.S.Yang
1972). However, the pre-subject NP cannot be considered the "subject" in its usual sense. The 'subject' is the main syntactic function used to define the state of affairs represented by the predicate, so the subject must be an entity either directly involved (i.e. the agent) in or directly affected (i.e. the patient) by the action or state of affairs which the predicate denotes. Thus, when a proposition contains either an agent or a patient, such arguments as instrument, goal, or locative cannot in Korean be chosen as the subject over the agent or the patient. For example, suppose that the speaker has the following propositions:

(56) kassta (Chelswu, Pusan) [Goal] went

(57) patasssta (Chelswu, Yenghi, i cha) [Instrument] hit this car

(58) phiessta (kkoch, cengwen) [Locative] in-bloom flower garden

What can be chosen as subjects in syntacticizing these propositions is either Chelswu, Yenghi, or kkoch 'flower', because they are the entities directly involved in or affected by the action (e.g. 'hitting' in (57)) or the state of affair (e.g. 'being in bloom' in (58)). Note that none of the underlined arguments can be chosen as subjects.

Now, it should be noted that the pre-subject NP is often not an argument directly involved or affected by the action or state of affairs the predicate denotes—e.g., na 'I' and i umsik 'this food' in (47-48). Thus, they cannot be
considered 'subject' in the usual sense of the term. Furthermore, the term "sentential subject" is not desirable. "Subject" is a notion used in contrast with "predicate" and a 'sentence' cannot be considered a 'predicate' in its usual sense as a grammatical term. Li and Thompson (1976) and Tsao (1977) give some other arguments against the view that the pre-subject NP is a 'subject.' (Problems with the notion "macro-subject" will be discussed in fn. 13.)

If the pre-subject NP is viewed as the 'topic', however, these problems do not arise. The topic need not be an argument of the predicate, and a sentence can function perfectly well as a 'comment.' Thus, it seems that 'topic' is the best syntactic role in which the pre-subject NP can fit. Based on these arguments, we will assume that the pre-subject NP is the topic in Korean.

At the outset of this section we argued that topic carries the meaning "aboutness." The question now is how the pre-subject NP node comes to be associated with the 'aboutness' meaning? Note that the pre-subject NP in sentences like (46-47) can be followed not only by nun but also by other delimiters like man 'only' and to 'also'. Thus, we cannot say that nun is the only contributing factor.

However, whatever the particle that follows the pre-subject NP, it is clear that that NP must be placed in a position that dominates the following expression. There are several pieces of evidence that this is the correct constituent structure. First, the semantic scope of the pre-
subject NP—e.g. Chelswu in (46a)—extends over the whole of the following sentence (also see 2.4.2.3). Second, a pause normally falls between the pre-subject NP and the following sentence. Third, the pre-subject NP and the following subject do not form a constituency (cf. Shibatani and Cotton 1977). Thus, the underlying structures of sentences like (46a-47a) must look like the following:

(59)  

Thus, the pre-subject NP which is most optimal for expressing topichood, occupies the "higher NP" position which dominates the following sentence S'.

We claim that the intuitive feeling of 'aboutness' (or 'topicness') we have from the pre-subject NP comes from its higher syntactic position. The 'higher NP' position allows the pre-subject NP to have the function of setting up the framework within which the following sentence holds true or false, or restricting the domain in which the following sentence is to be judged. It seems that this function gives rise to the 'aboutness' feeling we have from the pre-subject NP.

In conclusion, the syntactic aspect of the topic is that it occupies the 'higher NP' position which is also the first slot of the basic syntactic structure of Korean.
2.4.1.2 The other property of the topic which is also derivative from its axiomatic 'aboutness' feature is that topic must be semantically definite. Since it is inconceivable for the speaker to talk about something that he himself does not know, the topic must be something definite. This is again a logical consequence of the 'aboutness' feature of the topic. It might be argued that this is just a logical conclusion and that logical necessity does not always go parallel with linguistic phenomena. But in this case there is evidence that this logical necessity between topic and definiteness is actually reflected as such in language.

It would probably be such expressions as -ey tayhaye malhantamyen or -lo malhal kes kathumyen (, both of which can be translated to 'if we talk about' in English) that show in the clearest and most unmistakable way that something is a topic in Korean. We will call them "topic announcers" for convenience. If an element turns out to be incompatible, i.e., cannot co-occur, with topic announcers, it cannot be considered a possible topic. This is an entirely semantic consideration having nothing to do with syntactic position. However, if an element can co-occur with topic announcers, it can be a topic. 14

What is significant in this context is that indefinite specifics cannot co-occur with topic announcers. What we mean by 'indefinite specifics' is that the speaker is committed to the existence of some individual, but the actual identity of the individual is left unspecified, because the
speaker himself is not sure of its identity. The fact that indefinite specifics cannot co-occur with topic announcers is interpreted as showing that semantically indefinite noun phrases cannot be topics or that indefiniteness is incompatible with topicness. Consider

(60) *nwukwu-ey tayhaye malhantamyen, ku-nun cwukessta.
    who if-we-talk-about he TOP died
    'If we talk about who, he died.'

(61) *nwukwu-ey tayhaye malhamyen, ku-nun ecey wasanni?
    who if-we-talk-about he yesterday came
    'If we talk about someone, did he come yesterday?'
    'If we talk about who, did he come yesterday?'

(62) *mwues-ey tayhaye malhamyen, ne-nun kukes-ul ani?
    what if-we-talk-about you TOP that OM know
    'If we talk about what/something do you know it?'

(63) *etten salam-ey tayhaye malhantamyen, ku-nun wass-
    some man if-we-talk-about he TOP came
    ta/-ni?
    'If we talk about a man, (did) he (come) came.(?)'

(64) *sey salam-ey tayhaye malhamyen, kutul-un chakhata.
    three men if-we-talk-about they TOP good
    'If we talk about three men, they are good.'

(60-64) show that indefinites such as nwukwu 'who' (or 'someone') cannot co-occur with topic announcers. This fact clearly proves that the topic cannot be indefinite semantically, and that the logical correspondence of topichood and definiteness holds true in language, too. Note that when the NP is definite, it can co-occur freely with topic announcers, as in (65):

(65) Chelswu-ey tayhaye malhantamyen, ku-nun cwukessta.
    'If we talk about Chelswu, he died.'
2.4.2 Case Particles and the Particle nun.

Since the topic must occupy the 'higher NP' position and must be definite, we can, conversely, use these topic properties as test-frames to determine whether or not a certain element of the sentence is the topic. This section will examine the NP-nun and case-marked NP's (including the subject marker ka) in S-initial position, the constituents which I.S.Yang (1973) argued are topics. To put the conclusion first, case-marked NP's in S-initial position are not topics. To show this point we will examine three syntactic phenomena (reflexivization, sentence-embedding, and discourse function) and co-occurrence with indefinites.

2.4.2.1 As mentioned in 2.4.1.1, reflexivization is subject to the command condition—the antecedent must precede and command the reflexive pronoun in a simplex sentence. Thus, if the antecedent commands the reflexive, forward reflexivization is always acceptable in a simplex sentence, but its backward application is not, as in the following (66a-b):

(66)a Chelswu-ka cakicasin-i-ul haktayhako issta.
   SM self OM mistreat -ing
   'Chelswu is mistreating the self.'

b *Cakicasin-i Chelswu-lul haktayhako issta.
   'The self is mistreating Chelswu.'

Now, consider the following sentences: 16

(67)a *Chelswu-lul cakicasin-i haktayhako issta.
   OM self SM mistreat -ing
   'Chelswu, the self is mistreating him.'
Reflexivization in both (67a-b) applies in a forward direction. The fact that NP-lul in (67a) cannot reflexivize the following subject indicates that when the object NP is fronted, it does not command the subject. On the contrary, the fact that NP-nun in (67b) can trigger reflexivization shows that NP-nun, whether it is fronted (in a transformational analysis of topic) or not, commands the following subject. Thus, the surface constituent structures of (67a-b) must look like the following (68a-b):

Therefore, NP-nun in S-initial position occupies the higher NP position but NP-lul does not, and only NP-nun is the topic.

As for the relationship between the subject NP-ka and NP-nun, consider the following pair of sentences:
Reflexivizations in (69a-b) meet the 'subject' condition and both apply forward. Since both (82a) and (82b) are acceptable, it is not possible to say that only NP-nun in (69b) occupies the higher NP position. Note that it might be argued that (69a) has the following constituent structure:

However, (70) does not seem to be the correct constituent structure of (69a). Consider

If (70) were the correct constituent structure of (69a), there is no reason why (71a) should not be acceptable, as (71b) is. The only way to explain this difference in acceptability between (71a) and (71b) is to assume that the subject NP-ka does not occupy the higher NP position but that NP-nun
does. Therefore, the subject NP-ka in S-initial position is not the topic, but NP-nun is.

Other case-marked NP's in S-initial position do not pass this same reflexivization test for topics. Consider

(72)a *Chelswu_{i}-eykey caki_{i}-ka (cakicasin_{i}-eykey) mayngse-lul hayssta.

'To Chelswu, he swore to himself.'

b Chelswu_{i}-nun caki_{i}-ka cakicasin_{i}-eykey mayngse-lul hayssta.

'As for Chelswu, he swore to himself.'

(73)a *Chelswu_{i}-wa caki_{i}-ka (cakicasin_{i}-kwa) maumsok-eyse ssawuko issta.

'With Chelswu, he is fighting with himself mentally.'

b Chelswu_{i}-nun (caki_{i}-ka) cakicasin_{i}-kwa maumsok-eyse ssawuko issta.

'As for Chelswu, he is fighting with himself mentally.'

The case-marked NP's such as NP-eykey (the dative) and NP-wa (the comitative) in the (a) sentences above do not reflexivize the following subject, whereas the NP-nun's in the (b) sentences do. This fact again seems to support the conclusion that case-marked NP's do not occupy the higher NP position, but that NP-nun does.

Based on the reflexivization evidence, we conclude that
case-marked NP's in S-initial position are not topics, whereas NP-nun is.

2.4.2.2 A difference between case-marked NP's and NP-nun is also observed in complex sentences. Consider

(74)a John-ul seltukha-ki swiwessta.  
   OM persuade to easy  
   'To persuade John was easy.'

b John-un seltukha-ki swiwessta.  
   TOP  
   'As for John, he was easy to persuade.'

On the surface it is not clear whether NP-lul and NP-nun are different in semantic scope, but note that (74a) cannot be followed by (75a), but (74b) can be by (75b):

(75)a *(John-ul seltukha-ki swiwessta) kuliko Mary-lul and
   OM elyewessta.
   difficult
   'To persuade John was easy, and Mary was difficult.'

b (John-un seltukha-ki swiwessta) kuliko Mary-nun
   TOP elyewessta.
   'As for John, he was easy to persuade, and for Mary, it was difficult.'

The fact that Mary-lul of (75a) cannot be associated with the higher verb elyewessta 'difficult' shows that John-ul of (74a) is associated with the embedded verb seltukhata 'persuade' rather than the higher verb swiwessta 'easy'. In contrast, Mary-nun can be associated with the higher verb elyewessta in (75b), which indicates that John-un of (74b) is
associated with the higher verb *swiwssta* rather than the embedded verb *seltukhata*. These facts can be explained by giving (74a-b) the following constituent structures (76a-b) respectively:

\[
(76)a \quad (76)b
\]

\[
\begin{aligned}
S & \quad S' \\
NP & \quad NP \\
\text{swiwssta} & \quad \text{John-un} \\
NP & \quad NP \\
VP & \quad VP \\
\Delta \text{John-ul} & \quad \Delta \text{seltukhata}
\end{aligned}
\]

A comparison between (76a) and (76b) shows that only NP-nun, not case-marked NP's, occupies the higher NP position.

A similar phenomenon is observed between the subject NP-ka and NP-nun. Consider

\[
(77)a \quad \text{John-i uycang-ulo toynun kes-i yulihayessta.} \\
\quad \text{SM chairman as become to SM better} \\
\quad \text{"That John is elected the chairman was better."}
\]

b \text{John-un uycang-ulo toynun kes-i yulihayessta.} \\
\text{TCP} \quad \text{"As for John, that he is elected the chairman was better."}

Note that (77a) cannot be followed by (78a), but (77b) can be by (78b).

\[
(78)a \quad \text{*(John-i uycang-ulo toynun kes-i yulihayessta),} \\
\quad \text{kuliko Mary-ka pwullihayessta.} \\
\quad \text{and SM unfavorable}
\]
'That John is elected the chairman was better, and Mary was unfavorable.'

(78)b (John-un uycang-ulo toynun kes-i yulihayessta), kuliko Mary-nun pwullihayessta, and TOP unfavorable

'As for John, that he is elected the chairman was better, and, for Mary, it was unfavorable.'

Again, the unacceptability and acceptability of (78a-b) can be explained by positing the constituent structures (79a-b):

(79)a

(79)b

(79a-b) show that only NP-nun, not NP-ka, can qualify as topic, if the topic occupies the higher NP position.

2.4.2.3 Another piece of evidence for the higher NP position of the topic is that the functional role of the topic is not limited to a single sentence but is constant across sentences. Put the other way around, we might say that this discourse function of topic is derivative of its 'higher NP' position. The discourse function of topic is clearest in the so-called topic chain (Tsao 1977), in which the semantic domain of a single topic extends over more than one sentence:

(80) i mincok-un [han-tongan yekiceki hutickeyese] this race TOP for-a-while here-there scattered l
'As for this race, they lived scattered here and there, hunted animals, cultivated the land; then their power became strong, and they built a nation near the Taytong river.'

In (80) the topic \textit{i mincok} 'this race' is a common denominator of the following six sentences, and is essential to their semantic interpretations. The semantic domain of the topic in this discourse might be represented as follows:

\begin{center}
\begin{tikzpicture}
  \node (S) at (0,0) {S};
  \node (NP) at (-2,1) {NP \textit{i mincok}};
  \node (S1) at (-1,2) {S_1};
  \node (S2) at (0,2) {S_2};
  \node (S3) at (1,2) {S_3};
  \node (S4) at (2,2) {S'_4};
  \node (S5) at (3,2) {S_5};
  \node (S6) at (4,2) {S'_6};

  \draw (NP) -- (S);%\node (NP) at (-2,1) {NP \textit{i mincok}};
  \node (S1) at (-1,2) {S_1};
  \node (S2) at (0,2) {S_2};
  \node (S3) at (1,2) {S_3};
  \node (S4) at (2,2) {S'_4};
  \node (S5) at (3,2) {S_5};
  \node (S6) at (4,2) {S'_6};

  \end{tikzpicture}
\end{center}

where the topic occupies the higher NP node outside the individual sentences.

Again we might use this discourse function of topic as a test-frame to determine whether NP-nun and case-marked NP's are topics or not. Consider

\begin{itemize}
\item\textbf{(82)a} \textit{Chelswu-eykey Yenghi-ka ton-\text{ul} cwuess-una, pat-} to \textit{SM money gave but receiv} ci \textit{anhassta}.
\item\textbf{(82)b} \textit{Chelswu-nun Yenghi-ka ton-\text{ul} cwuess-una, pat-}\textit{TOP}
\end{itemize}
ci anhassta.
'As for Chelswu, Yenghi gave him money, but he did not receive it.'

(83)a  i  pyek-ey ttay-ka mwutese, ttakass-una,
this wall at stain SM is cleaned but
kkaykkusha-ci anhta.
clean not
'At this wall, stains are there, and so I cleaned it, but it is still not clean.'

b  i  pyek-un ttay-ka mwutese, ttakass-una, kkaykkusha-
TOP ha-ci anhta.
'As for this wall, stains are there, and so I cleaned it, but it is still not clean.'

The underlined NP's in (82a-83a) are dative and locative, respectively. These fronted case-marked NP's are all sentence-bound, and their semantic scopes cannot exceed a sentence. This becomes clear from the fact that they cannot be associated with the clauses following the first. Consider

(84) *Chelswu-eykey [pat-ci anhassta]$_{S_2}$
'to receive not 'To Chelswu, he did not receive it.'

(85)a  i  pyek-ey [ttakass-una]$_{S_2}$
this wall at cleaned 'At this wall, I cleaned, but'

b  i  pyek-ey [kkaykkusha-ci anhta]$_{S_3}$
clean not
'At this wall, it is still not clean.'

Thus, the functional role which a case-marked NP has is usually limited to the confines of a single sentence.

In contrast, the semantic domain of NP-nun in (82b-83b) extends equally over the following clauses. Compare (84-85)
with (86-87):

(86) Chelswu-nun [pat-ci anhassta]$_{S_2}$
    'As for Chelswu, he did not receive it.'

(87)a i pyek-un [ttakass-una]$_{S_2}$
    'As for this wall, I cleaned it, but'

    b i pyek-un [kkaykkusha-ci anhta]$_{S_3}$
    'As for this wall, it is still not clean.'

Unlike case-marked NP's, NP-nun can be associated with all of the following clauses.

To explain this difference between (84-85) and (86-87) in acceptability, we posit the following underlying structures (88a-b) for (83a-b), respectively:

(88)a

Note that the NP-nun occupies the higher NP node in (88b), but case-marked NP's do not, as in (88a).

2.4.2.4 In this section we will examine the topichood of
the NP-nun and case-marked NP's from the semantic point of view, that is, the view that the topic must be definite. Consider the following sentences, where the underlined NP's are all indefinite specifics.

(89)a \textit{nwukwu(nka)-ka wassta.} \textit{someone SM came 'Someone came.'}

(90)a \textit{mwues-ul ne-nun mekessni?} \textit{what OM you TOP ate 'What did you eat?'}

(91)a \textit{etten yeca-iul na-nun salanghako issta.} \textit{certain woman OM I TOP love 'A certain woman, I love.'}

(92)a \textit{nwukwu-eykey Chelswu-ka chayk-ul cwuessni?} \textit{who to SM book OM gave 'To whom did Chelswu give a book?'}

(93)a \textit{eti-eyse Yenghi-ka wassci?} \textit{where from SM came 'Where is Yenghi from?'}

(89a-93a) show that case markers such as \textit{ka}, \textit{lul}, \textit{eykey}, etc. can co-occur with indefinite specifics.

In contrast, the NP-nun in S-initial position cannot co-occur with indefinite specifics in normal contexts. Compare the (a) sentences with the following (b) sentences:

(89)b *\textit{nwukwu(nka)-nun wassta.}

(90)b *\textit{mwues-un ne-nun mekessni?}

(91)b *\textit{etten yeca-nun na-nun salanghako issta.}

(92)b *\textit{nwukwu-nun Chelswu-ka chayk-ul cwuessni?}

(93)b *\textit{eti(eyse)-nun Yenghi-ka wassci?}

The (a) sentences and the (b) sentences of (89-93) show that
case markers and the particle nun differ from each other in their occurrence with indefinites.

In 2.4.1.2 we argued that a topic announcer like -e tayhave malhantamyen 'if we talk about' cannot co-occur with indefinite specifics, because the 'aboutness' property of the topic is semantically incompatible with indefiniteness. Now, the fact that the particle nun cannot co-occur with indefinites may be interpreted as showing that nun has the semantic property of 'aboutness.' In contrast, the fact that case particles can freely co-occur with indefinite NP's indicates that they do not have the 'aboutness' property. Otherwise, it would be difficult to explain why only nun, but not case particles, cannot co-occur with indefinites. Thus, we hypothesize that the particle nun, but not the case particles, has the semantic property of 'aboutness.' That is, only nun is inherently related to topicness.

One further fact that seems to support this hypothesis is the lexical semantic feature of the particle nun. In 3.2.5 we will argue that the topic function of the particle nun is derivative of its use as a delimiter, i.e. a constituent focus marker with contrastive meaning. The basic meaning of the latter contrastive focal nun is that the speaker is committing himself to the validity of the assertion with regard to the entity preceding it, but not to other possible candidates (cf. Kuroda 1969:145). This meaning is very close to the function of the topic, i.e. picking out an entity out of other possible candidates for what is to be predicated.
Thus, it seems that the 'aboutness' feeling we get intuitively from the S-initial NP-nun comes in part from the lexical meaning of the particle nun itself.

To sum up, the NP-nun is clearly the topic, because the NP is always definite and the particle itself has the semantic property of 'aboutness.' In addition, NP-nun is always the 'higher NP' which occupies the optimal topic position. In contrast, case-marked NP's in S-initial position are not topics, whether definite or indefinite, because they do not appear in the higher NP position and because case-markers themselves do not imply 'aboutness.'

2.4.3 Non-nominal Elements

As for the topichood of non-nominal and non-referential elements such as adjectives, adverbs, particles, and verbs, it suffices to point out that they cannot co-occur with topic announcers:

(94) *alumtawun-ey tayhaye malhantamyen, ku kkoch-i beautiful if-we-talk-about the folower
     kulehta.
     so
     'If we talk about 'beautiful', the flower is so.'

(95) *ppalli-ey tayhaye malhantamyen, ku-ka kulehkey fast if-we-talk-about he SM so
     ttwinta.
     run
     'If we talk about 'fast', he can run like that.'

(96) *kulaysse-ey tayhaye malhantamyen, kulayse ku-ka so if-we-talk-about so he SM
wassta.
came
'If we talk about 'so', so he came.'

(97) *salang:hako issta-ey tayhaye malhantamyen, Chel-
love if-we-talk-about
swu-ka Yenghi-lul kulehkey hako issta.
SM  OM so do
'If we talk about 'love', Chelswu is doing so
with Yenghi.'

Therefore, contrary to what Magretta (1977) claimed, these
S-initial non-nominal elements cannot be topics.

2.4.4 Delimiters

It is well-known that the nun of topic marking is for­
mally identical to what has been known as the nun of con­
trastive focus, which is a member of the larger natural class
called "delimiters" (I.S.Yang 1972). The delimiters also in­
clude other particles such as man 'only', to 'also', kkaci
'even', etc. The relationship between the topic marking nun
(which we will represent as nun) and the delimiter nun (re­
presented as nūn) is a puzzling aspect of what is apparently
a single formative nun. In a similar fashion, the particle
ka also has dual functions: one (kā) for marking subjects, and
the other (ká) for expressing a "determinative focus"(see
3.2.5).

Now, it is not only NP-nūn and case-marked NP's but
also NP's followed by a delimiter (including nūn and kā)
that can appear in S-initial position. Thus, the question is
whether these delimiter-attached NP's in S-initial position
are also topics or are simple fronted NP's like case-marked NP's.

2.4.4.1 To determine the topichood of delimiter-attached NP's in S-initial position, we will apply the syntactic and semantic tests for topics discussed in 2.4.1 and 2.4.2. In this sub-section we will examine their syntactic properties.

First, consider the following so-called "double subject constructions" (henceforth DSC), which are actually "topic-subject constructions" (see 2.4.1.1). Note that the first NP can be followed by a delimiter, as in (98a):

(98)a Chelswu-to[-man,-kkaci,-nún,-ká] salam-i cohta.

\[\text{man SM good}\]

'Also [only, even, at least, just] Chelswu, he is good in personality.'

b Chelswu-

\[\text{TOP}\]

nún salam-i cohta.

'As for Chelswu, he is good in personality.'

The NP-nún in (98b) cannot be derived from a subject-predicate sentence (Sohn 1980); similarly, the delimiter-attached NP in (98a) does not have any derivational source of the subject-predicate type. Thus, the underlined NP Chelswu in (98a) must be generated directly as such in underlying structure and so it is syntactically the topic since it occupies the 'higher NP' position.

Consider the following non-DSC sentences.

(99)a pwulkoki-to[-man,-kkaci] Cosenok-ulo kaya tway.

\[\text{BBQ to went must}\]

'For the barbeque too, we have to go to Cosenok Restaurant.'
(99b) \textit{pwulkoki-nun Cosenok-ul kaya tway.} \\
'As for the barbecue, we have to go to Cosenok Restaurant.'

(100a) \textit{nakssi-to[-man,-kkaci] Kim-sensayngnim-ul mosica.} \\
*fishing Mr. OM invite* \\
'For fishing too, let us invite Mr. Kim.'

b \textit{nakssi-nun Kim-sensayngnim-ul mosica.} \\
'As for fishing, let's invite Mr. Kim.'

(101a) \textit{kwankwang-to[-man,-kkaci] Ceycwuto-lo} \\
*sightseeing cenghayessta. decided* \\
'For the sightseeing too, we decided on Ceycwu Island.'

b \textit{kwankwang-un Ceycwuto-lo cenghayessta.} \\
'As for the sightseeing, we decided on Ceycwu Island.'

Just like (98b), (99b-101b) also lack any derivational sources of the subject-predicate pattern (see 3.2.1). Thus, the NP-nun's must be generated directly as such in underlying structure and they occupy the 'higher NP' position. Similarly, (99a-101a) lack plausible derivational sources; therefore, the underlined delimiter-attached NP's occupy the 'higher NP' position in underlying structure. Thus, syntactically they are topics.

Second, just like the NP-nun, delimiter-attached NP's in the DSC all allow reflexivization:

(102a) \textit{Chelswu_i-to[-man,-kkaci,-ka] cakicasin_i-i} \\
*miwessta. hateful* \\
'Chelswu_i too, the self_i was hateful.'
(102)b Chelswuᵢ-nun cakicasinᵢ-i miwaressta.
'As for Chelswuᵢ, the selfᵢ was hateful.'

A similar phenomenon is observed in the following non-DSC sentence:

(103)a Chelswuᵢ-to[-man,-kkaci,-ka] cakicasinᵢ-i haktayhako isssta.
self SM
haktayhako isssta.
mistreat -ing
'Chelswuᵢ too, the selfᵢ is mistreating (him).'

b Chelswuᵢ-nun cakicasinᵢ-i haktayhako isssta.
'As for Chelswuᵢ, the selfᵢ is mistreating (him).'

The fact that the underlined delimiter-attached NP's all allow the reflexivization of the following subject seems to indicate that they occupy the 'higher NP' topic position.

Third, delimiter-attached NP's also exhibit the same behavior as the NP-nun in sentence-embedding. Consider

(104)a Chelswuᵢ-lul seltukha-ki swiwessta.
OM persuade to easy
'To persuade Chelswu was easy.'

b Chelswuᵢ-to[-man,-kkaci] seltukha-ki swiwessta.
'Chelswu too, to persuade him was easy.'

Note that (104a) cannot be followed by (105a) but that (104b) can be by (105b):

(105)a *(Chelswuᵢ-lul seltukha-ki swiwessta) kuliko and
Maryᵢ-lul elyewessta.
OM difficult
'To persuade Chelswu was easy and to persuade Mary, difficult.'
(105)b (Chelswu-to seltukha-si swiwessta) kuliko
Yenghi-to swiwessta.
easy
'Chelswu was easy to persuade, and so was
Yenghi.'

The unacceptability of (105a) and acceptability of (105b)
show that Chelswu-to is a higher NP, but Chelswu-iul is not.
Thus, just like the NP-nun, delimiter-attached NP's are
higher topics.

Fourth, delimiter-attached NP's also have the dis-
course function typical of the topic. Consider

(106) wuli kimchi-to[-man,-kkaci,-ka] [mas-i coh-ko]$_S$
taste good
our kimchi
[conglyu-to manh-ko]$_S$ [yangnyem-to yelekaci-ita]$_S$
kind also many 2 seasoning various
'Our kimchi too, its taste is good, its kinds
are numerous, and the seasonings are varied.'

(107) Chelswu-to[-man,-kkaci,-ka] [son-ul tachye]$_S$
hand hurt
[yak-ul chacass-una]$_S$ [amwu kes-to chac-ci mos
medicine looked-for 2 anything found not
ha-ca]$_S$ [pyengwen-ulo kassta]$_S$
so 3 hospital to went
'Also Chelswu got his hand hurt, looked for
medicine, but could not find anything, and
so went to hospital.'

(106) is a DSC sentence and (107) is a normal subject-predi-
cate sentence. Note that the semantic domains of the under-
lined delimiter-attached NP's extend over the following sen-
tences. This fact indicates that they also occupy the
'higher NP' position.
Finally, consider the following sentence.

(108) Chelswu-to swukcey-man hayssta.
also homework only did
'Also Chelswu did only homework.'

As Kuroda (1969) has argued for a Japanese sentence similar to (108), (108) implies that Chelswu—who may be the best student in the class—was expected to do more than merely the assigned piece of homework, but he only did the homework. Compare (108) with the following (109):

(109) swukcey-man Chelswu-to hayssta.
homework only also did
'Only homework, even Chelswu did.'

If (109) is assumed to have been derived from (108) by fronting the object swukcey-man, then, as Kuroda points out, the sentence is "mysteriously uninterpretable." However, if (109) is understood as (110)

(110) swukcey-man-un Chelswu-to hayssta.

the sentence becomes interpretable and the implication is that Chelswu—who may be the poorest student in the class—was expected to be among those who would not even finish the assigned homework, but, contrary to expectation, even he did the homework. Therefore, (108) and (110) are quite different in meaning.

Since (108) and (110) are so distinct in meaning, any transformational solution deriving the latter from the former is difficult to justify. It might be suggested that the
meaning difference mentioned above be taken care of by some surface semantic interpretation rules, but it is easy to see how complicated these rules would be because different combinations of delimiters produce different meaning changes. Therefore, the first delimiter-attached NP of (110) must be generated as such directly in underlying structure and so it must occupy the 'higher NP' position.

We have so far given five pieces of evidence, syntactic and semantic, which support the conclusion that S-initial delimiter-attached NP's occupy the higher NP position and so, just like NP-nun, they should be considered topics syntactically.

2.4.4.2 Let us now examine S-initial delimiter-attached NP's from the semantic point of view. As presupposition-bearing elements, delimiters can be used felicitously if and only if a certain specific presupposition compatible with its semantic content holds in the communicative situation. The required presupposition is usually provided by the prior discourse. 19

If the required presupposition holds, a delimiter can co-occur with either definites or indefinites. All of the examples (98-110) given in 2.4.4.1 have definite NP's. For the indefinite NP's, consider the following sentences.

(111) nwu-ka pay -ka aphyu-nyi?
      who stomach SM aching
      'Who has a stomachache?'
The fact that delimiters are compatible with indefinites, as in (111-114), seems to indicate that, like case markers, delimiters do not imply the semantic feature of 'aboutness.' If they did (like the topic announcer -ey tay-have malhantamyen or the particle nún), they could not co-occur with indefinites. Therefore, we hypothesize that delimiters (including nún) do not imply 'aboutness.'

It seems that S-initial delimiter-attached NP's fall somewhere between NP-nún and case-marked NP's in terms of topichood. We have found that NP-nún possesses both the syntactic and semantic properties of the topic, whereas case-marked NP's have neither of them. Delimiter-attached NP's in S-initial position possess the syntactic property of topic, because they occupy the higher NP position, but delimiters themselves do not imply 'aboutness.' It seems that this
is why the topichood of delimiter-attached NP's is not as clear as that of NP-nun, but is better than that of case-marked NP's.

2.4.5 Topic: A Continuum

As with other linguistic concepts such as "grammaticalcy," "word," "sentence," etc., the notion of 'topic' is conceptually primitive (i.e. intuitive) and theoretically derivative. For example, the goal of grammar in Chomsky (1957, 1965) is to define the native speaker's intuitions about 'grammaticality' (i.e. the grammar should generate all and only the grammatical sentences of the language), but the grammar makes use of the native speaker's intuition as to what is, and is not, grammatical. There are cases in the middle in which both intuition and analysis fail and fuzziness arises, where we cannot tell whether a sentence is grammatical or not.

The concept of topic seems to be just the same. The specification of topic is the goal of our analysis, but the analysis is based on the native speaker's intuitive knowledge of what is, and is not, topic. And, as with the notion of 'grammaticality', there are cases in which both intuition and analysis fail and fuzziness arises. In this area the topichood of a noun phrase becomes unclear. We argue that topic is not an all-or-none concept but a gradient phenomenon along a continuum of "topicality," which we define as "the relative strength of aboutness." Some NP's are
clearly topical and some are clearly non-topical, but in the middle there is a fuzzy area, typical examples of which are sentences like (111-114).

We propose that the degree of 'topicality' be determined by the interaction of the following three factors that contribute to 'aboutness'. The first factor is the syntactic 'higher NP position'. By providing the framework in which the following sentence is to be understood, this position gives rise to the 'aboutness' feeling (see 2.4.1.1). The second factor is the 'presence or absence of the particle or phrase that implies 'aboutness''. We argued that only the particle nun has this property, if we exclude from consideration such periphrastic topic announcers as -ev tay-haye malhantamyen. Third, it seems that 'definiteness' also contributes to 'aboutness'. As we will argue in 3.2.2, if a proposition to be conveyed contains a definite entity, there is a tendency to present the proposition linguistically with regard to (or, about) the definite entity. Thus, if an NP is definite, the NP is easy to talk about. Conversely, if an NP is indefinite, it is very difficult to interpret the NP as something about which the speaker is talking. In this sense, definiteness also contributes to the 'aboutness' feeling.

Based on the interaction of these three topic properties, the following topicality continuum is proposed.

(I) The NP-nun: unmarked topic
(115) **Chelswu-nun sengkyek-i myenglanghata.**

'As for Chelswu, his character is cheerful.'

(116) **ku haksayng-un kongpwu-lul cal hanta.**

'The student studies very well.'

NP-ᵣᵣ shows the strongest topicality. Its topicalhood is clearest. This is so because all three topic conditions are met: it occupies the higher NP position; the particle nᵣᵣ implies 'aboutness'; and the NP is always definite. The topicalhood of Chelswu in (115) is a little bit clearer than that of ku haksayng 'the student' in (116), because its higher NP position is evident on the surface.

We will call NP-ᵣᵣ the "unmarked topic" because, compared with the following "focus topic," NP-nᵣᵣ occurs freely as topic in discourse-initial (or, isolated) sentences. Moreover, NP-ᵣᵣ usually represents non-focal information within a discourse and achieves an 'unmarked' discourse transition (see 3.2.3.2).

(II) **Delimiter-attached definite NP: focus topic**

(117) **Chelswu-to[-man,-nᵣᵣ,-ka] sengkyek-i myenglanghata.**

'Chelswu too, his character is cheerful.'

(118) **ku haksayng-to[-man,-nᵣᵣ,-ka] kongpwu-lul cal hanta.**

'Also the student studies very well.'

Delimiter-attached definite NP's in S-initial position are relatively clear in topicalhood, because they occupy the higher NP position and are definite even though delimiters
themselves do not imply 'aboutness'. The NP Chelswu-to in (117) is clearer in topicality than ku haksayng-to in (118) because its higher NP position is shown as such on the surface.

Delimiter-attached NP's are highly context-dependent and restricted in occurrence. For example, they cannot occur as topics in discourse-initial (or, isolated) sentences. Consider the following sentence.

(119) * #Chelswu-to meli-ka aphuta.
   also head SM aching 'Chelswu too, he has a headache.'

Delimiter-attached NP's can occur only when some textual presupposition prevails in the discourse and the focal information is taken as the topic, as in the following sentences:

(120) Speaker A: #Chelswu-nun meli-ka aphuta.
   TOP head SM aching 'Chelswu has a headache.'

(121)a Speaker B: kulay? Yenghi-to meli-ka aphuntey.
   really also 'Really? Yenghi too, she has a headache.'

   only 'No. Only Yenghi has a headache.'

In this sense, delimiter-attached NP's are "marked" topics involving information focus.

(III) Delimiter-attached indefinite NP

(122) etten chengnyen-to[-man,-nun,-i] sengkyek-i
certain young man also character SM
myengkapghata.
cheerful
'A certain young man too, he is cheerful in character.'

(123) etten haksayng-to[-man,-nún,-ka] kongpwoo-lul cal
certain student also study 0M well
hanta.
do
'Also a student studies very well.'

The topic status of the underlined NP's in (122-123) is dubious. The NP's are indefinite and the delimiters do not have the semantic feature of 'aboutness'. The NP in (122) seems to be a little better than that of (123), and this again seems to be related to the fact that the NP in (122) occupies the higher NP position on the surface.

(IV) Case-marked NP's and Non-nominal elements

(124) etten kay-ka ttwie-kanta.
certain dog SM run go
'A dog is running.'

(125) Chelswu-ka yelsimhi kongpwoo-ko issta.
SM hard study Progressive
'Chelswu is studying hard.'

(126) alumtapkey, kwukhwa-ka phiessta.
beautifully chrysanthemum in bloom
'Beautifully, the chrysanthemum came into bloom.'

The case-marked NP's and non-nominal elements are not topics, because syntactically they are not in the higher NP position, and semantically case markers do not imply 'aboutness'.

As mentioned in 2.2, topic status can be affected a little bit by later grammatical processes. For example, if a non-topical element is shifted to before the topic by such
later transformational rules as Scrambling, the topicality of the topic, i.e., Yenghi in the following (127a), becomes a little obscured.

(127)a Chelswu-lul Yenghi-nun maywu salanghanta.
     OM TOP very love
     'Chelswu, Yenghi loves very much.'

The topicalhood of Yenghi is less clear in (127a) than in the corresponding unfronted version (127b):

(127)b Yenghi-nun Chelswu-lul maywu salanghanta.
     TOP OM
     'Yenghi loves Chelswu very much.'

Native speakers may differ as to the cut-off point at which an NP is no longer felt to be a topic. To me, it seems that NP-nun in (I) and delimiter-attached definite NP's in (II) can be topics in Korean. 21

The above discussion does not exhaust all the factors involved in the topicality continuum. In 3.2, we will show that, even among NP-nun forms, some are stronger as topics than others, depending upon the semantic and grammatical features and discourse-informational status of the NP.

3. Generation of the Topic

We defined the topic, which is a semantico-grammatical notion in Korean, as the NP-nun and the delimiter-attached definite NP in S-initial position. Then, the next question is how the topic is treated in grammar or how it is syntactically generated. Section 3.1 reviews the transformational
analysis of topic, and an alternative PS analysis is defended and elaborated.

3.1 Transformational Analysis: Topicalization

This classical analysis was first proposed by Ross (1967), who argued that the following sentences

(128) a The movie I haven't seen yet.
(129) a As for the book, we will discuss it next week.

are derived from their corresponding subject-predicate sentences

(128)b I haven't seen the movie yet.
(129)b We will discuss the book next week.

by fronting the underlined NP's by means of Topicalization rule (130)

(130) Topicalization (Ross 1967: 6.126)

\[
\begin{array}{ccc}
X & \NP & Y \\
1 & 2 & 3 \\
2 & \emptyset & 3 \\
& [+\text{pro}] & \\
\end{array}
\]

A similar transformational analysis was proposed by Oh (1970) and C.M.Lee(1973), who argued that the following sentences

(131) a Chelswu-nun chayk-ul sa-ko issta.
\text{TOP book OM buy Progressive}
'Chelswu is buying a book.'

(132) a Yenghi-eykey-nun Chelswu-ka chayk-ul cwuessta.
\text{TOP SM book OM gave}
'To Yenghi, Chelswu gave a book.'
are derived from their corresponding subject-predicate sentences

\[(131)b \ \text{Chelswu-ka chayk-ul sa-ko issta.} \]
\[\text{SM book OM buy Progressive} \]
'Chelswu is buying a book.'

\[(132)b \ \text{Chelswu-ka Yenghi-eykey chayk-ul cwuessta.} \]
\[\text{SM to book OM gave} \]
'Chelswu gave Yenghi a book.'

by moving the underlined NP (along with the optional particle) to the front of the sentence via a rule like (133)

\[(133) \text{Topicalization} \]
\[
\begin{array}{c}
X \ \underline{NP} \ \text{Particle} \ \ Y \\
1 \ (2 + 1) \ # \ (1, \ \emptyset, \ 3) \ (\text{Oh 1971:130})
\end{array}
\]

Below we will argue that this analysis is not plausible.

3.1.1 One of the most serious problems to the topicalization analysis is the DSC, where the first NP is the typical topic.

\[(134)a \ \text{Chelswu-nun meli-ka cohta.} \]
\[\text{TOP brain SM good} \]
'Chelswu is smart.'

It has been argued that sentences like (134a) are derived from a subject-predicate sentence like (134b) by Rule (133):

\[(134)b \ \text{Chelswu-uy meli-ka cohta.} \]
\[\text{Possessive} \]
'Chelswu's brain is good.'

However, this so-called genitive source analysis does not
work, because there are many DSC's that do not have a genitive source.

(135)a Yenghi-nun ton -i manhta.
   TOP money SM plenty
   'Yenghi has a lot of money.'

b *Yenghi-uy ton-i manhta.
   Possessive

Another objection to the topicalization transformation is that (134a) and (134b) differ in meaning: the former is about Chelswu and the latter is not. It might be argued that this kind of meaning difference is rather stylistic and negligible; or it may be taken care of by a later surface semantic interpretation rule. However, this way out cannot be maintained due to the following additional examples:

(136)a Chelswu-man meli-ka cohta.
   only
   'Only Chelswu, he is smart.'

b Chelswu-to meli-ka cohta.
   also
   'Chelswu too, he is smart.'

c Chelswu-ka meli-ka cohta.
   exactly
   'It is Chelswu that is smart.'

d Chelswu-kkaci meli-ka cohta.
   even
   'Even Chelswu, he is smart.'

(136a-d) all differ in meaning. If (134a) is derived from (134b) and the particle nun is transformationally introduced, as Rule (133) says, all of (136a-d) should also be derived from (134b) and the delimiters should later be introduced
transformationally. It is clear that this analysis cannot be maintained. First, the topicalization rule cannot predict which delimiter should be attached to the fronted NP. Second, the transformation will bring in a wide variety of meaning changes, because all the delimiters have their own semantic content. Note that the meaning differences between (134b) and (136a-d) are truth-functional. If these meaning changes are taken care of by the surface semantic interpretation rules, the rules should be numerous and extremely complicated, because more than one delimiter is allowed and their relative orders also affect the meaning of the sentence. For example, consider

(137)a Chelswu-kkaci-mace meli-ka cohta.
   even also

b Chelswu-mace-kkaci meli-ka cohta.

(137a-b) differ in meaning. Therefore, the cost this analysis should give to the grammar is so high that any interpretive solution cannot be entertained.

In order to avoid such wild meaning changes in sentences like (136a-d), we have to assume that the delimiters already existed in the underlying subject-predicate sentences. For example, the source structure of (136a) is not (134b), but something like (138a):

(138)a *Chelswu-man-uv meli-ka cohta.
   only Possessive

from which the delimiter-attached topic Chelswu-man is
fronted. However, the problem is that (138a) is not an acceptable sentence. The same is true of all other sentences in (138b):

\[ (138)b \text{ *Chelswu-to[-ka,-kkaci]-uv meli-ka cohta.} \]

It might be argued that the underlying structures of (136a-d) are not (138a-b) but the following (139):

\[ (139) \text{ Chelswu-uy meli-man[-to,-ka,-kkaci]-ka cohta. 'It is only Chelswu's brain that is good.'} \]

Aside from the dubious acceptability of this sentence, the topicalization analysis has no way to make the delimiters, which are attached to the genitive head NP meli in (139), be moved along with the genitive NP Chelswu when the latter is topicalized. Furthermore, (139) is again different in meaning from (136a-d). That is, the semantic scopes of the delimiters in the latter cover only Chelswu, but those in (139) cover the whole NP Chelswu-uy meli 'Chelswu's brain'.

Similarly, consider the following sentences:

\[ (140) \text{ Chelswu-man[-to,-ka,-kkaci] cakicasin-i miwessta. 'For only Chelswu, the self was hateful.'} \]

There are two possible underlying structures for (140). If the delimiter man is taken for illustration, the two source structures would be the following (141a-b):

\[ (141)a \text{ Chelswu}_{i}\text{-man-i Chelswu}_{i}\text{-eykey miwessta. 'Only SM to hateful} \]
(141)b Chelswu₁-ka Chelswu₁-eykey-man miwessta.

Of these two alternatives, (141a) is not the correct choice, because what should be topicalized is the dative, but man is attached to the subject in (141a). Thus, (141b) must be the correct underlying structure of (140). However, the problem is that (141b) is again different from (140) in meaning. The former implies that people other than Chelswu do not hate Chelswu, whereas the latter implies that people other than Chelswu do not hate themselves. Therefore, there is no plausible deep subject-predicate sentence for (140).

Incidentally, Chang (1972) has argued that the differences in the semantic scopes of such logical operators as only and even cannot be handled adequately by surface interpretive rules but should be incorporated in the semantic structure. Thus, we conclude that sentences like (136) and (140) should be the underlying structures as they stand.

3.1.2 The relationship which a topic may have with its following comment sentence is of two kinds. One is that the topic is an argument of the predicate of the following comment sentence and so is grammatically tied to it. (142-144) are illustrative of this relation:

(142) Chelswu-run hakkyo-ey kassta. [subject/agent] TOP school to went 'Chelswu went to school.'

(143) Chelswu-run Yenghi-ka ttaylyessta. [object/patient] TOP SM hit 'As for Chelswu, Yenghi hit him.'
In (142-144), the topic Chelswu is an obligatory argument required for the semantic completion of the predicate of the comment. Thus, the topic maintains one of the standard grammatical relations with the predicate.

Sometimes the topic is not an argument of the predicate, but it is still related to the following comment sentence grammatically. Genitive relationship is one such case.

In (145) Chelswu is not an argument of the predicate khuta, but it has a genitive relationship with the patient khi.

The topics in (142-145) may be called "intrinsic" because they have a standard grammatical relation with, and participate in, the grammatical organization of the comment sentences following them. The transformational solution normally works for the 'intrinsic' topics because a corresponding subject-predicate sentence is always available.

The other relationship is not grammatical but rather pragmatic. In this case, the topic is not an argument of the predicate, nor is it related to other arguments in certain clearly definable ways. It does not function as a grammatical constituent of the comment and is connected to it loosely on pragmatic grounds. Topics of this nature may be called
"extrinsic."

Consider the following sentences:

(146) *wu*li-nun i pangmyen-ey cengpo-ka epsta.
    we TOP this area at information SN none
    'As for us, there is no information in this field.'

(147) Chelswu-nun hankwuk-ey ton -i manhta.
    TOP Korea at money SM plenty
    'As for Chelswu, he has a large fortune in Korea.'

(148) yangsik -un eti-ey cahun cip -i ismnayo?
    western cuisine where good house SM exist
    'As for the western cuisine, where is the best place?'

The comments in (146-148) are full sentences. The verbs
manhta, epsta, and isssta require a patient and a locative
argument to complete its semantic make-up, and both of them
are found in the comments: e.g., ton (patient) and hankwuk-
ey (locative). Thus, the comments in (146-148) can stand as
complete sentences without their topics. If the topic wuli
in (146) is forced into the comment, as in (149-150)

(149) ??i pangmyen-ey cengpo-ka wuli-ev ey epsta.
    locative

(150) ??i pangmyen-ey wuli-uy cengpo-ka epsta.
    possessive

the resulting sentences become unacceptable. One might sug-
gest that wuli of (146) is a locative, as in (149). But, as
Fillmore (1968) has argued, a sentence cannot have more than
one of the same case relation, so Chelswu cannot be a loca-
tive, either. Thus, the grammatical relation that the topic
is assumed to have with the comment in (146-148) is difficult
to specify, and it seems that there are no plausible subject-predicate sources for these sentences.

We argue that the relationship which the topic has with the comment in (146-148) is not strictly grammatical. They are rather related on semantic or pragmatic grounds: e.g., in (146), that 'there is no information in this area' is somehow relevant to 'us' (see 3.2.4).

It is relatively easy to find many extrinsic topics, for which no plausible grammatical relationship, either with the predicate or one of its arguments, is established:

(151) \[ \text{talun} \, \text{pwun} \, \text{un} \, \text{sayngkak-i an nayo.} \]
\[
\text{other people TOP thought SM not come 'As for the other person, I cannot remember him.'}
\]

(152) \[ \text{sang} \, \text{un} \, \text{i} \, \text{pen-i cheum-ita.} \]
\[
\text{prize TOP this time SM first be 'As for the prize, this is the first time.'}
\]

(153) \[ \text{mwusep-ki-nun mwues-i mwusewe?} \]
\[
\text{being-afraid TOP what SM afraid? 'Why are you so afraid?'}
\]

(154) \[ \text{yenphil-un enu kes-i cohayo?} \]
\[
\text{pencil TOP which one SM good 'As for the pencil, which is good?'}
\]

(155) \[ \text{Chelswu-nun him -i cohta.} \]
\[
\text{TOP power SM good 'Chelswu is strong.'}
\]

None of the above sentences have a plausible subject-predicate source.

It might be argued that the topics in (154-155) have a kind of locative relation with the predicate, suggesting a source structure like (156) for (154):

(156) \[ \text{talun} \, \text{pwun} \, \text{un} \, \text{sayngkak-i an nayo.} \]
\[
\text{other people TOP thought SM not come 'As for the other person, I cannot remember him.'}
\]
However, as Kuno (1973:251-252) has correctly pointed out, the deletions of such lexically non-empty elements as cwung-eyse are difficult to justify on semantic grounds.

For a slightly different example, consider the following:

(157)a  na-kyengwu-nun cwung-hakkyo ttay-ka ceyil
        I case TOP middle school time SM most
        cohaseyoy.
        good
        'In my case, the middle school days were the best.'

If a subject-predicate source is resorted to for (157a), it should be something like (157b):

(157)b  cwung-hakkyo ttay-ka na-eykey ceyil cohaseyoy.
        I to

The problem is how the expression kyengwu 'case' is generated when the experiencer na 'I' is topicalized.

Extrinsic topics are not limited to the DSC but are also common in normal subject-predicate sentences:

(158)  khophi-nun cam -i an wa.
        coffee TOP sleep SM not come
        'As for coffee, I cannot sleep (if I drink it).'

(159)  kom-sanyang-un cili-san-ulo cenghayessta.
        bear hunting TOP Mt. Cili on decided
        'As for the bear-hunting, we decided on Mt.Cili.'

(160)  nakssi-nun Kim-sensaygnim-ul mosi-ca.
        fishing TOP Mr.Kim OM invite
        'As for fishing, let's invite Mr.Kim.'
Sentences like (158-160) clearly show that the topic can have a loose pragmatic relation with the comment. We might say that the topic and the comment in these sentences are almost two independent speech acts; the speaker identifies a referent, which can be considered a single speech act, and then he adds a proposition to it.

It is clear that a transformational derivation for extrinsic topics is hardly possible. Since a transformational analysis cannot offer a uniform treatment for the two types of topics, it should be valued less highly than an approach which can (see 3.2.1).

3.1.3 Since topicalization is a syntactic process operating on formal symbols like NP, PP, etc., it is blind to the meaning of the elements involved. But topic is very much a semantic notion. Thus, the syntactic operation brings about various meaning changes difficult to accommodate by later surface interpretation rules.

The meaning of a sentence may be defined, at least partially, in terms of the conditions under which the sentence is either true or false. There are two different views of this. One is truth-functional semantics based on two-valued logic, where the truth or falsity of a sentence is determined by the relation between the sentence and the external state it describes. The other is presuppositional semantics based on three-valued logic, where the ultimate truth or falsity of a sentence is determined relative to both the contexts and
and the speaker's beliefs. The difference between the two theories of meaning is revealed most clearly when the sentence involves a presupposition failure. For example, consider

(161) The King of France is bald.
(162) That Carter won the election pleased Mary.
(163) The man whom I met yesterday was crazy.

(161) is the classic example involving the "existential presupposition" of a referring expression. (162-163) involve a "factive presupposition" that Carter won the election or that the speaker met the man yesterday. If there is no King of France and if Carter did not win the election, (161-162) are simply 'false' in the strict truth-functional semantics, because the two presuppositions are just the entailments of (161-162). However, the same sentences would be 'neither true nor false' in the presuppositional semantics, because, without these presuppositions satisfied, there is no ground on which to judge the truth or falsity of the actual assertions of (161-162). In short, the presuppositional meaning, being a 'contingent and understood' one, does not affect the truth value of the whole proposition from a truth-functional viewpoint; but it is an 'essential and integral' meaning component affecting the truth value from a presuppositional semantics viewpoint.

To get at meaning in this thesis, we will adopt the position of presuppositional semantics. There are several
reasons for this preference. First, truth-functional seman-
tics is faced with a logical paradox in sentences like the
following:

(164)a The King of France is bald.
  b The King of France is not bald.

The traditional two-value logic is based on the following
rules of inference:

(165)a Every (declarative) sentence is either true or
false.
  b If a sentence is true, its negation must be
false.

According to (165a), (164a) is false because France does not
have a king. Then, (165b) predicts that the negation of
(164a), i.e. (164b), must be true. However, (164b) is equal-
ly false. Thus, we have a logical contradiction. Note that
this kind of contradiction is easily explained in presupposi-
tional semantics.

Second, topic is a notion which necessarily involves
the speaker's and the addressee's pragmatic assumptions in
the actual communicative situation. It has very little to do
with the abstract and strictly logical concept of truth.
Consider the following:

(166) Chelswu-wa Yenghi-nun kyelhon-ha-l kes ita.
      TOP marry will
    'Chelswu and Yenghi will get married.'

    waynyahamyen,
    'Because,'

(167)a Chelswu-nun Yenghi-lul salangha-ki ttaymwun-ita.
      TOP  OM love because
  'Because Chelswu loves Yenghi.'
(167)b Chelswu-ka Yenghi-lul salangha-ki ttaymwun-ita. 'Because it is Chelswu who loves Yenghi.'

c Yenghi-ka Chelswu-lul salangha-ki ttaymwun-ita. 'Because it is Yenghi who loves Chelswu.'

d Chelswu-ka salangha-nun salam-un Yenghi-i-ki  
     SM love     ing man TOP  be  
     ttaymwun-ita. 
     because
     'Because the one who Chelswu loves is Yenghi.'

(167a-d) are all identical in their strict truth-functional meanings--they are all true or all false, depending upon whether Chelswu and Yenghi do or do not love each other. However, (167a-d) are all different in linguistic meaning. Thus, strict truth-functional meaning is too restricted for our purpose. 23

Now, we will review the transformational accounts given by Kuno (1973) and Tonoike (1975) from the semantic viewpoint. Kuno has proposed two conflicting analyses for the syntactic generation of topic. He argued that some types of topic should be directly generated as such in deep structure (Kuno 1973: chapter 21). This group of topics includes those appearing in (154-155) and (168):

(168) sayngsen-un tomi -ka cohta.  
     fish TOP snapper SM good  
     'As for fish, snapper is good.'

For another group of topics such as those in (169-170)

(169)a Chelswu-nun khi -ka khuta.  
     TOP height SM tall  
     'As for Chelswu, he is tall.'
(170)a New York-un khun kenmwul-i manhta.
    TOP tall building SM many
    'As for New York, there are many skyscrapers.'

Kuno (1973: chap 3) has argued that they should be transformationally derived from the following respective deep structures

(169)b Chelswu-uy khi-ka khuta.
    possessive
    'Chelswu's height is tall.'

(170)b New York-ev khun kenmwul-i manhta.
    at
    'In New York there are many skyscrapers.'

via the following transformational rules and derivational steps:

(171)a Chelswu-uy khi-ka khuta. [deep structure]
    b Chelswu-ka khi-ka khuta. [Subjectivization]
    c Chelswu-ka khi-ka khuta. [Marking for Exhaustive Listing]
    d Chelswu-nun khi-ka khuta. [Topicalization]

Kuno limited his Subjectivization-Topicalization analysis to the sentences in which the topic bears either the possessive or locative grammatical relations.

Kuno (1973) also suggested that the following types of sentences

(172)a Chelswu-nun nayngmyen-i cohta.
    TOP cold noodle SM good
    'As for Chelswu, cold noodles are good (to him).'

(173)a Chelswu-nun paym -i mwusepta.
    TOP snake SM terrifying
    'As for Chelswu, snakes are terrifying.'
be derived from the following deep structures

(172)b Chelswu-nun nayngmyen-ul cohahanta.
   'As for Chelswu, he likes cold noodles.'

(173)b Chelswu-nun paym-ul mwuswe-hanta.
   'As for Chelswu, he is afraid of snakes.'

On the other hand, Tonoike (1975) argues that there is no motivation for a transformation such as that changing NP-lul to NP-ka. He proposes that the deep structures of (172a-173a) are closer to those for

(172)c Chelswu-eykey nayngmyen-i cohta.
   'To Chelswu, cold noodles are good.'

(173)c Chelswu-eykey paym-i mwusepta.
   'To Chelswu, snakes are terrifying.'

arguing that Kuno's Subjectivization rule be expanded to apply to any S-initial 'NP-particle NP VP' structure. According to Tonoike, (172c-173c) are converted to (172d-173d) by Subjectivization:

(172)d Chelswu-ka nayngmyen-i cohta.
   'It is Chelswu that likes cold noodles.'

(173)d Chelswu-ka paym-i mwusepta.
   'It is Chelswu that is afraid of snakes.'

Though Tonoike does not claim that (172d-173d) are converted to (172a-173a), there is no doubt that this is what he had in mind. Thus, Tonoike's derivation may be summarized as follows:
(174a) Chelswu-evkey nayngmyen-i cohta. [deep structure]
b Chelswu-ka nayngmyen-i cohta. [Subjectivization]
c Chelswu-nun nayngmyen-i cohta. [Topicalization]

Now, let us examine the validity of Kuno's and Tonoike's analyses in derivations (171) and (174), respectively. What we are particularly interested in here are the derivational steps (171c) → (171d) and (174b) → (174c), by which the exhaustive listing NP-ka is changed to the topic NP-nun by optional Topicalization.

First of all, note that (171c) and (174b) cannot be used as discourse-initial sentences. They can be used if and only if it is known or presupposed that there is somebody who is tall or that there is somebody who likes cold noodles. This is shown by the fact that these sentences are normally used as the answers to such questions as

(175) nwu-ka khi -ka khu-ni?
who SM height SM tall
'Who is tall?'

(176) nayngmyen-ul coha-ha-nun salam-i nwukwu-i-ni?
cold noodles like -ing man SM who be
'Who is the man that likes cold noodles?'

In other words, (171c) and (174b) presuppose that somebody is tall or that somebody likes cold noodles, and what is asserted is that the identity of that somebody is Chelswu.

Let us assume a universe of discourse in which Chelswu, Yenghi, and Yengswu are registered members. Let us further assume that it is understood by all the discourse participants that all three discourse entities are short in height
or that they all hate cold noodles. If one of the discourse participants utters either (17lc) or (174b) with regard to these people in this context, then these sentences are both 'neither true nor false' because the presuppositions these sentences carry are not satisfied. Note that (17lc) is identical in meaning to (177) following:

(177) khi-ka khun kes-un Chelswu-ita.
     height tall one TOP be
     'The one who is tall is Chelswu.'

If (177) is uttered in the same context, it is also 'neither true nor false'.

However, if the same speaker utters instead either (171d) or (174c) in the same context, then these sentences are simply 'false' because they do not assume such presuppositions. Therefore, we argue that (17lc) and (174b), on the one hand, and (171d) and (174c), on the other, are different in truth value in this context, and so Kuno's and Tonoike's analyses which derive the latter from the former are problematic because the two sentences are not synonymous (not mutually entailing) but different in meaning.

Kuno's and Tonoike's analyses are semantically equivalent to claiming that the two particles ka and nun are optionally interchangeable. This is clearly wrong. If their analyses are just meant to be syntactic analyses, we have to say that this kind of a mechanical syntactic derivation blind to meaning is not interesting at all and offers no explanation for the nun-ka distinction. Therefore, we conclude that
Kuno's and Tonoike's transformational analyses of topic are not acceptable.

Ree (1969) noted an interesting correspondence between the particles nun and ka in Korean and the articles the and a in English. The following Korean sentences (178a-b)

\[
\begin{align*}
(178)a & \quad \text{ikes-i mwuncey-ita.} \\
& \quad \text{this problem be} \\
& \quad b \quad \text{ikes-un mwuncey-ita.} \\
& \quad \text{TOP}
\end{align*}
\]

may be rendered into English by (179a-b):

\[
\begin{align*}
(179)a & \quad \text{This is the problem.} \\
& \quad b \quad \text{This is a problem.}
\end{align*}
\]

(179a-b) are distinct semantically---(179a) presupposes that something is problematic but (179b) does not. The same is true of (178a-b). If the presupposition is not satisfied, both (b) sentences are simply 'false', but the (a) sentences are 'neither true nor false'. They are thus distinct in truth value. It is interesting to note that no proposal has been made in English to derive (179a) from (179b). This presuppositional difference between the (a) sentences and the (b) sentences must be reflected in underlying structure (see Morgan 1973a).

The following sentences show a different kind of meaning change involved in topicalization.
\( (180)a \) yeça-ka pang-ulō tule-wassta.
woman SM room into entered
'A woman entered the room.'

\( b \) yeça-nun pang-ulō tule-wassta.
TOP
'As for the woman, she entered the room.'

\( (181)a \) chayk-i chayksang-wi-ey issta.
book SM desk on-top-of exist
'There is a book on the desk.'

\( b \) chayk-un chayksang-wi-ey issta.
TOP
'As for the book, it is on the desk.'

The \( (a) \) sentences above normally imply that the 'woman' and 'book' are indefinite entities, whereas the same 'woman' and 'book' in the \( (b) \) sentences imply that they are definite. The difference between the two sentences may be represented as follows:

\( (182)a \) \( \exists x (x:woman) \cdot \text{ENTER} (x, \text{room}) \)

\( b \) \( \forall x (x:woman) \cdot \text{ENTER} (x, \text{room}) \)

The existential and the iota operators have distinct meanings.

Similarly, consider the following sentences:

\( (183)a \) namnye-konghak-ul ha-nun hakkyo-ka cohta.
co-ed study OM do-ing school SM good
'Schools which co-educate are good.'

\( b \) hakkyo-nun namnye-konghak-ul ha-nun kos-i cohta
school TOP co-ed study OM do-ing place good
'As for the school, those which co-educate are good.'

In the transformational analysis \( (183b) \) is derived from \( (183a) \) by preposing the relativized head NP 'school'. But the two sentences differ in meaning; \( (183b) \) is about schools
in general, while (183a) concerns certain co-educational schools.

Sentences (17lc-d) and (178-183) show that if NP-nun is derived from NP-ka by topicalization, we are faced with various meaning changes. It might be suggested that these meaning changes be taken care of by later surface interpretive rules. However, this is not possible. Note that the (a) sentences and the (b) sentences in (178-182) are minimal pairs with the only difference being the particles ka and nun. Thus, it is clear that these meaning differences between the two are due to the semantic differences between the two particles, and that no surface interpretive rules based on the complementary contextual environments in which they occur are possible.

The real and important issue we should address at this moment is what kinds of syntactic phenomena should be treated transformationally at the expense of meaning? And why? Surface interpretive rules are semantic adjustment rules which restore the semantic deviations (additions or losses of some semantic features) caused by syntactic operations called transformations. It is often said that if the semantic deviation is serious enough to cause a difference in the truth-functional meaning, transformations are not allowed. However, we have seen that the truth-functional meaning falls far short of the linguistic meaning. It seems that the semantic deviations caused by transformations are different in degree rather than kind, forming a continuum. The problem is
at what point of this continuum the surface semantic adjustment rules should, or should not, be allowed to operate and why? There is no reason why (179a-b) should not be related by a transformation, e.g. the-insertion rule, if a surface interpretive rule adds the meaning of 'definiteness' to any constituent to which the has been added. Similarly, why can we not posit a si-insertion rule to accommodate the following sentences:

(184)a  sensayngnim-i wassta.
       teacher SM came
       'The teacher came.'

b  sensayngnim-i o-si-essta.
    HON
       'The teacher(HON) came.'

if a surface interpretive rule is assumed to the effect that any constituent to which si is attached gets an additional 'honorific' meaning.

It seems that the syntactic operation we call 'topicalization' is motivated by a desire to devise a maximally simple syntactic generative mechanism. For example, nothing will be simpler than to have a topicalization rule to generate (131a-132a) and (131b-132b) together, or to generate the sentences in (183). However, we believe that the question of how a set of data can be generated with the least amount of generative machinery, as in Kuno's derivation (171), is really irrelevant to explaining how various linguistic forms are linked to their meanings and how they differ from each other in communicative function. Therefore, the mechanical
syntactic operation of topicalization is very uninteresting, at best, and has little to offer in explaining the ka-nun distinction, and the notion and communicative function of topic.

Finally, for those who believe that the meaning differences which really count as evidence against transformations are assertive, not presuppositional ones, there is evidence that topicalization sometimes brings about a change even in the assertive meaning.

(185)a i cip - uy cengwen-i khu-ko cohta.
this house of garden SM big and good
'The garden of this house is big and good.'

b i cip-un cengwen-i khu-ko cohta.
this house TOP garden SM big and good
'As for this house, the garden is big and good.'
'As for this house, the garden is big and the house is good.'

(185a) has only one reading, but its topicalized counterpart has two readings. The second reading of (185b), in which what is good is 'the house', asserts something different from (185a), where what is good is 'the garden'.

To recapitulate, the syntactic operation of 'topicalization' brings about a variety of meaning changes, either in presuppositions or assertions. These meaning changes are difficult to accommodate by surface interpretive rules. Nor is it desirable to handle them that way. Therefore, the topicalization analysis is not plausible.

3.1.4 From the standpoint of speech production, it seems
correct to say that the speaker often determines and produces the topic even before he has a clear idea of what sort of predication he is going to make about it. Hesitations which commonly fall between the topic and the following comment sentence indicate that this is exactly what is taking place in actual speech production. The topicalization hypothesis, in which the topic is extracted from an already developed subject-predicate sentence is either incompatible with or irrelevant to actual speech production. Other things being equal, an analysis which fits in with actual speech production is to be more highly valued than one that does not.

Furthermore, the topic seems to be outside the performative modality of the following comment sentence. Since modalities are expressed in S-final position in Korean, we have not found a good example showing this point, but consider the following English example:

(186) As for Chelswu, I promise you that he cannot do the work.

Under the topicalization analysis, the rule would move the topic to outside the performative structure. This would be an exceptional kind of movement, because topicalization is a sentence-bound rule.

We conclude, then, that transformational generation of topic is not a feasible choice for a well-motivated explanation of the phenomena presented.
3.2 A Proposed Analysis

3.2.1 We have so far argued that topichood in Korean optimally has three components: syntactically, the 'higher NP' position; morphologically, the particle nun; and semantically, definiteness. We have also argued that topic cannot be considered an element extracted from the subject-predicate structure by a movement rule. What we suggest in this study is that topic, as a constituent of the underlying structure, is generated directly by the basic formation rules of the language. For example, the following sentences with topic noun phrases:

   [Chelswu too,] his wife is pretty
   'As for Chelswu, his wife is pretty.'

(188)a Chelswu-[nūn] chayk-ul sa-ko issta.
   [Chelswu too,] he is buying a book.
   'As for Chelswu, he is buying a book.'

can be generated by PS rules beginning thus:

(189)a S → (NP) S'
b S' → NP VP
c VP → (NP) (PP)...

What we will be mostly concerned about in this section is PS rule (189a). From (189a-b), the topic and the comment can be defined as [NP]S and [S']S, respectively, and the subject as [NP]S'. This means that, unlike as in English, the semantic concept of topic has been syntacticized as [NP]S in
The PS analysis of topic is preferable to the transformational analysis both for theoretical and empirical reasons. First, all the meaning changes the topicalization rule brings about in 3.1.1 and 3.1.3 do not arise in the alternative analysis, because nun and other delimiters are directly attached to the topic in underlying structure.

Second, 'extrinsic' topics such as (151-160) cause no problems for this analysis because, as PS rule (189a) shows, topic choice is independent of the particular noun phrases in the comment sentence. As for the 'intrinsic' topics, we maintain that when the speaker establishes the topic, he sometimes anticipates the syntactic role it will play in the grammatical structure of the comment sentence. Thus, the PS analysis is empirically more adequate than the transformational analysis since it allows that the topic may or may not participate in the grammatical organization of the comment sentence. The PS analysis is also theoretically more plausible than the transformational one because it offers a more general and unified account of topics.

Third, the PS analysis, in which topic is generated even before the comment sentence is formed, is a conceptual model that is closer to, or more compatible with, actual speech production than the transformational analysis.

3.2.2 PS rule (189a) shows the topic as an optional constituent, which means that some sentences have a topic and
some do not. The question is now what sentences have topics (NP-nūn) and what sentences do not. This is related to the controversial question concerning the distribution and occurrence of the particles nun and ka. Furthermore, if the sentence has a topic, how has the topic actually been established? This is important to the PS analysis, since we have rejected the simple extraction hypothesis. We will consider the first question here and take up the second in the next section.

We claim that there are basically two ways of describing a given event or state of affairs linguistically. One is to describe the event or state of affairs objectively as it is, or as it is perceived by the speaker. The other is to describe the event or state of affairs from a certain perspective, i.e., to relate the event or state of affairs with regard to a certain entity that is somehow associated with the event or state of affairs to be described.

To use a metaphor, it is just like taking a picture of a scene with a camera. We can take a picture of the entire scene with the focus of the camera given evenly to all the things involved in it—a object, foreground, background, and all. We might say that the camera is placed directly in front of the scene, e.g. something like the following:
This will result in a rather unfocused objective description of the entire scene.

Or, we can take a picture of the same scene with the focus of the camera centered mainly on a certain entity. We might say that the camera is placed in a position where the angle is in direct line with the chosen subject, but not with the entire scene. To use a diagram, this may look like

The focus or perspective from which the event or state of affairs is described may be called the 'topic'.

If the former mode of description is chosen, the sentence will have no topic, and if the latter is chosen, it will have a topic. The former sentence type will be called "topic-less" and the latter "topical" (or a "topic-comment sentence"). Note that, in this view, whether a sentence will have a topic or not is ultimately dependent upon the speaker's communicative intent or attitude in describing a given event or state of affairs. Though the use or non-use of a topic is ultimately up to the speaker, this is not totally arbitrary but is predictable to a certain extent. This is the main subject of this section.

According to Kuroda (1972), propositions ("statements" in his term) can be classified into two types according to
their semantic nature: "specific" and "generic." 'Specific' propositions are ones which "refer to a particular occurrence of an event or state of affairs" (1972: 160) at a specific time and location. Such propositions contrast with those describing some general event or constant state of affairs. Now, the first point we want to make is that, if the speaker has the communicative intent of describing the specific occurrence of an event or state of affairs objectively (i.e., with the focus given evenly on the entire situation), he must express the proposition in a 'topic-less' sentence. The clearest case in which this generalization works is when the speaker describes an on-going action or currently existing state of affairs in the here-and-now context objectively as he perceives it. The event or state of affairs may be undergoing a change at the time of utterance. For instance, consider

(190)a  o! palam-i/*un siwenha-kwuna.
    oh wind    cool
    'Oh! The wind is cool.'

(191)a (Looking at the moon)
    a! tal -i/*un palkta.
    ah moon    bright
    'Ah! The moon is bright.'

(192)a (Putting the hand into water)
    mwul-i/*un cha-kwuna.
    water    cool
    'The water is cool.'

(190a-192a) all represent the speaker's direct and single perception of a currently existing state of affairs.
(193)a a! talk-i/*un wu-nunkwuna.
rooster cry
'Ah! a rooster is crowing.'

(194)a (Looking out of the window)
yâ Chelswu-ka/*nun o -ko issta.
wow come Progressive
'Wow! Chelswu is coming.'

(193a-194a) represent the speaker's direct perception of an event that is taking place before his eyes. The state of affairs expressed by (190a-192a) and the events in (193a-194a) are specific to this moment, e.g., the wind in (190a) may not blow at all tomorrow, and Chelswu's activity in (194a) may terminate very soon. We argue that the specific sentences of this nature cannot have topics. 29

Also related to this hypothesis are sentences whose verbs represent an 'emergence into' or 'existence in' the current scene. Consider

(195) ie John-i/*un nathanassta.
soon appeared
'Soon John came out.'

(196) ku ttay, kwunin-tul-i/*un pang-an-ulo tulewassta.
then soldiers room into entered
'Then, soldiers came into the room.'

It seems that in (195-196) the focus is given to the entire scene rather than to a certain entity like 'John' or 'soldiers'.

Of course, direct descriptions need not be limited to the currently on-going events in the here-and-now context. The speaker can also describe an event of the past, as a
whole, as it was. Consider

(197) #ecey pi -ka/*nun wassta.
yesterday rain came
'Yesterday it rained.'

(198) a #Chelswu-ka/?nun ecey cwookessta.
yesterday died
'Chelswu died yesterday.'

In (197-198) the speaker is making a single assertion on the event which occurred in the past, and specific sentences of this type are common in discourse-initial or isolated sentences. 29

Another situation in which an objective description is used is when there is no definite object involved in the event to be described. Since topics must be semantically definite, these propositions normally have no topics. Consider the following sentences.

(199) etten namca-ka/*nun moca-lul mantul-ko issessta.
certain man hat OM make Progressive
'A man was making a hat.'

(200) sey myeng-uy chengnyen-i/*un han yeca-lul
three young man one woman OM
hit
'ttaylyessta.
hit
'Three young men hit a woman.'

Note that the particle nun is not appropriate in (199-200).

In contrast to the objective description, the speaker can choose the option of describing an event or state of affairs from a certain angle, i.e. with regard to an entity involved in the event or state of affairs. This will result
in a 'topical' (or 'topic-comment') sentence. One of the situations in which the use of topic is mandatory (regardless of the speaker's communicative intent) is when the proposition to be conveyed is semantically "generic," where 'generic' means that the proposition refers to "general, habitual, or constant state of affairs of some sort." Following Kuroda (1972), we contend that 'generic' propositions must be expressed in a topical sentence. For example, with a few modifications, all the sentences (190a-194a) can be rendered into generic propositions; then, they must be expressed by means of a topical sentence. Compare (190a-194a) with the following:

(190)b yelum palam-un/*i siwenhata.  
summer wind cool  
'The summer wind is cool.'

(191)b tal -un/*i polum -ey palkta.  
moon full-moon at bright  
'The moon is bright when it is full.'

(192)b mwul -un/*i enceyna chata.  
water always cool  
'Water is always cool.'

(193)b talk -un/*i saypyek-ey wunta.  
rooster morning at cry  
'The rooster crows in the morning.'

(194)b Chelswu-nun/*ka mayil yeses-si-ey o-nta. everyday six o'clock come  
'Chelswu comes at six everyday.'

Note that (190b-194b) are all 'generic' and they must have topics.

Most representative of the 'generic' sentences are "definitional sentences" with the copula ita 'be' as its main
verb. Sentences expressing a general truth or an habitual action are also typically 'generic'. All of these sentences must have a topic.

(201) Chelswu-nun/*ka uysa-ita.
    doctor be
    'Chelswu is a doctor.'

(202) 'paper'-nun/*ka congi-ita.
    paper be
    'Paper' is 'congi' (in Korean).'

(203) Seoul-un/*ka hankwuk-uy swuto-myeng-ita.
    Korea of capital name be
    'Seoul is the name of the capital of Korea.'

(204) mwul -un/*i swuso-wa sanso-lo toye-issta.
    water hydrogen and oxygen with become
    'Water consists of hydrogen and oxygen.'

(205) cikwu-nun/*ka twungkul-ta.
    earth round
    'The earth is round.'

(206) Chelswu-nun/*ka pyengwen-eyse ilhanta.
    hospital at work
    'Chelswu is working at a hospital.'

(201-204) are all definitional sentences, and (205-206) express either a truth or a habitual action. Note that NP-ka is not acceptable with them.

Of course there is no reason why generic propositions should be expressed in a topic-comment structure, but this is a linguistic fact. Our explanation for this phenomenon is that the semantic feature of 'genericness' is always associated with a certain quality or property and that a quality or property always requires an object which carries it or to which it is attributed. Therefore, a generic proposition must be presented linguistically in the bipartite form
of topic and comment.

When a proposition is 'specific', however, it can be expressed either topiclessly or topically. While discussing the former option, we mentioned that specific events of the past can be described objectively as in (198a). But the same event can also be described topically with regard to Chelswu. Consider

(198)b (Chelswu-ka ettehkey toyessni?)
SM how became
'What happened to Chelswu?
Chelswu-nun/*ka ecey cwukessta.
  yesterday died
  'Chelswu died yesterday.'

Note that (198a) is felt to be a rather objective report of what happened yesterday, i.e. a certain death in which Chelswu is involved, whereas (198b) is felt to be a description of a death with a particular reference to Chelswu. It seems that the only difference between (198a) and (198b) is the absence and presence of 'aboutness' respectively. 30

We argued in 2.4.5 that definiteness contributes to topicality. When a definite entity is involved in the event or state of affairs to be described, there is a tendency to describe the event topically with special reference to the definite entity, rather than describing it objectively. Thus, when a definite entity is involved, even a specific event which is taking place before the eyes of the speaker in the here-and-now context can be described topically. For example, compare (207a) with (207b):
Another case in which the use of a topic is almost obligatory is when a specific proposition occurs within a discourse. When the event involves an entity already mentioned in the previous discourse, the event is normally described topically with regard to the entity. Consider

(208)a #Chelswu-ka chayk-ul ilk-ko isssta.
book OM read Progressive
'Chelswu is reading a book.'

b ku-nun/*ka pelsse han sikan-ccay ilk-ko isssta.
he already one hour read -ing
'He has already been reading for an hour.'

(209)a nay-ka Yenghi-lul mannass-ul ttay,
I SM OM met when
'When I met Yenghi,'

b na-nun/*ka cim -ul wunpanha-ko issessta.
I luggage OM carry Progressive
'I was carrying luggage.'

(210)a (Two men are watching Chelswu pass by them.)

#Chelswu-ka eti-ey ka-nayo?
SM where to go
'Where is Chelswu going?'

b Chelswu-nun/*ka hakkyo-ey ka-ko isseyo.
school to go Progressive
'Chelswu is going to school.'

(211)a #etten keci-ka kilkeli-eyse se-issessupnita.
certain beggar street on stand -ing
'A beggar was standing on the street.'
Note that the propositions represented by the (b) sentences must be expressed topically. Our explanation for this fact is this: once an entity like Chelswu or etten keci 'a beggar' has been mentioned in the previous discourse and has thereby become anaphoric or definite, the entity is put into an epistemologically privileged position. It is thus almost impossible to describe the event or state of affairs objectively.

Note that sentences (190a-192a), which do not allow the topic NP-nun, are all meteorological expressions describing such objective natural phenomena as a snowfall, rainfall, or moonlight. It is almost impossible to describe these natural phenomena with the focus centered on any entity involved in them. This seems to be why the sentences describing a natural phenomenon do not have topic-comment structure.

To sum up, the speaker has the option of describing an event or state of affairs in two ways: one is to describe the whole event objectively as it stands, and the other is to describe the event with special focus given to an entity involved in it. The former option leads to a 'topic-less' sentence, and the latter to a 'topical' sentence. Which mode of description is used depends ultimately upon the speaker's communicative intent, but it can in part be predicted if we consider such factors as 'generic-specific', 'discourse-
anaphoric', and 'definite'.

3.2.3 There are other interesting problems concerning PS rule (189a). Even though the speaker has decided to put the proposition into a topical sentence, he may find that the underlying proposition contains more than one candidate qualifying for choice as the topic. Normally, only one of them is actually chosen as the topic, and there are several factors--semantic, syntactic, discourse-informational--which are involved in, and also govern, the actual establishment of the topic. We will call these factors involved in topic choice "topic hierarchies." Topic hierarchies exist because the topic is what the speaker wants to talk about, and there are universal and language-specific tendencies with regard to what people are likely to talk about.

3.2.3.1 We argued that a generic proposition must have a topic. If the speaker has the following messages:

(212)a mwusepta (paym) 
terrifying snake

(213)a yeypputa (cangmi) 
beautiful rose

(214)a manhta (ton) 
plenty money

the first two, being generic, must be presented linguistically in a topic-comment structure. Thus, (212a-213a) will be realized as (212b-213b) respectively:
(212)b  paym-un/*i mwusepta.
    snake     terrifying
     'The snake is terrifying.'

(213)b  cangmi-nun/*ka yeypputa.
    rose       pretty
     'The rose is beautiful.'

(214a) does not allow the generic interpretation; so it is
realized syntactically as follows:

(214)b  ton-i/*un manhta.
    money    plenty
     'There is a lot of money.'

As (212b-214b) show, when there is only one patient argument,
it must be chosen either as the topic or as the subject de-
pending upon the semantic property of the proposition.

However, let us assume that the underlying propositions
are something like (215-217):

(215)a  mwusepta (paym, na)
    terrifying snake I

(216)a  yeypputa (cangmi, kkoch)
    beautiful rose   flower

(217)a  manhta (ton, Chelswu)
    plenty money

where the first argument is the patient and the second either
an experiencer or a locative. Since (215-217) are all gener-
ic, they must be realized syntactically as a topic-comment
structure. The question, then, is which of the two arguments
will be chosen as topic. It seems that the experiencer and
the locative are higher in topicality (or, the degree of ease
with which they are talked about) than the patient.
Therefore, the choice of the former as the topic is much more natural than that of the latter. Consider

(215)b  na-nun paym-i mwusepta.
     'For me, snakes are terrifying.'
  c  ?? paym-un nay-ka mwusepta.
     'As for snakes, I am afraid of them.'

(216)b  kkoch-un cangmi-ka yeyppta.
     'Among flowers, the rose is beautiful.'
  c  *cangmi-nun kkoch-i yeyppta.
     'As for roses, the flower is beautiful.'

(217)b  Chelswu-nun ton-i manhta.
     'As for Chelswu, money is plenty for him.'
  c  *ton-un Chelswu-ka manhta.
     'As for money, Chelswu has a lot of it.'

The (b) sentences are unmarked, so to speak, that is, without contextual restrictions. But the (c) sentences are highly marked or even unacceptable. The latter can be used only if certain textual presuppositions hold. For example, (215c) may be acceptable only in such a discourse as the following:

(218)  Yenghi-nun paym-i mwusepta. anya ______.
     'Yenghi is afraid of snakes. No.'

Thus, we propose the following semantic topicality hierarchy:

(219) \{Experiencer\} \supseteq \{Patient
     \{Locative

This means that the experiencer or locative normally takes precedence over the patient in topic choice. It seems that this hierarchy reflects a pragmatic tendency for someone
experiencing an emotion to be in an epistemologically more privileged position to be talked about than the object triggering the emotion.

3.2.3.2 The speaker can express the message or proposition which he wants to convey in various ways. Which form of expression is used is largely dependent upon two decisions. One is how the state of affairs represented by the proposition is defined. This task is accomplished by subject choice, i.e. by how the syntactic function 'subject' is assigned (see Fillmore 1977; Dik 1978 for details).

The other decision is how the proposition is connected to what has already been said. Otherwise, as Grice's "Maxim of Relation" predicts, communication might fail. Since topic is what the sentence is about, it is natural that the topic carries the main responsibility of achieving the desired discourse connection. Normally (i.e., when one does not aim at producing a specific communicative effect such as 'emphasis') the speaker wants an easy and smooth transition from what has been said. This is achieved by taking as the topic an element already mentioned in the previous discourse and thus easily identifiable. In this sense, the informational structure of the discourse can play an important role in topic choice, and the topic functions as a bridge between the prior discourse and what the speaker is going to assert.

For example, let us assume that one has the following message to convey:
(220) ssessta ( cakka, sosel )
  wrote  writer  novel

where the two arguments cakka and sosel are on equal terms in competing to be the topic.

If the discourse previous to (220) was the following

(221)a na-nun ecey  sosel-ul ilkessta.
  I TOP yesterday novel read
  'I read a novel yesterday.'

b kulentey __________________
  and

then, out of the two arguments cakka and sosel in (220), the latter, rather than the former, will get assigned the topic function. This is because, sosel, having already been mentioned, makes an easy transition from (221a). Depending upon which of the two arguments gets assigned the subject function, (220) will be realized syntactically as either (222a) or (222b):

(222)a (ku) sosel-un etten cakka-ka ssessta.
  [+topic]  [+subject]
  'As for the novel, a writer wrote it.'

b (ku) sosel-un etten cakka-ey uyhaye ssuyecyessta.
  [+topic/+subject]
  'As for the novel, it was written by a writer.'

In the context (221), however, cakka will normally not be chosen as the topic:

(223) *etten cakka-nun ku sosel-ul ssessta.
  'A writer wrote the novel.'

However, if the previous discourse was the following
(224) a na-nun ecey etten cakka-lul mannassta.
    I TOP yesterday writer OM met
    'I met a writer yesterday.'

    b kulentey ___________
    and

cakka, rather than sosel, will be chosen as the topic, be-
cause cakka is discourse-anaphoric. If this argument is
simultaneously chosen as the subject, (220) will be realized
syntactically as follows:

(225) ku cakka-nun sosel-ul ssessta.
    [+topic/+subject]
    'As for the writer, he wrote a novel.'

But, in this context, the choice of sosel for the topic will
be inappropriate:

(226) *sosel-un ku cakka-ka ssessta.
    'As for the novel, the writer wrote it.'

Consider the following discourse from a children's
story.

(227) a aki twayci yel-twu mali-ka sophwung-ul kassupnita.
    baby pig twelve SM excursion OM went
    'Twelve piggies went on a picnic.'

    b _________________

After (227) the speaker has the following message to convey:

(228) ssa-kaciko kassta (kutul, cemsim)
    prepare-and-went they lunch

where kutul is discourse-anaphoric (thus an "activated given"
in Gundel's (1977) terms) and cemsim is not (thus an
"unactivated given"). Note that only the former can be chosen for the topic, as in (229a-b):

(229) a kutul-un cemsim-ul ssa-kaciko-kassupnita.
'\text{They went with a lunch box.}'

b *cemsim-un kutul-i ssa-kaciko-kassupnita.
'\text{As for lunch, they went with it.'}

Thus, we propose the following discourse-informational topicality hierarchy:

(230) Discourse anaphor \textgreater Discourse non-anaphoric

This hierarchy means that it is easier to talk about an object that has already been mentioned in the discourse than an object that has not been. In Gundel's (1977) terms, 'activated given' information is higher in topicality than 'unactivated given' information. Note that it is actually this hierarchy that is responsible for the phenomena we observed in sentences (208-211) in 3.2.2.

It seems that the informational hierarchy operates without exception when the discourse-anaphoric entity has already been the topic of a previous question. Consider

(231) a Speaker A: John-un cocciy muey-ul hayssni?
\text{TOP yesterday what did}
'As for John, what did he do yesterday?'

b Speaker B: __________

Let us assume that Speaker B wants to say the following proposition after (231a):
Since John and na are two co-participants in the act of studying, they are on equal terms in competing to be the topic. Moreover, as we will see in 3.2.3.5, the first person is generally higher in topicality than the third person. In spite of that, it is the third person John, not the first person na, that is chosen as the topic of (231b) in the context of (231a):

(232)a John-un na-hako kathi kongpwuhayssta. 'For John, he studied with me.'

b *na-nun John-hako kathi kongpwuhayssta. 'For me, I studied with John.'

It seems that the discourse-informational hierarchy is stronger than other hierarchies to be discussed in this section. For example, this hierarchy often takes precedence over other hierarchies when it conflicts with the others, such as with the "animacy hierarchy" (3.2.3.4) and the "person hierarchy" (3.2.3.5), as in (232).

3.2.3.3 When the topic maintains a grammatical relationship with the following comment sentence, the topic choice is also affected by the grammatical role the topic would play in the comment. Keenan and Comrie (1977) and Johnson (1977) propose a hierarchy of grammatical relations, which applies to the grammatical processes having the effect of making an NP the most prominent element in a clause. Since topic
choice is just such a process, their NP accessibility hier-
archies are just a sort of topicality hierarchy formulated
in terms of grammatical relations. Keenan and Comrie's hier-
archy is formulated as follows:

(233) Subject ≥ Direct object ≥ Indirect object ≥ Object
    of Preposition ≥ Possessor NP ≥ Object of Com-
    parative particle

Keenan and Comrie's hierarchy, as it stands, does not
fit the Korean data. For instance, the possessor NP is high-
ly accessible to topic choice, and the 'object of preposi-
tion' should be further subdivided according to the particles
used. We propose a revised NP accessibility hierarchy for
Korean as follows: 32

(234) Subject ≥ Direct object ≥ Indirect object ≥ Indirect object
    \{ Locative \} ≥ \{ Instrument \} ≥ \{ Comparative \}
    \{ Temporal \} ≥ \{ Reason \} ≥ \{ Comitative \}
    \{ by-passive agent \}

We will call this a "topicality hierarchy of grammatical re-
lations" (or simply "grammatical hierarchy"). The grammati-
cal hierarchy says that the subject is easier to talk about
than the object, which is in turn easier to talk about than
the indirect object, and so on. One thing that should be
pointed out is that there is a cut-off point at which an NP
cannot be accessible to topichood, and, as we will discuss
in Chap 4, the cut-off point must be between the locative/
temporal/source and the instrument/reason. This means that
any NP which is lower in accessibility than the locative cannot be chosen as topics.

Now, let us go back to sentences (224-225). We repeat them here for ease of reference. The previous discourse was

(224)a na-nun ecey etten cakka-lul mannassta.
I TOP yesterday a writer OM met
'I met a writer yesterday.'

and the speaker had the following message to convey:

(220) ssessta ( cakka, sosel )
wrote writer novel

Since cakka is discourse-anaphoric, it is more eligible for topic choice than sosel. Depending upon which of the two arguments is assigned the subject function, (220) will be realized syntactically either as (225) or (235):

(225) *ku cakka-nun sosel-ul ssessta.
[+topic/+subject]
'As for the writer, he wrote a novel.'

(235) *ku cakka-nun sosel-i (ku-ey uyhaye) ssuyecyessta.
[+topic] [+subject]
'As for the writer, the book was written by him.'

Note that (225) is acceptable but (235) is not.

The acceptability of (225) and unacceptability of (235) reveal an interaction of the discourse-informational hierarchy and the grammatical hierarchy. The two hierarchies are in conflict in (235). The informational hierarchy will choose cakka over sosel for the topic, but the grammatical hierarchy predicts that sosel is more accessible to topichood
than cakka. As mentioned in 3.2.3.2, the informational hierarchy tends to overrule other hierarchies, including the grammatical hierarchy. For example, there are many cases in which the grammatical object is chosen over the subject as the topic, because the former is discourse-anaphoric. However, this overruling is allowed only within the topic accessibility limit of the grammatical hierarchy. Any NP lower than the cut-off point of the grammatical hierarchy cannot be chosen as the topic, whatever informational status it may have. (225) is all right because the subject is a grammatical relation accessible to topic, but (235) is not because by-passive agent is a grammatical relation in the lowest position of the hierarchy, below the cut-off point. Thus, (235) creates an unresolvable conflict between the two hierarchies, and this explains why (235) is not acceptable.

3.2.3.4 Another factor to be considered in topic choice is "animacy." Since we can more easily project ourselves into the point of view of an animate being than that of an inanimate object, animates tend to be chosen more often as topic than inanimates. Consider

(236)a na-nun oppa-eykey kkoch-path-ul mantule-I TOP brother to flower bed OM make

talla-ko collassta.
give asked
'I asked the brother to make a flower bed.'

b kulayssteni 
then
Let us suppose that the speaker has the following message after this:

\[(236)b\] mantul-e cwuta (oppa, kkoch path, cengwen) make-for brother flower-bed garden

where oppa is an animate agent, kkoch path, an inanimate patient, and cengwen an inanimate locative. The discourse-informational hierarchy predicts that the locative, being discourse non-anaphoric, is out of the question, and that the animate agent and the inanimate patient are on equal terms in the competition to be topic. In this case, however, only the animate agent can be chosen:

\[(237)a\] oppa-nun cengwen-ey kkochpath-ul mantule cwuessta. 'My brother made a flower bed in the garden.'
\[b\] ??kkochpath-un oppa-ka cengwen-ey mantule swuessta. 'The flower bed, my brother made it in the garden.'

The precedence of animacy over non-animacy (whatever semantic functions are associated with the latter) is clear in the following examples:

\[(238)a\] ku yeca-nun i sako -lo cwukkessta. the woman this accident in died 'The woman died in this accident.'
\[b\] *i sako-nun ku yeca-lul cwukyessta. 'This accident killed the woman/'

\[(239)a\] Chelswu-nun ku nyusu-ey nollassta. the news at surprised 'Chelswu was surprised at the news.'
\[b\] *ku nyusu-nun Chelswu-lul nolla-key hayssta. 'The news surprised Chelswu.'
(240)a Chelswu-nun ecey ton-i wassta.
   yesterday money SM came
   'As for Chelswu, money came yesterday.'

b ?? ton-un ecey Chelswu-eykey wassta.
   'As for money, it came to Chelswu yesterday.'

In (238), ku yece is the animate patient and i sako is the inanimate agent. In (239), Chelswu is the animate experiencer and ku nyusu is the inanimate agent. Note that it is the 'animate' entity that is normally chosen to be the topic, even though the agent is generally higher in topicality than either the patient or the experiencer. In (240), Chelswu is the animate locative and ton is the inanimate patient. Again, the animate is easier to be chosen as the topic than the inanimate. Thus, we propose the following animacy topicality hierarchy:

(241) Animacy $\geq$ Inanimacy

Animacy is often associated with agency, and the resulting animate agent seems to be strongest in topicality. For example, (237a-b) show that the animate agent is higher in topicality than the inanimate patient. Since we also found in 3.2.3.1 that experiencers and locatives are higher than patients in topicality, it would be useful to put all of these semantic functions together and examine their relative strength of topicality.

It seems that not only locatives but also temporal arguments are higher in topicality than patients. Consider
(242)a  #ecey-nun pi-ka naylyessta.  
yesterday rain SM fell  
'Yesterday it rained.'

b  #pi-nun ecey naylyessta.  
'As for the rain, it came yesterday.'

(243)a  #i maul-un onul cang -i senta.  
this village today market SM open  
'As for this village, the market opens today.'

b  #cang-un onul i maul-eyse senta.  
'As for the market, it opens in this village today.'

(242a-243a) have temporal and locative NP's as topics, whereas (242b-243b) have patients. In normal contexts, the (a) sentences are more natural than the (b) sentences. This seems to indicate that temporal and locative arguments are more accessible to topichood than patients.

Between experiencers and temporal/locative NP's, the former seem to be higher in topicality than the latter. Consider the following:

(244)a  Chelswu-nun onul maywu sulphessta.  
today very sad  
'Chelswu was very sad today.'

b  onul-un Chelswu-ka maywu sulphessta.  
'For today, Chelswu was very sad.'

(245)a  Yenghi-nun hakkyo-eyse mopsi hungpwunhayssta.  
school at very excited  
'Yenghi got excited very much at school.'

b  hakkyo-(eyse)-nun Yenghi-ka mopsi hungpwunhayssta  
'At school Yenghi got excited very much.'

(244a-245a), which have experiencers as topics, are more natural than their (b) counterparts, which have temporal or locative NP's as topics.
Between animate agents and experiencers, it is the agents that are higher in topicality. Consider

(246)a Chelswu-nun Yenghi-lul sulphukey hayessta. OM sad made 'Chelswu made Yenghi feel sad.'

  b Yenghi-nun Chelswu-ka sulphukey hayessta. SM sad made 'As for Yenghi, Chelswu made her feel sad.'

where Chelswu is the animate agent and Yenghi the experiencer. It seems that (246a) is more natural than (246b) in normal contexts.

So far we have argued for three binary hierarchical relations: Animate agent ≥ Experiencer; Experiencer ≥ Temporal/Locative; and Temporal/Locative ≥ (Inanimate) Patient. Assuming such relations are transitive, we can hypothesize that these three binary hierarchies will constitute the following linear implicational schema, which we may call a "topicality hierarchy of semantic functions."

(247) Agent ≥ Experiencer ≥ \{Temporal\} ≥ Patient

The correctness of this semantic hierarchy depends upon whether the assumed transitivity relation among the binary hierarchies holds true, and it seems to hold true. We have already shown in 3.2.3.1 that experiencers are higher than patients and sentences (237a-b) show that agents are higher than patients. The remaining relation to be examined is that between agents and temporal/locative NP's. Consider
The (a) sentences with agent topics seem more natural than the (b) sentences with temporal or locative topics in normal contexts, i.e. when they are unaffected by prior discourse or without specific communicative effects (e.g. emphasis) intended by the speaker. Thus, in neutral contexts, agents seem to be higher than temporal or locative NP's.

The semantic hierarchy (247) suggests that, unless no specific communicative effect is intended, the animate agent NP is the easiest to talk about, the experiencer the next easiest, and the patient NP is the most difficult to talk about. Thus, the process of topic choice is sensitive to this semantic hierarchy.

3.2.3.5 Another topicality hierarchy may be called the "person hierarchy" (Kuno 1976; Givon 1976).

(250) Speaker ≥ Hearer ≥ The third person

As Givon points out, this hierarchy "expresses the egocentric character of discourse, where the speaker tends to be the universal point of reference and the most highly
presupposed argument" (1976:153). It seems likely that there is a distinction between the speaker and the hearer in topicality, but conclusive examples are not easy to come by. The following discourse may seem relevant:

(251)a na-nun tangsin-ul salanghapnita.
   I TOP you ON love
   'I love you.'

   b *tangsin-un na-lul salanghapnita.

But note that the acceptability of these sentences will be reversed in questions. So, sentences like (251) support neither side.

   In contrast, the distinction between the speech roles (the speaker and the hearer) and the third person in terms of topicality seems to be a valid one. There is evidence that the speech roles take precedence over the third person in topic choice.

(252)a na-nun ecey Chelswu-wa ssawessta.
   I TOP yesterday with fought
   'I fought with Chelswu yesterday.'

   b Chelswu-nun ecey na-wa ssawessta.
   'Chelswu fought with me yesterday.'

(253)a #ne-nun Chelswu-lul mannassni?
   you TOP OM met
   'Did you meet Chelswu?'

   b #Chelswu-nun ne-lul mannassni?
   'Did Chelswu meet you?'

(254)a na-nun Yenghi-wa kyelhonhanta.
   I TOP with marry
   'I am going to marry Yenghi.'

   b #Yenghi-nun na-wa kyelhonhanta.
   'Yenghi is going to marry me.'
The verbs in (252-254) are all reciprocal verbs, so that na or ne, on the one hand, and Chelswu and Yenghi, on the other, are on equal terms as potential topics. Note that the (a) sentences, which have speech role topics, are more natural than the (b) sentences, which have third person topics. Thus, we propose the following hierarchical relation for Korean:

\[(255) \text{Speech roles } \geq \text{Third person}\]

Again, this hierarchy receives pragmatic motivation: when the speaker describes an event in which he himself is involved, he naturally does it from his point of view rather than someone else's.

The last topicality hierarchy we will discuss is related to the possessive relationship between two NP's. It seems that the possessor NP (more correctly, the referent denoted by the possessor NP) is higher in topicality than the possessed NP. Therefore, it is normally the possessor, not the possessed, that is chosen for the topic. Let us assume that the speaker has the following message:

\[(256) \text{khuta (khi, Chelswu)}\]

\[\text{tall height}\]

where khi and Chelswu are in a genitive relationship. Since the proposition is generic, it must be realized syntactically in a topic-comment structure. In this case, the possessor NP Chelswu is the better candidate for topic:
(257)a Chelswu-nun khi-ka khuta.
'As for Chelswu, his height is tall.'

b khi-nun Chelswu-ka khuta.
'As for height, Chelswu is tall.'

(257a) is more natural than (257b). Thus, we propose the following hierarchy:

(258) The possessor \( \geq \) The possessed

This hierarchy also seems to be related to the nature of the topic. One of the functions of topic is to identify an entity in order to talk about it. Thus, topic must be a referential entity. When two NP's are in a genitive relationship, the possessor usually has autonomous reference, but the possessed NP is usually defined with reference to the possessor. In other words, the latter does not have autonomous reference but is dependent upon the possessor NP for identification. Therefore, the possessor NP is easier to talk about than the possessed NP.

To conclude this section, the linearization process involved in PS rule (189a) is constrained by such topicality-related hierarchies as we discussed above. Secondly, the greater the number of topicality hierarchies that an NP is involved in (assuming that it is at the top of the hierarchies), the stronger its topicality becomes and thus the greater its eligibility as topic. For example, the first person animate agent that is discourse-anaphoric may be strongest in topicality. Consider
(259)a Speaker A: ne-nun mwues-ul hallay?  
you TOP what OM do-will    
'What are you going to do?'

b Speaker B: ____________________.

If Speaker B wants to say the following after (259a):

(259)b pangmwunhanta ( na, Yenghi, Chelswu(-hako) ) 
visit I with

he will normally choose na 'I' as the topic rather than other arguments in (259b).

3.2.4 In the PS analysis, topic has two functions. One is that topic, as a referring expression, identifies an entity from the context as the subject of the discourse. Topic thus chosen is assumed to be put under the higher NP node in our analysis. We may call this the "referring function of topic." This has been our major concern in the preceding section.

The other function of topic is to set up the domain or framework under which the main predication holds. Therefore, the topic affects the grammatical organization of the following comment sentence either directly or indirectly. In this section we will discuss this "grammatical function of topic," i.e., how the topic is related to the following comment.

In the transformational analysis, topic is always syntactically linked to the comment sentence, because topic is always a grammatical constituent of the comment. There is,
thus, nothing particular to say about the relationship between the topic and the comment. In the PS analysis, however, topic is established independently of the following comment, even before the latter is formed. Therefore, if there is no constraint between the topic and the comment, we have to allow the generation of sentences like (260-261):

(260) *hanul-un pap-mas-i cohta.
sky TOP appetite SM good
'As for the sky, the appetite is good.'

(261) *hwahak-un nalssi-ka alumptapta.
chemistry TOP weather SM beautiful
'As for chemistry, the weather is beautiful.'

which are wildly unacceptable.

It is clear that the relationship between the topic and the comment cannot be constrained syntactically, because there is no grammatical relationship between the two in the case of 'extrinsic' topics. We suggest that the relationship between the topic and the comment be characterized in terms of the degree of semantic distance between the two.

In the case of intrinsic topics, the topic is grammatically tied to the comment, so that the relationship between the two can be shown by the use of a variable in the comment, as Gundel (1974) has proposed:

the book NP wrote

However, this kind of syntactic binding is not possible for extrinsic topics such as the following:
The relationships between the topic and the comment in (263-264) are simply semantic. The semantic distance between the two is quite close, because the involved semantic relations are clearly definable in such ways as "class-member", "owner-owned", and so on.

Now, let us consider another set of extrinsic topics:

(265) nakssi-nun Kim-sensayngnim-ul mosica.
    fishing TOP Mr.Kim OM invite
    'For fishing, let's invite Mr.Kim.'

(266) khophi-nun cam -i an wa.
    coffee TOP sleep SM not come
    'For coffee, I cannot sleep.'

    TOP SM expensive
    'As for New York, Tiffany's is expensive.'

What relates the topic to the comment in (265-267) is not semantic; it is rather some pragmatic presupposition that is shared by the discourse participants. (265) is acceptable only if it is understood, for instance, that Mr.Kim is an authority on fishing; (266) is acceptable when it is understood that coffee can disturb some people's sleep; and the plausibility of (267) is based on the real world knowledge that Tiffany's is located in New York. We should say that the semantic distance between the topic and the comment in
(265-267) is rather great.

When the semantic distance between the topic and the comment is so great that the addressee cannot see any connection between the two, the sentence becomes unacceptable. This is the case in (260-261), where 'the sky' and 'the appetite' on the one hand, and 'chemistry' and 'weather' on the other seem totally irrelevant to each other. There is no semantic or pragmatic relation whatsoever to connect the two. Hence, the unacceptability of (260-261).

Thus, to constrain the relationship between the topic and the comment, we adopt the following Relevancy Principle from Gundel (1974) which was originally proposed by Strawson (1964):

In order for a comment, S', to be successfully predicated of a topic, NP₁, NP₁ must be of a type or category such that it is logically possible for S' to be true or false of NP₁. (Gundel 1974:66)

3.2.5 It has been known that the particle nun has two functions: marking the topic of the sentence (in unstressed S-initial position) and expressing the meaning of "contrastive focus" (stressed or elsewhere). It is legitimate to ask what is the relationship between the form and functions of the particle nun.

D.W. Yang has argued that "the difference between topic and contrastive focus seems to be primarily the presence and absence of the contrastiveness, and the other factors seem to remain the same" (1975:118). But the bipartite functional
division of the particle nun into topic marking, which is assumed to be non-contrastive, and contrastive focus is incorrect. The nun which unambiguously marks a topic can imply contrast in appropriate contexts. For example, let us assume that I am in a restaurant. The waitress comes to ask what I am going to order. If I were alone at that time, I would say:

(268)a  nayngmyen cwuseyyo.
       cold noodles give
       'Give me cold noodles.'

(268a) is derived apparently from (268b):

(268)b  na-nun nayngmyen cwuseyyo.
       'For me, give me cold noodles.'

The topic na-nun has been deleted because it is so obvious. However, if I were with friends at the time, I would probably say:

(269) na-nun nayngmyen cwuseyyo.
      'For me, give me cold noodles.'

Note that the topic na-nun in (269) is clearly contrastive but the same topic in (268b) is not.

The only significant difference between (268b) and (269) is that, in the latter, I am conscious at the time of utterance of my friends who might want to order different dishes, and in the former I am not. If I were thinking about something else and temporarily forgot the existence of my friends, I might have uttered (268a) rather than (269), even
though I am with my friends physically and am ordering only for myself. Thus, (268-269) clearly show that the topic can be contrastive or non-contrastive and the "only consistent factor [for contrastiveness] seems to be that the speaker assumes that a limited number of candidates is available" (Chafe 1976:34) in his mind.

Let me give one more example. Let us assume there is a father with two children, Chelswu and Yenghi. Consider the following discourse.

(Father: nehuy-tul-un mwues-ul hallay? )  
you Plural TOP what OM do-will  
'What are you going to do?'

(270)a Chelswu: na-nun kongpwuhallayyo.  
I TOP study-will  
'I will study.'

b Yenghi: na-nun nollayyo.  
I TOP play  
'I will play.'

The topics na-nun in (270a-b) are clearly contrastive. If the same father had only one boy, Chelswu, and asked the same question, Chelswu would probably say:

(271)a kongpwuhallayyo.

which is derived from (271b):

(271)b na-nun kongpwuhallayyo.

Again, note that the topic na-nun in (270) is clearly contrastive and the same topic in (271b) is not.

Therefore, we claim that contrastiveness is a property
shared by the topic marking nun and the contrastively focal nun. The former picks up contrastiveness whenever there are more than one candidate in the mind of the speaker (or in the universe of discourse) for the role taken by the topic NP-nun.

Though the topic nun can also imply contrastiveness, the two nun forms do not have exactly the same contrastive meaning. We will first make explicit the difference between the two nun forms and then present a hypothesis about the relationship between the two.

Assume that the speaker has the following message:

(272) wassta (pom )
came spring

The proposition (272) normally has only a 'specific' interpretation, and so it is expressed in a 'topic-less' sentence:

(273)a pom -i wassta.
'Spring has come.'

However, if the speaker utters (273b) using nun instead of ka

(273)b pom -un wassta.

then (273b) normally takes the meaning of contrastive focus. We assume that nun is used as a delimiter in (273b) and pom-un is a contrastive focal topic (see 2.4.5).

Sentence (273b) asserts that spring has come, but it
also implies that something else has not come. For instance, (273b) may imply that the speaker has the following statements in mind:

(274) a kulena nim-un o-ci an-nun-ta.
    but lover come not
    'But my lover has not come.'

b kulena ceypi-nun tola-o-ci an-nun-ta.
    but swallow come-back not
    'But the swallows have not come back.'

Note that the presuppositions (274a-b) carry a negative an. The delimiter nun in (273b) implies that, among the members registered in this particular universe of discourse, only pom 'spring' has come and the other members of the set has not come. Thus, (273b) may be represented as follows:

(275) wassta (pom) .∞(wassta (∞pom) )

Thus, we argue that nun as a delimiter entails the negation of the predicate of the sentence with regard to the complementary members of the set, of which the NP-nun is a member.

It might be argued that the meaning of 'negation' is not necessarily implied, because the speaker of (273b) may presuppose statements like the following:

(276) a kulena acik nalssi-nun chwup-ta.
    but yet weather cold
    'But the weather is still cold.'

b kulena nwun-i acik ssahye-issta.
    but snow yet piled is
    'But the snow is still piled up.'

where there is no negative morpheme. Even in these cases,
however, it seems that the speaker expected that the weather would be warm or that the snow would melt away, but contrary to his expectations, such things did not happen. Thus, even in these cases, sort of negative meanings contrasting with the actual assertion is implied. The sequence of (273b) and (276a-b) might be represented as (277a-b):

(277)a wassta (pom) \sim ttattushata (nalssi)
came spring warm weather

b wassta (pom) \sim nokassta (nwun)
came spring melted snow

Therefore, we argue that \textit{nun}, used as a delimiter, always implies the negation of a certain predicate (e.g., \textit{wassta} in (275) or \textit{ttattushata/nokassta} in (277a-b)), with regard to the complementary members of the set.

In contrast, the \textit{nun} used as topic marking does not carry such a negative implication as the delimiter \textit{nun} does. For example,

(278)a Yenghi-ka eti-ey kasseyo?
SM where to went
'Where did Yenghi go?'

b Yenghi-nun tosekwan-ey kasseyo.
TOP library to went
'Yenghi went to the library.'

the topic \textit{Yenghi} of (278b) does not seem to imply contrastiveness at all. This seems to be so because (278a) mentions only \textit{Yenghi} and so \textit{Yenghi} is the only member registered in this particular universe of discourse. However, the mentioning of \textit{Yenghi} in (278a) may bring into the
consciousness of the speaker of (278b) another entity somehow related to Yenghi on pragmatic grounds, a brother or sister perhaps. Then, the speaker of (278a) may add the following (278c) after (278b):

\[(278)c \text{ kulena Chelswu-nun kukcang-ey kassta.} \]
\[\text{but TOP cenema to went} \]
\[\text{'But Chelswu went to the movies.'}\]

where Chelswu is pronounced without stress. Now, the topic of (278c), Chelswu, is clearly contrastive (e.g., with Yenghi). This is because there are two members registered in this particular universe of discourse, Chelswu and Yenghi, and the speaker of (278c) is referring only to the former. Note, however, that there is no implication either in (278b) or (278c) that the speaker is negating a certain predicate (e.g., kassta 'went') with respect to the complementary members of the set registered in this universe of discourse.

Therefore, we argue that the contrastive meanings implied by the topic nūn and the delimiter nūn may be characterized in terms of the absence and presence of the meaning of 'negation' of the predicate with regard to the complementary members registered in the universe of discourse. We also maintain that this meaning difference between the two nun forms is due, not to the absence or presence of 'contrastiveness', but to the absence or presence of 'focus'. In other words, the lack of an explicit negation accompanying topic nūn is due not to the lack of 'contrastiveness', but to the loss of the informational focus carried by the
contrastive focal nun.

The remaining question, then, is why the focal meaning of nun is lost when used to mark the topic in S-initial position? For this, we will present a hypothesis based on the historically attested development of the subject particle ka.

But first we would like to point out a parallelism between the two particles nun and ka. Like nun, ka also has two functions: it marks the subject of the sentence, and it expresses so-called "exclusive focus." But this parallelism has been obscured by the assumption that ka is always a mere structural marker, a marker only of subject position, restricted in occurrence and without meaning. Even when ka is understood as implying an exclusive focus, it has been discussed as part of the subject function. For example, Kuno's (1973) examples in Japanese are all of this sort.

(279) kolay-ka phoyu-tongmwl-ita.
whale mammal be
'The whale is a mammal.'

(280) John-i na-uy chinkwu-ita.
I of friend be
'John is my friend.'

What we would like to point out is that the particle ka is more than a subject marker; it actually behaves more or less like a delimiter. There are two pieces of evidence for this, one synchronic and the other comparative. Synchronically ka is not limited to the subject in distribution.
It should be clear that ka does not mark the subject of the sentences in (281-286). Note that it need not even be limited in occurrence to nominals. Therefore, like nun, the particle ka can also function as a delimiter, i.e. a constituent focus marker.

As for the comparative side, Kim (1956) notes a phenomenon in other Altaic languages which is similar to the use of the particle ka in Korean. In Manchu, the subject particle inu, which is the cognate with Korean i (the older form of ka), can be used as an adverb to modify another adverb or a verb.

(287) inu damu duin tanggu cooka gamaha. really only 4 100 soldiers led
Thus, the subject particle *INU* in Manchu also seems to behave like a constituent focus marker. Based on these facts, we will assume that both *KA* and *NU* behave similarly in the sense that both of them function as delimiters on the one hand and function as respective markers of the subject and the topic on the other. 37

If the particle *KA* can function as a delimiter, it should have a meaning because delimiters usually have their own semantic content. It has been assumed that the particle *KA* has the meaning of 'exclusive focus' (Im 1974; D.W. Yang 1975), but it seems that *KA* is basically a "determinative" particle—it "identifies" or "designates" an object with emphasis for a given assertion. Consider

(289) Chelswu-*ka* pay-*ka* aphuta.
    stomach SM aching
    'Just Chelswu, he has a stomach ache.'

where the underlined *KA* has the so-called exclusive focal interpretation. Let us assume that there are three people, John, Mary, and Chelswu in this particular universe of discourse. What the speaker of (289) is asserting is that it is Chelswu that has a stomachache. The implication that the others do not have a stomachache is not what is being asserted here; the others are just of no interest to the speaker of (289) in this regard. The implication is just an invited inference or contingent meaning, not an asserted one. We might
say that the speaker is just identifying Chelswu, with emphasis, as the person who has a stomachache.

This point will be clear if we compare ka with another delimiter man. Compare (289) with (290) following:

(290) Chelswu-man pay-ka aphuta.

What the speaker of (290) asserts is that Chelswu has a stomachache and also that the rest of the people in the universe of discourse do not have one. The delimiter man has 'exclusiveness' as its asserted meaning.

Note that the particle ka in (281-286) has no exclusive meaning, either. It just has the function of a strong and emphatic identification. Thus, we argue that the meaning of the particle ka is 'determinative'.

Now, returning to the main issue of this section, what is the form-function relationship of the two particles ka and nun? Particularly, why is the focal meaning of nun lost when it is used to mark the topic in S-initial position? For this question we suggest the following hypothesis. It seems that the two particles were originally used as constituent focus markers, i.e. to 'identify' or 'contrast' any constituent of a sentence with emphasis. Then, there may have arisen among the speakers a psycholinguistic desire or need to formally codify the subject ([NP]ₘ) and the topic ([NP]ₜ). This hypothetical linguistic process is quite understandable. First, the subject and the topic are the two most important syntactic roles in the sentence and their functional loads
are very high. Second, Korean allows three NP’s to occur before the main verb; so if these syntactic functions are not surface-marked, there might be some confusion. The surface-coding of these syntactic functions may have been initiated for the effectiveness of communication. 38

Later, perhaps, the particles ka and nun came to be chosen for marking the subject and the topic respectively, the choices being motivated on semantic grounds. The subject has the function of identifying or designating an entity with regard to the predicate. ka with the determinative meaning must have been the optimal choice, because ka is semantically neutral without any 'inclusive' or 'exclusive' biases. It actually has the very similar function of 'identifying' or 'designating'.

Similarly, the choice of the particle nun for topic marking must also have been appropriate. As Kuroda (1969) has correctly pointed out for Japanese wa, the basic meaning of the delimiter nun is to assert that the speaker is committing himself to the validity of the assertion only with regard to the entity preceding it, but not to the other possible candidates to which the same assertion can apply. That is, the speaker asserts that he is only concerned about the entity he has in mind. On the other hand, the function of topic is to "pick out, from all the [possible] referents that might be categorized in that way [i.e. by what is being asserted in the comment], the one [the speaker] has in mind" (Chafe 1976:39). Thus, the basic meaning of the delimiter
nun is semantically very close to or almost identical to that of the topic.

After the particles ka and nun were associated with the two syntactic positions, their occurrences in these positions may have come to be grammaticalized via a process of semantic bleaching or fading. Since the subject need not always represent the informational focus of the sentence, gradually the particle ka must have lost its original focal meaning. By the same token, the topic need not be, and often is not, the informational focus of the sentence; so, nun also came to lose its original focal meaning. As the grammaticalization process became more and more perfect, the two particles may have been fossilized eventually as syntactic position markers. On the other hand, in other non-subject and non-topic (or stressed) positions, they may have maintained their original meanings of 'determinative' and 'contrastive' focus, respectively.

It would be no easy task to confirm such speculations. But there is a historical fact that provides some support. Kim (1956) points out that the subject particle ka (in its older form i) was always pronounced with the high tone (sang-seng) in Middle Korean. Kim speculates that when the particle i came to be used as the subject marker, it must have carried the meaning of "emphasis", which seems to be similar to what we mean by 'determinative focus', and he argues that this focal meaning came to be lost later.

Unfortunately, however, we have found no such
historical support for our speculations concerning the particle nun. But it would not be unreasonable to hypothesize a similar developmental process—the loss of the focal meaning came about through some kind of semantic bleaching while the delimiter nun became grammaticalized as a marker of the syntactic topic. In any case, this hypothesis is much preferable to another possible hypothesis such as Kuno's (1973)—that the particle nun is flatly a contrastive particle—because the latter is not compatible with the obvious fact that nun is a member of the set of delimiters, which are constituent focus markers.
Consider the following sentences:

(i) Speaker A: What did John give to Bill?
(ii) Speaker B: John gave a BOOK to Bill.

There are two problems with Chomsky's proposal. First, Chomsky predicts that the topic (presupposition) of (i) is 'John gave what to somebody', but this is a self-contradiction because the presupposition itself contains a focus 'what'. Second, Chomsky predicts that (i) and (ii) must share the same presupposition because the latter is a natural response to the former (cf. Chomsky's statement quoted before). But (i) and (ii) do not share the same presupposition.

Presupposition is a blanket term covering a wide range of linguistic phenomena (see Morgan 1973a; Kempson 1975; Fillmore and Langendoen 1971), some of which have nothing to do with the notion of 'topic'.

This may be acceptable for English, where topic is not surface-marked.

We will ignore marginal periphrastic expressions such as 'as for'.

Ree (1969) for assuming that (42b) underlies (42a).

The observation that the topichood of Chelswu in (39a-42a) is less clear than that of Chelswu in (39b-42b) seems due to the following. In 2.4.1 we will argue that the intuitive feeling of aboutness we get from the topic comes mainly from its "higher NP" position. The topichood of an NP is also affected by its surface word order, and S-initial position is highly significant. Thus, when an NP appears in S-initial higher NP position, its topichood is clearest.

In (39a-42a), the Preposing rule has placed a non-topicual element before the topic, thus creating a surface constituent structure like (i)

(i) 

In (i), the topic Chelswu no longer appears in the surface S-initial position, and this is why its topichood has been
a little bit obscured and is unclear.

7 This argument may not apply to Yang (1973), because a comma, which he uses to show the topic of a sentence, does not occur after Yenghi.

8 Note that 'topic' here is not meant to be a linguist's made-up notion that what is 'given' in the prior discourse (e.g., pav-ka kophuta in (45)) is what the sentence is predicated about, i.e. the topic. Our notion of 'topic' is based on the native speakers' unsophisticated intuition that (45) is felt to be about Chelswu, whatever its informational status is.

9 Note that we are assuming that the particle ka is functioning as a delimiter in (45). It seems that, just like nun, ka has dual functions: it is a subject marker and a delimiter. See 3.2.5 for details.

10 However, later syntactic operations such as Scrambling or Deletion may result in the topic not appearing as such in S-initial position in the surface form. We have already discussed some relevant examples in 2.2, where sentences like (39a-42a) and (43b), which have a non-topical element in the surface S-initial position, were argued to have been derived from the corresponding topic-initial underlying structure by Scrambling or Deletion.

We will give one more example here. Consider

(i) Chelswu-man kwiyeweha-sye, wuli sensayngnim-un.

'He likes only Chelswu, our teacher.'

Instead of having the usual function of identifying an entity to talk about, the S-final topic in (i) has the function of adding the topic as an after-thought. Apparently, the speaker did not mention the topic due to its assumed obviousness, but later decided to add it for clarity. Then, it is perfectly justified to assume that the S-final topic in (i) was originally in S-initial position and was removed from that position by a later Topic Postposing (or Right dislocation) rule, which is part of Scrambling. If the topic is a surface concept, as in the syntactic definition, the topic of (i) would be Chelswu-man. But this is clearly wrong.

11 One can argue that English has the pattern (NP (SVO)S), as Gundel (1974) and Chomsky (1977) seem to. and that the topic function is carried by the initial NP in English. I would accept this argument. However, it is also true that (NP (SVO)S) pattern is not as basic as (SVO)S. That is, I assume that the so-called left-dislocated sentences using such periphrastic expressions as 'as for', 'concerning', etc. are relatively minor in English.
The claim that topic must occupy the pre-subject NP node does not necessarily mean that any NP which occupies this position is a topic (see 2.4.5).

Yang's "macro-micro" analysis is inadequate for dealing with double subject constructions in general. We will point out several problems here. For other problems, see Im (1974). First, Yang's macro-micro relation is based on the same case relation, but in the following sentences, the case relations of the macro and micro NP's are not the same:

(i) John-un/i Mary-ka cohta.
   "John likes Mary"

(ii) na-nun/ka apeci-ka aphi-ki-ta.
   "For me, Father is sick."

Second, Yang's constituent structure for (46a) should look like the following:

(iii) \[
\begin{array}{c}
S \\
\downarrow \\
NP \\
\downarrow \\
\text{Chelswu} \\
\downarrow \\
cohta \\
\end{array}
\]

\[
\begin{array}{c}
\text{NP} \\
\downarrow \\
\text{NP} \\
\end{array}
\]

which is incorrect.

Third, there are many macro-micro sentences that are not acceptable:

(iv) *mikwuk-i Texas-ka khuta.
   America big

(v) *Hankwuk-i Seoul-i cohta.
   Korea good

Note that if the macro NP is followed by nun, (iv-v) are acceptable.

Co-occurrence with the topic announcer is a necessary, but not a sufficient, condition for an element to qualify as topic. That is, the fact that an element can co-occur with the topic announcer does not necessarily mean that the element is the topic of the sentence, as in (i):

(i) Chelswu-ey tayhaye malhantamyen, na-nun ku-ka cohta.

Sometimes the actual identity of the individual may have been unspecified, because the speaker thought that it
it is no import in the particular communication, even though he knows it. This is the so-called "partially consummated reference" (Searle 1969:82-89; Magretta 1977:87-88), which means that a less than unique description is given by the speaker in a sloppy manner. Indefinite NP's of this nature can be topics. Consider

(i) nay-ka ecey masi-n etten pwulkun swul-un tokhata.  
I SM yesterday drank a red liquor TOP strong  
'A red liquor that I drank yesterday was strong.'

(ii) nay-ka a-nun etten salam-un sako-lo cwukessta.  
I SM know a man TOP accident died  
'An acquaintance of mine died in an accident.'

(iii) nay chinkwu hana-nun New York-ey santa.  
my friend one TOP at live  
'One of my friends is living in New York.'

(iv) kutul cwung-uy han myeng-un ton-ul hwumchyessta.  
they among of one TOP money stole  
'One of these guys stole money.'

Note that the topics in (i-iv) are all indefinite specifics. But the referents of these topics are definite at least to the speaker, and he can provide a complete and unambiguous identification of them to the addressee on demand. Thus, we argue that 'definiteness' must be understood in terms of the speaker's point of view.

There are two other surfacy counter-examples to our claim; one is generic NP's, which may be said to be indefinite non-specifics, and the other is the existence of the implicational definites. We do not have space here to discuss them in detail. For the generics, see Kuroda (1965:73, 1972: 167) and for implicational definites, see the detailed discussion of Allerton (1978).

This argument about (67a-b) is due to C.M.Lee (1973: 94-99). Note that the 'subject' condition is not relevant to (67a-b) because it is the subject itself that has undergone reflexivization.

Also note that case-marked NP's cannot co-occur with topic announcers, another piece of evidence that they are not topics.

(i) #Chelswu-ekey-ey tayhaye malhantamyen, na-nun (ku-to if-we-talk-about I TOP he ekey) ton -ul cwuessta.  
to money OM gave  
'If we talk about 'to Chelswu', I gave him money.'
Case markers express certain relations. The main function of topic is to identify a referent, but case-marked NP's serve other than a purely referential and identifying function, particularly when they are semantically non-empty (e.g., wa 'with', pote 'than', etc.). This seems to be why case-marked NP's cannot be topics. Thus, we assume that the underlined case-marked NP's in (iii-iv) are not topics but simply fronted NP's:

(iii) Chelswu-evkey na-nun ton-ul cwuessta.
    To Chelswu, I gave money.'

(iv) i khal-lo na-nun namwu-lul peyessta.
    with 'With this knife, I cut the tree.'

For example, the underlying structure of (10lb) looks like (i):

(i) \[ \text{[[kwankwang]}_{\text{TOP}}[[\Delta]\text{NP}[\text{Cewuto-lo cenghayessta}]_{\text{VP}}]} \]

where the empty node is the subject.

However, as I.S. Yang (1973) has correctly pointed out, it can be purely pragmatic.

What I cannot understand with the topic marking nun and the delimiter nun is that if the two constitute a single formative nun, which I believe is true (see 3.2.5), both of them should not co-occur with indefinites because both of them seem to share the same assertive meaning. But nun can co-occur with indefinites and nun cannot (see 2.4.1.2). This may be so because nun implies more than nun (see 3.2.5). Anyway, it is still a hypothesis that nun does not have the 'aboutness' feature.

Some speakers of Korean may accept only NP-nun as topic.

Though presuppositional semantics is preferred over the strict truth-functional semantics, both theories of meaning are far short of representing the full linguistic meaning of a sentence (cf. Morgan 1973a; Shopen 1972b; Strawson
1964). For example, consider the following sentences:

(i) They are parents.
(ii) They are not children.
(iii) Next week we will be visiting relatives.

In both theories of meaning, (i) and (ii) are identical in meaning, but it is clear that they are different in linguistic meaning. There are many examples like (i-ii), which involve such lexical items as buy/sell, defeat/lose, etc. (iii) has two linguistic meanings, depending upon the grammatical status of visiting; but truth-functionally there is only one, because 'a visiting relative' is always 'visiting a relative'.

See Strawson (1964) for other advantages of presuppositional semantics and Kempson (1975) for its disadvantages.

Tonoike’s analysis is actually quite arbitrary, because there is no reason why the experiencer should be placed before the patient in deep structure. If the experiencer occurs after the patient in deep structure, his analysis does not work. As for the locative, Kuno (1973) argued that it might occur before the patient in deep structure on semantic grounds. However, it is an arbitrary syntactic proposal that any 'NP + particle' occurs before the patient in deep structure.

What we are claiming here is that the truth value of a sentence is ultimately context-sensitive. The same position was upheld by Strawson (1964) and G. Lakoff (1971a). There have been opposing views. For partially opposing views, see Morgan (1973a) and Dretzke (1972). For total opposition, see Kempson (1975) and Katz and Langendoen (1976).

The controversy about the justifiability of the passive transformation is the typical case. Also see ft. 22.

What we mean by 'focus' here is not "informational focus" widely used in topic-comment theories (e.g. Gundel 1974; Chomsky 1972) and in the informational definition of topic, but simply "the speaker's center of attention."

This distinction is similar to Kuroda's (1972) "thetatic" and "categorical." Our discussion in this section is influenced by Kuroda's (1965, 1972, 1976) analyses.

However, when the 'specific' proposition contains a definite entity, as in (194a), it is possible to describe the event with regard to the definite entity. This is so because when there is a definite entity, people have a tendency to
describe the event with regard to the entity rather than objectively. Compare (194a) with (207b). This is why we argued that definiteness contributes to topicality in 2.4.5. The same thing is observed in (198a), which is to be compared with (198b) on p.293.

There have been three other alternative explanations about the distinction between NP-ka and NP-nun in sentences like (198a-b)

One is Kuno's analysis, where NP-nun is assumed to be "old" informationally and NP-ka is "new" (Kuno 1972). As we will see later in 3.2.3.2, topic is usually given or old in informational status within the discourse. But this explanation often fails when we are dealing with discourse-initial sentences.

Though we put (198b) in a discourse context in which Chelswu has already been mentioned, the same sentence can occur as a discourse-initial sentence, as in (i):

(i) #emma! Chelswu-nun ecey cwukesstayyo.

There is no difference between the two Chelswu in (198a) and (i) in informational status. Kuno might say that the relationship of Chelswu with regard to the predicate is 'old' in (i) but it is 'new' in (198a). But there is no doubt that this argument is not convincing. Consider the following sentences:

(ii) #ikes-i elma-yo?
   this SM how much
   'How much is this?'

(iii) #ikes-un elma-yo?
   TOP

Note that ikes in (ii) and (iii) are equal in informational status. This is why I do not accept Kuno's analysis, though his theory is correct in part.

Another explanation is 'lexical', i.e. that NP-nun is 'contrastive' and NP-ka is not. This analysis has an intuitive plausibility in distinguishing sentences like (198a-b) or (ii-iii) above. However, the explanatory power of the lexical analysis is quite limited. For example, 'generic' propositions must have the topic, but note that there is no reason why they cannot be expressed 'non-contrastively'. By the same token, there is no reason why sentences like (208b-211b), which we will discuss shortly, cannot be expressed 'non-contrastively'. Furthermore, all the facts which we will discuss in 3.2.3 cannot be explained by the lexical analysis where the distribution of nun and ka is explained solely in terms of 'contrast' or 'non-contrast'.

The third explanation is D.W.Yang's (1974). He argues that NP-nun and NP-ka are always "generic" and "specific"
respectively. His analysis is based on a controversial definition of 'generic' and 'specific'. For example, Yang argues that (ii) is 'specific' and (iii) is 'generic'. This is very difficult to justify. His notion of 'generic' depends upon the presence or absence of mental reflection on the part of the speaker, which is very vague.

This phenomenon seems to be related to what Keenan and Schieffelin (1976) called "topic collaboration," a common phenomenon in question-answer dyads.

A full-scale justification of this hierarchy in Korean is beyond the scope of this study. For noun phrase accessibility in Japanese, see Inoue (1975).

That is, when the sentence is not affected by the previous discourse. This is why we put the marker # before the sentence.

It is not possible to say about spring that it has the property of 'having come'.

However, the S-medial delimiter nun is not the topic in our analysis:

1. na-nun pom-un silhta.
   TOP spring dislike
   'I dislike at least spring.'

where pom-un is not the topic.

Note that the adverb acik is a negative-polarity word. If acik is substituted for by the positive-polarity counterpart peisse, sentences (276a-b) become unacceptable.

H.M. Sohn (personal communication) has correctly pointed out that the distinction between the two functions of the particle ka is not always clear.

1. ceki Chelswu-ka kanta.
   there SM go
   'There goes Chelswu.'

2. Yenghi-ka pap-ul mek-ko issta.
   SM rice OM eat  Progressive
   'Yenghi is eating rice.'

3. Chelswu-ka papo-ita.
   fool be
   'Chelswu is a fool.'
(iv) (nwu-ka wasni?) Chelswu-ka wassta.
    who came came
    'Who came?' Chelswu came.

(v) Chelswu-ka pay-ka aphuta.
    stomach aching
    'Just Chelswu, he has a stomachache.'

(vi) Chelswu-ka seltukha-ki-ka swipta.
    persuade to SM easy
    'Just Chelswu, he is easy to persuade.'

(vii) Chelswu-nun wenlay-ka papo-ita.
    by-nature fool be
    'Chelswu is a fool by nature.'

Ka in (i-ii) is used clearly to mark the subject, and ka in (vii) is used as a delimiter. As for ka in (v-vi), we argued that it is not the subject marker because the NP-ka is not the subject but the topic. We assumed that ka in (v-vi) is used as a delimiter. But what about the ka in (iii-iv)? It is not clear whether it is used as the subject marker or as a delimiter, probably as both.

38 Altaic languages originally did not mark the subject. The use of the subject marker is a later development.

39 "Grammaticalization" means the shift of an element from being a full root with an independent meaning to being a mere grammatical marker of some kind, e.g. to being affixes, postpositions, particles, and so on. (cf. R. Anttila. An Introduction to Historical and Comparative Linguistics (1972, Macmillan.) pp.149-152)
CHAPTER IV
RELATIVE CONSTRUCTIONS

1. Introduction

1.1 Relativization is one of the well-described areas of Korean syntax (Cook 1968; Ree 1969; I. S. Yang 1972; D. W. Yang 1975; H. B. Lee 1975). However, all of these previous analyses have been concerned with simple relative constructions with a single embedded sentence. In this chapter, we shall deal with both simple and multiple relative clauses, and we will describe relative clauses within the framework of topic-comment structure.

The chapter is organized into three sections. Section 1 is an introduction in which we point out some controversial issues in Korean relativization and outline the basic syntactic rules involved in the formation of relative clauses, both simple and multiple. After claiming that the relative head NP is the topic of the underlying embedded sentence, we examine in Section 2 the structures of relative clauses in terms of topic-comment articulation. A hypothesis is presented that relative clauses are a syntactically compressed or frozen version of text. Section 3 reviews Kuno's argument for the relationship between thematization and relativization based on particle deletion. We propose that Kuno's formulation involving particle deletion should be replaced by a recoverability condition.
1.2 We start with a brief sketch of relative clauses in Korean. First, the relative clause precedes its head on the surface. Second, there are no overt formal relative markers (e.g., who, which, where in English) which indicate the boundary of a relative clause. What happens instead is that the forms expressing the sentence modality in a relative clause (e.g., ssta in sentence (1)) are adjusted (e.g., -n in sentences (2a-c)) to indicate that the clause is not an independent sentence (I.S. Yang 1972 for details). Third, unlike English, where the preposition of a relativized NP is either pied-piped or remains orphaned in its original position, the postposition of a relativized NP is simply deleted along with the noun phrase. These three characteristics are shown in the following example:

(1) i pwun-i (nay-ka ku salam-ey tayhaye malhay-ssta) salam-ita. [Underlying structure]
    Past Decl man be
    'This is the man (I talked about the man).'

(2)a *i pwun-i (nay-ka malha-n ø-ey tayhaye) salam-ita. Adj. Mod.
    'This is the man about whom I talked.'

b *i pwun-i (nay-ka ø-ey tayhaye malha-n) salam-ita. Adj. Mod.
    'This is the man whom I talked about.'

c i pwun-i (nay-ka ø malha-n) salam-ita. Adj. Mod.
    *'This is the man I talked.' [Surface forms]

There are two controversial issues in the analysis of Korean relativization. One is whether relativization is a
movement process (Ross 1967; Muraki 1970), a process of pro-
formation followed by pro-deletion (Sanders and Tai 1969; D.
W. Yang 1975), or merely a simple deletion (Cook 1968; Ree
1969; I.S.Yang 1972; H.B.Lee 1975). The second issue is
whether the relative clause precedes or follows the head NP
in underlying structure. Cook (1968), Muraki (1970),
and D.W.Yang (1975) have argued that, in underlying struc-
ture, the relative clause follows the head. We might call
this a "post-nominal" analysis. Others have argued for a
"pre-nominal" analysis--the relative clause precedes the

We consider relative clause formation rule in Korean
to be a simple deletion process. The implausibility of the
movement analysis has been discussed by D.W.Yang (1975:148-
154). Moreover, the movement analysis does not fit well with
has noted that the absence of a rule moving relativized con-
stituents correlates with the position of the clause in
front of the modified noun phrase. Schwartz (1971) points
out that the absence of a movement rule for relativized con-
stituents follows from the absence of a rule moving wh-ele-
ments in questions, this lack itself being a necessary pro-
that the lack of movement follows from the clause-final po-
sition of the complementizer.

Yang's alternative to movement--"pro-formation and pro-
deletion"--is not plausible, either. In this analysis, the
coreferential NP in the underlying embedded sentence is first pronominalized and then deleted. However, simple deletion is the basic strategy for forming relative clauses in Korean. Leaving a pronominal copy in a relative clause is an exceptional and extremely rare auxiliary device. Thus, it is difficult to specify when and why most pronominal copies should be deleted (see Section 3). As C.M. Lee correctly points out, "the grammatical cost of the obligatory-ness of deletion of the pronominal copy in undefinable vast environments and other related problems is unsurmountable" (1973:111).

Since neither the movement nor the pro-formation and pro-deletion analysis seems plausible, we will assume Korean relativization to be a simple deletion process.

As for the second issue—whether the relative clause precedes or follows the head—first note that the relative clause is a modifier of the head NP. One general, in fact exceptionless, pattern of Korean is that every modifier precedes the form modified. If the post-nominal analysis were valid, this would constitute a single exception to this general pattern. Such an analysis should be provided with very strong motivation indeed, since it goes against the grain of the language. H.B. Lee (1975) has shown the inadequacy of the justifications provided for the post-nominal analysis. While this analysis may be plausible for English, where the relative clause follows the head in the surface structure, it is difficult to justify for Korean, which is otherwise a
consistently left-branching language. We, therefore, assume the correctness of the pre-nominal analysis.

Since a relative clause precedes its head in underlying structure and since it undergoes a simple deletion of the coreferential NP, we can formulate the relative clause formation rule in Korean as follows:

(3) Relative clause formation (Deletion)

\[ SD: (W) [S X NP Y ] NP Z \]

\[ SC: 1 2 3 4 5 6 (if 3 = 5) \]

It is not necessary here to go into the detailed house-keeping rules following Deletion rule (I.S. Yang (1972) for details). Rule (3) generates simple relative clauses such as (4a):

(4a) [tulek-i pat-un] kay-ka cwukessta.
truck SM hit dog SM died
'The dog that the truck hit died.'

from their corresponding underlying structures like (4b):

(4b)

\[ S \rightarrow NP \rightarrow S_0 \rightarrow VP \]

\[ NP \rightarrow \text{tulek} \rightarrow VP \rightarrow \text{cwukessta} \]

\[ NP \rightarrow \text{kay}_1^{-ka} \]

\[ NP \rightarrow \text{kay}_1^{-lul} \text{ patassta} \]

Now, let us consider a somewhat more complex relative
clause:

(5a) [Chelswu-ka salangha-nun [tulek-i pat-un]] kay-ka SM love -ing truck SM hit-ing dog SM cwukessta.
died.
'The dog that the truck hit that Chelswu loved died.'

(5a) is an example of so-called stacked relative clauses, in which any number of relative clauses can be added to the left. Cook (1968) noted that a stacked relative clause such as (5a) is an endo-centric construction, in which the inner relative clause, tulek-i patun in (5a), is more deeply embedded than the outer relative clause, Chelswu-ka salangha-nun, in surface structure. Cook argued that the depth of embedding in surface structure should be reflected as such in underlying structure. If he is correct, the underlying structure of (5a) will look like (5b):

(5b)

```
S0  VP
    /   \\
  S1   VP
    /     \\
  NP   NP
   /     / \\
Chelswu-ka NP  kay'-ka
 /       / \   /
NP VP salanghyssta kay'-lul
    /       /   /
   VP       V   kay'-lul patassta
```
In (5b), the depth of embedding on the surface is correctly reflected. Deletion rule (3) will apply to (5b) to produce an intermediate structure (5c):

\[
(5c)_{S_1} \text{Chelswu-ka}_{S_2} \text{tulek-i patun] kay-lul salang-}
\text{hayssta] kay-ka cwukessta.}
\]

Since (5c) still meets the SD of Deletion rule (3), as the above proper analysis shows, (5c) should undergo another deletion, generating (5d):

\[
(5d)_{S_1} \text{Chelswu-ka}_{S_2} \text{tulek-i patun] salangha-nun] kay-}
\text{ka cwukessta.}
\]

The depth of embedding of the underlying structure is still maintained in (5d), but (5d) is not acceptable as it stands. This shows that Deletion rule (3) is not enough to generate well-formed multiple relative clauses like (5a). (5d) is unacceptable because the two underlined constituents of \(S_1\) are separated from each other. We apparently need another rule to place together the two separated constituents of \(S_1\) in (5d). We will call this rule "Relative clause formation (Reordering)"; and formulate it as follows:

\[
(6) \text{Relative clause formation (Reordering)}
\]

<table>
<thead>
<tr>
<th>SD:</th>
<th>(V_{S_1}(W) [S_2 X] Y] NP Z)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2+4 3 4 5 6</td>
</tr>
<tr>
<td>SC:</td>
<td>1 3 (\emptyset) 5 6</td>
</tr>
</tbody>
</table>
(Condition: (a) Term 2 can be null. (b) $S_1$ and $S_2$ are relative clauses.)

The application of Reordering rule (6) to the intermediate structure (5d) yields the correct surface form (5e):

$$(5e) [S_1 \text{Chelswu-ka salangha-nun} [S_2 \text{tulek-i patun}] \text{Kay-ka cwukessta.}$$

Note that it is not Term 3, which represents the relative clause $S_2$ created at the previous stage of derivation, that moves rightward over Term 4 in the Reordering rule (6). The previously formed relative clause does not move; instead, Term 4 (e.g., salangha-nun of (5d)) is moved leftward over Term 3 (e.g., $S_2$ in (5d)). The motivation of Reordering rule (6) and the direction of the movement in it will be further discussed in 2.2.4 and 2.2.5.

To summarize, the relative clause formation process in Korean consists of two syntactic operations: Deletion of the coreferential downstairs NP in the pre-nominal (left-branching) underlying structure and subsequent Reordering.

2. The Thematic Structures of Relative Clauses

2.1 It has been reported for many languages that relative clauses are closely related to topic-comment structures. To take an example, there are languages which only allow relativization of the topic: Malagasy (Keenan 1972) and Dyirbal. In Hittite (Justus 1976) the topic marker is formally
identical to the relative marker, which developed historically from the former. For Japanese, Kuno claimed that "what is relativized in a relative clause is not an ordinary noun phrase but a noun phrase followed by the thematic particle wa" (1973:243). In a similar vein, Gundel (1974) claimed that a relative clause is underlyingly a topic-comment structure in English. In this sub-section, we still support the Gundel-Kuno claim that a relative clause is a topic-comment structure in underlying structure. We will first review the arguments Kuno presented in support of his claim, along with another piece of evidence given by C. M. Lee (1973). Then we will present our own argument.

Kuno (1973:243-254) based his claim on four syntactic parallels between topical sentences and relative clauses in Japanese.

(a) **Parallelism in particle deletion.** Compare the following sentences (7-8):

(7a) ku maul -eyse-n John-i wassta. [topical]
the village from TOP SM came
'From that village, John came.'

b *ku maul-un John-i wassta. [particle deleted]
'As for the village, John came.'

c *(John-i o-n) ku maul 6 [Relativized]
'the village that John came (from)'

(7b) shows that the source particle eyse 'from' cannot be deleted in thematization, and, correspondingly, its corresponding relativization (7c) is unacceptable.
In (8) the goal particle can be deleted in the topical sentence; the corresponding relative clause is acceptable.

While there is a partial correlation between the two structures in particle deletion, this generalization is not entirely correct. There is some discrepancy between topic-comment sentences and relative clauses in terms of particle deletion, which is the main topic of Section 3.

(b) *Parallelism in leaving a pronominal copy*. The topic sometimes leaves a pronominal trace behind and it is exactly in these cases that relativization does the same.

(9)a (ku ai-ka cohaha-nun) kay-ka cwukessta.[topicless]
the boy love dog SM died
'The dog that the boy loves died.'

b ku ai₁-nun (cakī₁-ka cohaha-nun) kay-ka cwukessta.
'self
'As for the boy, the dog (he) himself loved died.'
[topicless]

b [. (cakī₁-ka cohaha-nun) kay-ka cwuk-un] ai₁,.
'the boy who the dog that (he) loved died.'
[relativized]

The topical counterpart of (9a) can leave a pronominal trace, as in (9b), and its corresponding relative clause with the same pronominal trace is acceptable, as in (9c). In
contrast, compare (9) with (10):

(10)a khokkili-uy kho-ka khuta. [topicless]
elephant of nose SM big
'The trunk of an elephant is long.'

b *khokkili1-nun caki-uy kho-ka khuta. [topical]
TOP self
'As for the elephant, the trunk is long.'

c *(caki-uy kho-ka khun) khokkili [relativized]
'the elephant whose trunk is long.'

Kuno argues that when the topical sentence, as in (10b), is unacceptable, the corresponding relative clause is also unacceptable, as in (10c).

(c) Parallelism in violating island constraints. Korean can freely violate island constraints; topic-comment structures and relative clauses behave the same with respect to island constraints. Compare (11) and (12):

(11)a (ku salam-i ip-un) os-i telepta. [topicless]
the man SM wore colthes dirty
'The clothes that the man wore are dirty.'

b ku salam-un (ip-un) os-i telepta. [topical]
TOP
'As for the man, the clothes he wore are dirty.'

c [(ip-un) os-i telewun] ku salam [relativized]
dirty
'the man who the clothes he wore are dirty'

(12)a John-i tul-e oko, nay-ka nakassta. [topicless]
SM enter and I SM went-out
'John came in and I went out.'

b *John-un tul-e oko, nay-ka nakassta. [topical]
TOP
'As for John, he came in and I went out.'

c *(tul-e oko nay-ka naka-n) John [relativized]
'John who came and I went out'
(12c) seems to be unacceptable because (12b) is unacceptable, i.e. because John cannot be the topic of (12a).

(d) Finally, there are some topical sentences for which no corresponding topicless sentences are found. For these sentences it is possible to construct corresponding relative clauses. This suggests that the topical sentence is the only source of relativization for these sentences.

(13a) *mwullihak-eyse chwicik-i elyepta. [topicless]
physics in employment difficult
'In physics, finding a job is difficult.'

b mwullihak-un chwicik-i elyepta. [topical]
TOP
'As for physics, finding a job is difficult.'

c (chwicik-i elyewun) mwullihak [relativized]
'physics where finding a job is difficult'

(13a) is unacceptable in its strictly locative reading; so the underlying source of (13c) can only be the topical sentence (13b).

However, this last argument must also face some counterexamples. Some of the topical sentences which have no corresponding topicless sentences cannot undergo relativization.

(14a) *sayngsen-eyse tomi-ka cohta. [topicless]
fish at snapper SM good
'Among fish, snapper is good.'

b sayngsen-un tomi-ka cohta. [topical]
TOP
'As for fish, snapper is good.'

c * (tomi-ka coh-un) sayngsen [relativized]
'fish, among which snapper is good'
(14b) is a topical sentence, where no topicless counterpart is found, as in (14a). But (14b) cannot undergo relativization.

Though Kuno's four syntactic parallels between topic-comment structures and relative clauses strongly suggest that the two structures might profitably be grouped together in a syntactic description, they do not prove conclusively that the relative head NP must actually be the topic of the underlying embedded sentence. An argument based on parallelism has an inherent weakness similar to that found in statistical studies using correlations as evidence. Parallelisms between two structures do not mean a cause and effect relationship between the two. Thus, we need stronger support.

C.M. Lee (1973:106-109) provides a more convincing argument. According to Lee, the reflexive pronoun caki 'self' is coreferential with either the subject or the topic NP when they command caki. For example, in (15), caki is not coreferential with ku namca 'the man'.

(15) *(caki, -ka a-nun) yeca-ka ku namca,l-ul cwukyessta.
    self SM know woman the man OM killed
    'the woman that he knows killed the man.'

because ku namca is neither the subject nor the topic. However, if we relativize the main clause object ku namca, as in (16):
(16) ai-ka [(caki₁-ka a-nun) yeca-ka cwuk-in] ku nam-
boy SM the man
ca₁-lul poassta. 9
'The boy looked at the man whom the woman that he knows killed.'

then, caki which was not coreferential with ku namca in (15), becomes coreferential with ku namca, which is the head NP of the relative clause, as in (17):

(17)

If ku namca came from the deleted object (∅), then caki and ku namca cannot be coreferential, as we see in (15). But caki and ku namca are coreferential in (17). Since reflexivization is triggered in Lee's analysis only when caki is coreferential either with the subject or the topic, Lee argues that the head noun ku namca did not come directly from object position, but from the topic position of S₂. He posits the following derivation:

(18)a [deep structure]:ai-ka [(caki₁-lul a-nun) yeca-
ka ku namcaₐ-lul cwukyessta] ku namca-lul
The basic motivation for claiming that topics and relative clauses are related is a semantic one. In a topic-comment structure, the topic is what the comment is said "about," and in a noun phrase consisting of a relative clause and its head, the head is what the relative clause is said "about." Thus, both structures share the same semantic property of predicating something of an entity. The comment and the relative clause are predicated "about" the topic and the relative head, respectively.

What is significant in this context is the fact that the topic (NP-nun) cannot appear within a relative clause. Consider (19):

(19) John-un hakkyo-ey kassta.
    TOP school to went
    'John went to school.'

where John is the topic. If we embed (19) into the following sentence:

(20)a  hakkyo-ka phakoytøyessta.
    school SM destroyed
    'The school was destroyed.'
as a relative clause, then we will have either (20b) or (21b):

(20)b *(John-un ka-n) hakkyo-ka phakoytoyessta.
'The school to which John(TOP) went was destroyed.'

(21)b (hakkyo-ey ka-n) John-i epsecyessta.
'John who went to school disappeared.'

Note that if the topic is relativized, as in (21b), the resulting sentence is acceptable. However, if some element other than the topic is relativized, as in (20b), the result is unacceptable.

This phenomenon is not unique to Korean. Gundel (1974: 80) notes that in English "left dislocation [which is a topic-comment structure] in a relative clause results in a completely unintelligible construction." Keenan (1972) also points out that the topic cannot appear in a relative clause in many other languages. Thus, it seems that non-occurrence of the topic in a relative clause is a cross-linguistic phenomenon.

The difference in acceptability between (20b) and (21b) is due to the following fact. In (20b), the clause within the relative clause is "about" John, whereas the relative clause itself is "about" hakkyo, the head NP. A relative clause and its head constitute a single semantic unit, because they are two component constituents of a single NP.
Thus, in (20b), the relative clause plus the head, which forms a single semantic unit, contains two conflicting "about-nesses" or "points of view": one "about" John and the other "about" the school. This conflict in topichood in a single semantic unit causes the unintelligibility of (20b), a kind of "topic-comment split" (see 2.2.4) which may be represented as follows:

\[
(22) \quad \begin{array}{c}
\text{(John-un TOP ka-n) hakkyo } \text{ } \text{NP} \\
\text{T_1} & \text{TOP} & \text{C_1} \\
\text{C_2} & \rightarrow & \text{T_2}
\end{array}
\]

where part of a single comment (e.g., ka-n) is predicated simultaneously about two distinct topics. Hence, the unacceptability of (20b).

In contrast, both the clause within the relative clause and the relative clause itself in (21b) are "about" John. Thus, there is no conflict between the two in terms of "about-ness" in the single semantic unit of the relative clause and the head. (21b) may be represented as follows:

\[
(23) \quad \begin{array}{c}
\text{(hawksyo-ey ka-n) John } \text{ } \text{NP} \\
\text{C} & \rightarrow & \text{T}
\end{array}
\]

Hence, the acceptability of (21b).

We claim that if both the clause underlying the relative clause and the relative clause itself are not predicated about the same entity (which is the relativized head NP), they always lead to unintelligible and unacceptable sentences because of a split in topichood. This fact not only
explains why the topic cannot appear in a relative clause across languages but also indicates that the relative head NP must be the topic of the underlying embedded sentence, if the relative clause is to be well-formed semantically.

We will discuss one more example from the same perspective. Suppose that the speaker has the following two propositions (24-25):

(24) capassta (swunkyeng, ku totwuk)
    caught policeman the thief
(25) sang-ul patassta (swunkyeng)
    was rewarded policeman

(24) can have three possible topic-comment articulations:

(26)a swunkyeng-un ku totwuk-ul capassta.
    policeman TOP
    'The policeman caught the thief.'
(27)a ku totwuk-un swunkyeng-i capassta.
    the thief TOP
    'As for the thief, the policeman caught him.'
(28)a ku totwuk-un swunkyeng-eykey cap-hi-essta.
    TOP by Passive
    'As for the thief, he was caught by the policeman.'

(26a) is predicated 'about' the policeman, and (27a-28a) are 'about' the thief. If (26a) is embedded into (25) as a relative clause, we will get (26b):

(26)b (ku totwuk-ul cap-un) swunkyeng-i sang-ul patassta.
    'The policeman who caught the thief was rewarded.'

But, if (27a-28a) are relativized, the following two
unacceptable sentences are obtained:

(27)b *(ku totwuk-un \(\text{TOP}\) cap-un) swunkyeng-i sang-ul patassta.

(28)b *(ku totwuk-un/i cap-hi-n) swunkyeng-i sang-ul \(\text{TOP}\) patassta.

(26b) is acceptable because there is no conflict in 'aboutness' within the semantic unit of the relative clause plus the head; both the clause underlying the relative clause and the relative clause itself are 'about' the policeman. (27b-28b) are unacceptable because the clauses underlying the relative clauses are 'about' the thief but the relative clauses themselves are 'about' the policeman.

The unacceptability of (27b) can be improved by raising the embedded topic into the main clause.

(29) "ku totwuk-un \(\text{TOP}\) (ku-lul cap-un) swunkyeng-i thief him OM caught sang-ul patassta.

'As for the thief, the policeman who caught him was rewarded.'

The result is much less natural than (26b). In this case, however, the topic within the relative clause is not the thief, but the policeman. Thus, there is no conflict in topichood in the underlined expression, so (29) is acceptable.

However, there is no way to save the unacceptable (28b). For example, even though the embedded topic is raised
to the main clause, as in (30):

(30) *ku totwuk-un (ku-ka cap-hi-n) swunkyeng-i
    thief TOP he SM was-caught
    sang-ul patassta.
    'As for the thief, the policeman by whom he was
    caught was rewarded.'

The result is still unacceptable. This is because the topic
within the relative clause in (30) is the policeman. which
functions as the passive by-agent in the embedded sentence.
As already mentioned in Chap 3: 3.2.3.3, however, the pas­sive by-agent cannot be the topic. Hence, the unacceptabil­ity of (30).

In short, the relative clause must be 'about' its head.
If some element other than the head NP appears as the topic
within the relative clause, this creates a conflict in terms
of 'aboutness' within the semantic unit of the realltive
clause and the head NP, yielding an unacceptable result.
Therefore, the relative head NP must be the topic of the em­bedded sentence underlying the relative clause.

Thus, the conclusion in the preceding section that re­lativization is the simple deletion of a coreferential NP
should be revised to say that relative clauses are derived
by the simple Deletion of the topic of the underlying embed­ded sentence, along with subsequent Reordering.

2.2.1 Since a relative clause is a topic-comment sentence
in underlying structure, we will examine in this section
various types of Korean relative clauses in terms of topic-comment articulation, that is, the thematic structures of relative clauses. Since English is not an exception to the afore-mentioned generalization, we will also examine English relative clauses from the same perspective.

Let us start with simple relative clauses. Consider

(31) (Chelswu-ka mannan) Yenghi-nun nay tongsayng-ita.
    SM met TOP my sister be
    'Yenghi whom Chelswu met is my sister.'

(32) na-nun (Chelswu-ka mannan) Yenghi-lul salanghanta.
    I TOP SM met OM love
    'I love Yenghi whom Chelswu met.'

In (31-32), the relative clause Chelswu-ka mannan is the comment to the topic Yenghi, which in turn functions either as the topic or part of the comment of the main clause. Thus, (31-32) have the following topic-comment structures (33-34), respectively:

(33) \[ T_1 \rightarrow C_1 \]  \[ C_2 \leftarrow T_2 (=T_1) \]

(34) \[ T_1 \rightarrow C_1 \]  \[ C_2 \leftarrow T_2 (=C_1) \]

where the subscripts 1, 2, etc. refer respectively to \( S_1 \), \( S_2 \), etc. If the main clause is taken out of consideration, the thematic structure of a simple relative clause is represented schematically as (35):

(35) \[ T_1 \text{ or } C_1 \text{(main clause)} \]
    \[ C_2 \leftarrow T_2 \text{ (= } T_1 \text{ or } C_1 \text{)} \]
Simple relative clauses in English have much the same properties as (33-34), except that the direction of branching is reversed. Consider

(36) Mary, whom John met yesterday, is pretty.
(37) I love Mary, whom John met yesterday.

(36-37) have the following thematic structures:

\[(\text{T}_1 \rightarrow \text{C}_1 \rightarrow \text{C}_2)\]

Multiple relative clauses show somewhat different patterns in thematic structure. Before we start analyzing multiple relative clauses, a brief comment on the different types of multiple relative clauses is in order. The recursive characteristic of relativization makes it possible to generate an infinite number of sentences with a finite number of rules. Chomsky (1965:12-14) treats recursion as consisting of five types according to their formal properties: nesting, multiple branching, left-branching, right-branching, and self-embedding. Nesting and multiple branching are not directly relevant to relativization, so we will be concerned here with left-branching, right-branching and self-embedding as they relate to multiple relative clauses.

2.2.2 First we will take up left-branching (Korean) and right-branching (English) multiple relative clauses. A relative clause "A is left-recursive if there is non-null B such
that $A \xrightarrow{\alpha} AB$" (Chomsky 1961:10) with the following phrase marker:

(40)

In contrast, the relative clause "A is right-recursive if there is a non-null B such that $A \xrightarrow{\alpha} BA$" (Chomsky 1961:10) with the following phrase marker:

(41)

English is basically a right-branching language, while Korean is a left-branching language in which relativization can be extended indefinitely leftward. Consider

(42)a ikes-un [(kim cwukun) nmača-lul salanghanun]11 this TOP already died man OM loved

yeca-lul ttaylin} Chelswu-lul kulin] chayk ita.
woman OM hit OM dealt book be

'This is a book that deals with Chelswu who hit the woman who loved a man who is already dead.'

(43) \{(saylo chilhan) mwun-yeph-ey issnun\} chayk-

newly painted door side at placed desk

sang-wi-ey nohin} chayk-un John-uy kes ita.

Top at placed book TOP of thing be

'The book that was on the table that was near the door that was newly-painted belongs to John.'
Sentences like (42a-43) are common in Korean, and they are all acceptable. The thematic structure of a multiple left-branching relative clause in Korean looks like (44a):

\[(44)a \quad \text{T}_1 \text{ or C}_1 \text{ (main clause)} \]
\[\downarrow \]
\[\text{C}_2 \quad \text{T}_2 = \text{T}_1 \text{ or C}_1 \]
\[\downarrow \]
\[\text{C}_3 \quad \text{T}_3 = \text{C}_2 \]
\[\downarrow \]
\[\text{C}_4 \quad \text{T}_4 = \text{C}_3 \]
\[\downarrow \]
\[\text{C}_n \quad \text{T}_n = \text{C}_{n-1} \]

As an illustration, (42a) fits well into the scheme (44a):

\[(42)b \quad \text{ikes-un} \left\{ \langle \text{imi cwukun} \rangle \text{ namca-lul salanghanun} \right\} \quad \text{T}_5 = \text{C}_4 \]
\[\text{yeca-lul ttaylin} \left\{ \langle \text{Chelswu-lul kulin} \rangle \text{ chayk ita.} \right\} \quad \text{T}_3 = \text{C}_2 \quad \text{T}_2 = \text{C}_1 \]

Note that the left-branching multiple relative clauses exhibit a chain of topic-comment nexuses in an orderly manner. That is, relativization proceeds by successively thematizing the previous comments.

The English multiple relative clauses which correspond to Korean (42a-43) have a right-branching structure, where relativization can be extended rightward indefinitely.

\[(45)a \quad \text{This is the dog} \left\{ \langle \text{that chased the cat} \rangle \text{ that killed the rat} \langle \text{that ate the cheese} \langle \text{that was rotten} \rangle \rangle \right\} .\]
(46) The book {that was on the table (that was near the door (that was newly-painted))} belongs to John.

Sentences like (45a-46) are all acceptable in English and have the following thematic structure, one identical to Korean (44a) except for the direction of branching:

\[
\begin{align*}
(44)b & \quad T_1 \text{ or } C_1 \text{ (main clause)} \\
& \quad (T_1 \text{ or } C_1 =) T_2 \quad \downarrow \\
& \quad (T_1 \text{ or } C_1 =) T_2 \quad \rightarrow C_2 \\
& \quad (C_2 =) T_3 \quad \downarrow \\
& \quad (C_2 =) T_3 \quad \rightarrow C_3 \\
& \quad (C_3 =) T_4 \quad \downarrow \\
& \quad (C_3 =) T_4 \quad \rightarrow C_4 \\
& \quad (C_{n-1} =) T_n \quad \rightarrow C_n
\end{align*}
\]

For example, the topic-comment organization of (45a) can be shown as (45b):

\[
(45)b \quad \text{This is the dog } [\text{that chased the cat } \{\text{that killed the rat } \langle \text{that ate the cheese (that was rotten) } \rangle \}].
\]

Like their Korean counterparts, English right-branching multiple relative clauses structure a chain of topic-comment nexuses in an orderly manner by thematizing the previous comments successively and linearly. This thematic structure seems to be the most typical way of forming a relative
clause in many languages. Each topic is developed naturally from the foregoing comment.

What attracts our attention in this connection is the striking similarity between this feature of relative clauses and the thematic structure of text. Daneš (1974:118-120, 1970), investigating the organization of text, found that the topic-comment distinction plays a key role in text organization. He suggested, as organizing principles for text, three major types of what he calls "thematic progression." One is "simple linear thematic progression," which has the following structure:

\[
\begin{array}{c}
T_1 \rightarrow R_1 \text{(Rheme)} \\
\downarrow \\
T_2 (=R_1) \rightarrow R_2 \\
\downarrow \\
T_3 (=R_2) \rightarrow R_3
\end{array}
\]

(47) Simple linear thematic progression (Daneš 1973: 118)

(47) is the most elementary and basic thematic progression, found not only in written text (Daneš 1974) but also in spontaneous narrative texts (Scinto 1977). It is not difficult to find many texts with this thematic structure. We will quote an English example from Daneš (1974:118):

(48) The first of the antibiotics was discovered by Sir Alexander Flemming in 1928. He was busy at that time investigating a certain species of germ.
which is responsible for boils and other troubles.

The following example is taken from a Korean testbook: 12

(49) na-nun khun hak-ul thako wuli nala-uy Seoul-ev 
I TOP big crane rode our nation's to
T₁

ka-poasse. wuli nala-uy Seoul-un khuko alumtawa. 
went C₁

T₂ (=C₁) (C₂)

kil-to nelpe. kil-keli-nun bus-wa taxi-ka cwul-
street wide street TOP and SM line
C₂

T₃ (=C₂)

cie taniko issesse.
in go Progressive C₃

'I went to Seoul, the capital of our country, on a crane. The capital of our country is large and beautiful. The streets are wide. On the street there were many buses and taxis going.'

Note that the English and Korean texts exhibit the same chain of topic-comment structures in the same orderly progression. The same phenomena are observed in Czech and German (Danes' 1974:118), which indicates that what we have observed above has some cross-linguistic validity.

2.2.3 It is not always the case that a language is either exclusively left-branching or right-branching. Though Korean and English are dominantly left-branching and right-branching respectively, both languages allow recursion in the opposite direction in the so-called stacked relative clauses. That is, stacked relative clauses in Korean possess a right-branching structure, while English counterparts have a left-
branching structure. Consider the following Korean data:

(50) apeci-ka 〈kay-lul ttaylin 〈mok-i malun (phikon-father SM dog OM hit throat dry tired hay poinun)〉 〈John-ul kkwucicessta.

'Father scolded John, who looked tired who was thirsty who hit the dog.'

(51) 〈Mary-ka salanghanun (sikan-tang 50 mile-ul talli- SM love hour per OM run-nun)〉 kay-nun tokil-san ita.

dog TOP German-born be

'The dog which makes 50 miles an hour which Mary loves is German-born.'

(50-51) are stacked relative clauses, in which the inner-most relative clause modifies the head, and the next inner-most clause further modifies the same head already modified by the inner-most clause, and so on. The thematic structure of stacked relative clauses can be represented as (52a):

\[
\begin{array}{c}
C_2 \leftarrow C_3 \leftarrow C_4 \leftarrow \cdots \leftarrow C_n \leftarrow T_n(T_2) \\
\downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\
T_1 \text{ or } C_1 & T_2(=T_1 \text{ or } C_1) & T_3(=T_2) & T_4(=T_2) & \cdots \\
\end{array}
\]

Linearly arranged, (52a) has the following thematic structure:

\[
\begin{array}{c}
C_2 \leftarrow C_2 \left\{ C_3 \left( C_4, (C_5) \right) \right\} \leftarrow T_2(=T_3=T_4=T_5) \\
\end{array}
\]
in which a chain of different comments (i.e., relative clauses) is linked up to the same topic and the progression of topic-comment nexuses is linear and orderly.

English stacked relative clauses are similar to those in Korean (50-51).

(53) John found the dog «that Mary loved) that the cat chased».

(54) The man «that she loved) that cooked rice» had a stomachache.

The same thematic structure as that of Korean is observed in (53-54):

\[ \begin{align*}
T_1 \text{ or } C_1 & \quad T_2 \quad C_2 \\
\downarrow & \downarrow & \\
T_2 = T_3 & \quad C_3 \\
\downarrow & \quad \vdots \\
T_2 = T_n & \quad C_n
\end{align*} \]

A similar type of textual organization in terms of topic-comment structure was found by Dane's, who called it "thematic progression with a continuous (constant) theme."

\[ \begin{align*}
T_1 & \quad C_1 \\
\downarrow & \downarrow \\
T_2 & \quad C_2 \\
\downarrow & \downarrow \\
T_3 & \quad C_3 \\
\downarrow & \downarrow \\
T_n & \quad C_n
\end{align*} \]
A cursory look at any piece of the Korean texts will reveal that this thematic structure is very productive, too.

(56) ayki twayci yeltwu mali-ka sophwung-ul kassupnita.13
baby piggy twelve SM picnic OM went

Ø cemsim-ul ssa-kaciko kassupnita. Ø nolay-lul
lunch OM prepare and went C₁ T₂
song OM

pwulumiense cwul-ul cie kassupnita. Ø sinaymwl-
singing line OM make went C₂ T₃
stream

ul kernekey toyessupnita.
OM cross happen to C₃

‘Twelve baby piggies went on a picnic. (They) went with a lunch box. (They) went in a line singing a song. (They) happened to cross a stream.’

Daniel gives similar examples from English, German and Czech (1974:119), from which we will quote an English example:

(57) The Rousseauist especially feels an inner kin-
ship with Prometheus and other Titans. He is

fascinated by any form of insurgency... He must

show an elementary energy in his explosion

against the established order and at the same
time a boundless sympathy for the victims of it.
Further the Rousseauist is ever ready to discover

beauty of soul in any one who is under the re-
probation of society.
Note that one and the same topic appears in a series of utterances, and that different comments are linked to it, just as in the stacked relative clauses. 14

2.2.4 We will next discuss self-embedded multiple relative clauses. Chomsky defines self-embedding recursion as having the following configurational properties: a structure of category A is self-embedded in B if A is preceded and followed by non-null strings of B, and if A is a phrase of the same type as B. The phrase of the same type refers to identical syntactic categories such as NP, S, and VP. The phrase structure looks like (58):

(58)

Both Korean and English allow self-embedding recursion in relative clauses. In Korean, self-embedded multiple relative clauses always occur when more than two relative clauses are attached to the grammatical object. This is due to the fact that Korean has an SOV pattern.

(59)a Chelswu-nun Yenghi-ka Yengswu-ka (Swunhi-ka salanghan) na-eykey cwun mongtwungi-lo ttay-
SM loved I to gave bat with hit
'Chelswu hated the dog that Yenghi hit with the bat that Yengswu gave me whom Swunhi loves.'

'Father met the teacher to whom Mother handed over the book that the clerk gave.'

In (59a-60a), the parenthesized relative clause modifies an element of the angle-bracketed relative clause with some non-null elements on both sides, and the angle-bracketed relative clause modifies an element of the curly-bracketed relative clause and so on. So, these sentences contain multiply self-embedded relative clauses.

(59a-60a) are very different from the relative clauses we have so far considered. They are either unacceptable or ungrammatical to some native speakers. This is not an idiosyncratic feature of Korean; the same phenomenon is observed in English, too.

The bread that the rat that the cat (that Mary loved) killed ate was rotten.

The nurse that the cook (that the maid met) saw heard the doctor.

It is clear that a self-embedded multiple relative clause is fundamentally different, in some sense, from either left-branching or right-branching relative clauses. The latter two generally pose no difficulty for comprehension, no matter
how much they are expanded, but self-embedding recursion, even beyond two or three embeddings, drastically reduces the degree of comprehensibility (cf. Fillenbaum 1973:42-51).

It has been suggested that psycholinguistic factors such as "limited memory span" or "limited information processing capacity" may account for this phenomenon but, as Kuno (1974:120-121) has pointed out,

there are no hypotheses, (at the moment), that psycholinguists agree upon that would account for the fact that center-embedding [i.e., self-embedding] but not right-embedding or left-embedding, hinders speech comprehension.... I will simply assume, without explanation, that it is a universal fact of language. (Emphasis added.)

It will be argued below that a partial linguistic explanation, rather than a psychological one such as "memory span", may be given for this phenomenon by investigating the self-embedded relative clauses in terms of their thematic structures.

The thematic organization of the multiply self-embedded Korean relative clauses may be schematically represented as

(63a):

\[
\begin{align*}
\text{C}_2 & \leftarrow \ldots \leftarrow \text{C}_2 \leftarrow \text{T}_2 (= \text{T}_1 \text{ or } \text{C}_1) \\
\text{C}_3 & \leftarrow \ldots \leftarrow \text{C}_3 \leftarrow \text{T}_3 (= \text{C}_2) \\
\text{C}_4 & \leftarrow \ldots \leftarrow \text{C}_4 \leftarrow \text{T}_4 (= \text{C}_3) \\
\text{C}_n & \leftarrow \text{T}_n (= \text{C}_{n-1})
\end{align*}
\]
We can arrange the topic-comment nexuses in (63a) in linear order as (63b):

\[(63)b \quad [C_2 \langle C_3 \langle C_4 \langle C_5 \ T_5 \rangle \ C_4T_4 \rangle \ C_3T_3 \rangle \ C_2T_2] \ T_1 \text{ or } C_1\]

The thematic structure (63b) may look similar to (44) or (52), but there is one important difference. In (63b), the comment of each relative clause is split into two parts, which may be called a "topic-comment disjunction," except in the most deeply embedded clause (e.g., \(C_5 - T_5\) in (63b)).

It seems that the topic-comment nexus is an essential communicative unit, functioning as a building block in communication. This means that a sentence is intelligible only if the hearer can identify what the speaker is talking about and what is actually being said about it, and can perceive a plausible connection between the two, i.e. recognize the topic and the comment. Relative clauses with a thematic structure like (63) are unintelligible because the hearer cannot recognize plausible connections between the topics and the comments. Note that this kind of topic-comment split does not appear in relative clauses with the thematic structures (44) or (52). Hence, (44) and (52) are perfectly intelligible. Although other psychological factors, memory span for example, may be involved, we maintain that, linguistically, the splits between the topic and the comment (or one part of the comment) are the primary reason for the reduced intelligibility of multiply self-embedded relative
For illustration, we take (59a), which has the following thematic structure (59b):

\[(59)b\quad \text{Chelswu-nun} \begin{cases} Yenghi-ka \\ Yengwu-ka \end{cases} (Swunhi-ka salanghan) na-eykey cwun > mongtwungi-lo ttaylin
\begin{cases} \text{T}_1 & \text{C}_2(a) \\ \text{C}_3(a) & \text{C}_4 \\ \text{T}_4 = \text{C}_3(b) \\ \text{T}_3 = \text{C}_2(b) \end{cases}
\text{kay-lul miwehayssta.}
\text{T}_2 = \text{C}_1
\]

Note that \(C_2(a)\) is split from both \(C_2(b)\) and \(T_2\) by two relative clauses, and that \(C_3(a)\) is split from \(C_3(b)\) and \(T_3\) by one relative clause. Because of these interruptions and the consequent unclearness as to the connections between the topic and the comment, the hearer may not recognize the topic-comment nexuses. Therefore, (59b) becomes completely incomprehensible. A similar explanation should hold for English self-embedded relative clauses.

Not surprisingly, in the literature of text analysis there has been no report of a text with a hypothetical thematic progression such as (64):

\[(64)\quad \text{T}_1 \quad \text{or} \quad \text{C}_1
\begin{cases} \text{T}_1, \text{C}_1 = \text{T}_2 \\ (\text{C}_2 =) \text{T}_3 \\ (\text{C}_3 =) \text{T}_4 \end{cases} \text{C}_1 \longrightarrow \text{C}_2 \longrightarrow \text{C}_3 \longrightarrow \text{C}_4
\]

where one part of the comment is split both from the other
part of the same comment, and also from the topic. If we work out a hypothetical text such as (65), it would be easy to see why no such texts are found:

(65) *halapeci-nun yongmalwu-ev kyeyvesstta. vong-\[15
grandfather TOP main floor at was main-
T_1\[1]
c_1malwu-nun otongnamwu-lo tove-issessko, ku otong-
floor TOP paulownia with made-of and the pau-
T_2\[2]c_2
namwu-nun cheum cip-ul ciul ttay, Pwulthay-san-
lownia TOP first house built when Mt.Pwulthay
T_3

evse nalla-on kes ilako hanta. Pwulthay-san-un
from brought that is-said Mt.Pwulthay TOP
C_3
10-li-na ttelecye-issess-nuntey, motwu cikey-lo
mile far-away and all A-frame by
C_4
nalla wassta hamye, hangsang pantulpantul ywun-i
brought came is-said always glossy gloss
C_3
na-issesstta.
shining
C_2

'Grandfather was sitting on the main floor. The main floor was made of paulownia trees, and the paulownia trees were brought from Mt.Pwulthay when the house was first built. Mt.Pwulthay was 10-li away, and (the paulownia trees) were brought by A-frames, and (the main floor) had a smooth and shining gloss.'
This hypothetical text is quite incoherent because the comments $C_2$ and $C_3$ are split into two. Note that it is impossible to associate the last underlined $C_2$ with its topic $T_2$.

Since splits in the thematic structure of self-embedded relative clauses impair the orderly progression of communicative units, i.e. strict topic-comment nexuses, it is not surprising that languages may have syntactic devices to repair such splits. For example, some languages—e.g. Czech—have morphological markers such as agreement endings, that help the hearer trace which topic is related to which comment. Some languages rearrange the split topics and comments so that each topic and comment is correctly aligned without a disjunction. Korean resorts to the latter strategy.

In Section 1, we claimed that Korean relativization consists of two separate processes: Deletion and subsequent Reordering. We noted that Reordering (6) moves one of the two split constituents of the relative clause so that they are adjacent. We now argue that Reordering is a syntactic process which functions to repair slips in the thematic structure of multiple relative clauses, a process which may have the same function in some other languages.

Note that the topic-comment splits in the self-embedded relative clause (59b) can be repaired by the Reordering rule (6), a rule formulated originally for stacked relative clauses. However, Reordering (6) will derive, from the input (59b), the wrong output (59c):
Although (59c) is ambiguous between the two readings given above, neither of them corresponds to the meaning of (59a):

'Chelswu hated the dog which Swunhi loved whom Yengswu gave me whom Yenghi hit with a bat.'

(i) 'Chelswu hated the dog which Swunhi loved whom Yengswu gave me whom Yenghi hit with a bat.'

(ii) 'Chelswu hated the dog which Swunhi loved which Yengswu gave me which Yenghi hit with a bat.'

In a number of languages, either left or right branching—but not both—are allowed in the surface structure of acceptable multiple embedded relative clauses. Since there are but two choices, either left or right, and since right-branching leads to 'stacked' relative clauses (in Korean), self-embedded relative clauses which are 'non-stacked' must be changed to have the left-branching structure on the surface. We have seen in Section 1 that if a stacked relative clause is to maintain the correct depth of embedding of the underlying structure on the surface, the direction of Reordering must be leftward. That is, to maintain the surface right-branching structure, Reordering must proceed leftward in 'stacked' relative clauses. Then, in 'non-stacked'
self-embedded relative clauses such as (59a-b), Reordering must take place rightward, so that iterative application of Reordering to the right results in the surface left-branching structure. Thus, we will now reformulate our Reordering rule in terms of topic and comment:

(66) Relativization (Reordering)

\[
SD: X \left[ S_1 C_1 (S_2 C_2 + T_2) C_1 \right] T_1 + Y
\]

(66a) SC: 1 \emptyset 3 4+2 5 [self-embedded]

(66b) SC: 1 2+4 3 \emptyset 5 [stacked]

(66a) is for self-embedded non-stacked relative clauses, and (66b) for stacked relative clauses.

Rule (66b), which operates on the stacked relative clauses, will be tested in 2.2.5. Let us see how (66a) operates in non-stacked self-embedded relative clauses like (59b). If we apply (66a) to (59b), we get the following output (59d), the correct surface form:

(59)d Chelswu-nun \{<<(Swunhi-ka salanghan) na-eykey
\]
\[
T_1 = C_4 T_4 = C_3
\]
\[
\text{Yengswu-ka }cwun\rangle \text{ mongtwungi-lo Yenghi-ka}
\]
\[
T_3 = C_2
\]
\[
ttaylin\rangle \text{ kay-lul miwehayssta.}
\]
\[
T_2 = C_1
\]

'Chelcwu hated the dog that Yenghi hit with the bat that Yengswu gave me whom Swunhi loved.'

Note that the problematic topic-comment splits of (59b) are now repaired in (59d), where there is instead an orderly and
linear progression of the topic-comment nexuses, and that
(59d) exhibits left-branching structure on the surface. It
also should be noted that the thematic structure of (59d) is
exactly identical to that of (44), which has linear themati-
izations of the previous comments. This confirms that the
thematic structure (44) is a basic and unmarked topic-com-
ment structure for relativization.

To summarize this sub-section, first, the primary mo-
tivation of the Reordering rule in Korean relativization is
to rearrange the split topics and comments such that there
is a linear progression of the strict topic-comment adjunc-
tions. Second, Reordering is subject to a sort of output
constraint, i.e., Reordering proceeds leftward in the stacked
relative clauses of the surface right-branching structure,
and proceeds rightward in the non-stacked self-embedded relative clauses which must have left-branching structure on the
surface. 17

Whereas reordering is the major syntactic device which
Korean employs to achieve the correct topic-comment align-
ment, English seems to use a different syntactic device.
Passivization has a similar function in English (see Hinds
1975). For example, (61a) has the following thematic struc-
ture with several topic-comment splits, as in (61b):

\[
(61b) \text{The bread} \langle \text{that the rat} \langle \text{that the cat} (\text{that Mary loved}) \text{killed} \rangle \text{ate} \rangle \text{was rotten.}
\]

\[
T_1 T_2 C_2 T_3 C_3 T_4 C_4
\]

\[
C_3 C_2 C_1
\]
Passivization converts (61b) to (61c):

\[(61)c \text{ The bread } \{ \text{ that was eaten by the rat } \{ \text{ that was killed by the cat } \{ \text{ that Mary loved } \}\} \text{ was rotten.} \]

(61c) is fully acceptable and does not have the splits in its thematic structure that its counterpart (61b) does; there is an orderly progression of topic-comment nexuses in (61c). Note that, just as in Korean, the thematic organization of (61c) is identical to that of (44b).

2.2.5 The last type of relative clause we examine has both the formal properties of self-embedding and right (Korean) or left (English) branching recursions. Due to the latter property, it is a stacked relative clause. We might call them self-embedded right-branching (Korean) and left-branching (English) stacked relative clauses. Self-embedded stacked relative clauses can occur in Korean when more than two relative clauses are attached to the same subject.

(67)a ??? {apeci-ka <emeni-ka (ai-tul-i cohahanun) saon> father SM mother SM boys SM love-ing buy
pelin }kwaca-nun pissata.
threw cake TOP expensive
'The cakes that boys love that Mother bought that Father threw away are expensive.'

(68)a ??? Swunhi-ka {Chelswu-ka <Yengswu-ka (Yenghi-ka salanghanun) ttaylin> cwukin } koyangi-lul miwe-loved hit killed cat OM hated
Self-embedded stacked relative clauses have the following thematic structure:

\[(69)\]  
\[
\begin{array}{c}
C_2 & - & - & - & - & C_2 \\
\downarrow & & & \downarrow & \text{T}_2 (= \text{T}_1 \text{ or } C_1) \\
C_3 & - & - & - & - & C_3 \\
\downarrow & & & \downarrow & \text{T}_3 (= \text{T}_2) \\
C_4 & - & - & - & - & C_4 \\
\downarrow & \quad & \quad & \quad & \downarrow & \text{T}_4 (= \text{T}_2) \\
C_n & \quad & \quad & \quad & \quad & \text{T}_n (= \text{T}_2)
\end{array}
\]

\[(69a)\] may be represented linearly as \[(69b)\]:

\[(69b)\]  
\[
\begin{array}{c}
\{C_2 \{C_3 \{C_4 \{C_5\}\} \} \} \text{ C}_2 \} \text{ T}_2 (= \text{T}_3 = \text{T}_4 = \text{T}_5) \\
\end{array}
\]

which is a combination of the two thematic structures \((52)\) and \((63)\). That is, one and the same topic continues throughout a series of topic-comment nexuses, and different comments are linked to the topic. But each comment is split into two parts. For example, \((67a)\) has the following topic-comment structure \((67b)\):

\[(67b)\]  
\[
\begin{array}{c}
\{\text{apeci-ka} \{\text{emeni-ka} \{\text{aitul-i cohahanun} \} \text{ saon}\} \} \text{ C}_2 \} \text{ C}_3 \} \text{ C}_4 \} \text{ C}_3 \\
\text{ pelin} \} \text{ kwaca-nun pissa} \text{.a} \} \text{ C}_2 \} \text{ T}_1 (= \text{T}_2 = \text{T}_3 = \text{T}_4) \} \text{ C}_1
\end{array}
\]
As we have seen, splits in the thematic structure interfere with the orderly progression of strict topic-comment adjunc-
tions. This interference is the main source of the incom-
prehensibility of (67b). Reordering rule (66b) can save the sentence from incomprehensibility. Since stacked relative clauses show right-branching on the surface, Reordering must proceed leftward. Reordering rule (66b) applies to (67b) 18 to produce the acceptable alternative (67c):

\[(67)c \{\text{apeci-ka pelin } \text{emeni-ka saon (aitul-i coha-} \begin{align*}
C_2 & \text{hanun)} \} \text{kwaca-nun pissata}. \\
C_3 & \\
C_4 & \end{align*}
\]

\[T_1 (\overset{=}{T_2} \overset{=}{T_3} \overset{=}{T_4}) C_1\]

Note that the thematic structure exhibited by (67c) is identical to that of (52), another main type of well-formed topic-comment organization in both relativization and text. It is also significant that no text is found with a thematic progression corresponding to (69).

English self-embedded left-branching relative clauses are also stacked and show a similar behavior:

\[(70)a \ *\text{The dog } \text{that (that chased the cat) bit John} \}\text{ belongs to me.}\]

\[(71)a \ *\text{Mary loves the hunter } \text{that (that is popular) lives in Boston} \].\]

(70-71) have the following linear topic-comment organization:

\[(69)c \ T_1 \{T_2 \langle T_3 (T_4 \text{C}_4) \text{C}_3 \text{C}_2 \text{C}_1 \}
\]
For an illustration, (70a) will be analyzed in its thematic structure:

(70)b The dog (that (that chased the cat) bit John) 
         T₁  T₂  T₃  C₃  C₂
   belongs to me.
   C₁

(70b) has not only a split between the topic (T₂) and the comment (C₂), but also a juxtaposition of two relative markers, which makes (70b) ungrammatical. (70b) may look different from Korean (67b), because (70b) is ungrammatical while (67b) is just unacceptable. The ungrammaticality of (70b) arises from the juxtaposition of the relative markers that that that. If Korean had such relative markers as English, Korean might well exhibit structural patterns similar to those of English (70b). A relative marker is a kind of conjunction, and Kuno (1974:126-128) has pointed out that, in many languages, any juxtaposition of more than one and the same conjunction leads to ungrammaticality. We claim, therefore, that English (70b) and Korean (67b) are identical in their thematic structures.

Passivization cannot help straighten out the splits in the thematic structure of (70b). As in Korean, reordering seems to be the only syntactic device available in English to repair these problematic cases. For example, (70a-71a) may be repaired by reordering the relative clauses in a linear order:
(70)c The dog (that chased the cat) that bit John belongs to me.

(71)b Mary loves the hunter (that is popular) that lives in Boston.

Note that (70c) has become ambiguous after reordering, but there seems to be no way to avoid this ambiguity. The resulting thematic structure corresponds well with that of (52c), one of the basic and unmarked thematic organizations in relative clauses and text.

2.2.6 Three significant implications can be drawn from the foregoing discussion. First, there is a plausible explanation for the unacceptability of sentences in which the topic and the comment or the two parts of the same comment are disjoined. While elements such as NP or VP are the formal building blocks of a sentence, elements such as topic and comment are the functional building blocks of communication. A communicative unit, which is presumably processed as a block by the listener, may be assumed to consist of the topic, what the speaker is talking about, and the comment, what is said about it. Communication proceeds in a chain of such communicative units; if the topic is separated from the comment, then the listener must first reconstruct the topic-comment communicative unit in order to understand the sentence. In contrast, if there is no such split in the communicative unit, reconstruction is not necessary and communication proceeds in a smooth way. We are thus able to offer a functional linguistic explanation for the processing.
difficulty associated with self-embedded relative clauses, an explanation utilizing the notion of topic and comment.

Secondly, it has been found that relative clauses and text sequences are quite similar in thematic structure. Note that the only acceptable thematic structures of relative clauses are (44) and (52), which are identical to those Daneš (1970, 1974) has found in text. Other unacceptable thematic structures of relative clauses, such as (63) and (69), must be adjusted to become either (44) or (52) by language-particular syntactic devices. Moreover, unacceptable thematic structures for relative clauses do not have counterparts in text. We claim that (44) and (52) are the basic and unmarked thematic structures for both relative clauses and text sequences, and hypothesize that a relative clause is a miniature representation of a text. That is, the well-formedness both of a relative clause and a text is governed by the same principle—an orderly and linear progression of communicative units consisting of a topic-comment nexus. In this connection Danes's remarks are noteworthy:

Any utterance, even a long and complex one, should contain, in essence, one theme-rheme nexus only, as the intelligibility of any utterance depends to a great extent on hearer's recognition of what the speaker is talking about and what he says about it. The same requirement is valid, analogously, even for longer portions of text, paragraphs: the text should be worded in such a way that the reader might without difficulties follow the thematic progression (Daneš 1970:139).

Our hypothesis that the relative clause is a syntactically compressed version of text seems to gain support from the findings of Sankoff and Brown (1976) as to the
development of relative clauses in New Guinea Tok Pisin. They claim that relative clause marking in Tok Pisin originated historically from discourse and that "relativization is only a special instance of the application of general 'bracketing' devices used in the organization of information" (p.631) in discourse sequencing.19

Thirdly, it seems that certain syntactic processes may operate to restructure sentences according to the above-mentioned principle. Languages may develop certain transformational processes to repair split thematic structures and achieve the topic-comment alignment necessary for successful communication. We have found that Korean Reordering (6) and English Passivization are two such syntactic devices.

3. Particle Deletion in Topics and Relative Clauses

In 2.1 we briefly reviewed Kuno's arguments for the hypothesis that topic-comment structures and relative clauses are related, a relationship based on four syntactic parallels between the two structures. The first argument involved a parallelism in particle deletion.

In English, when the object of a preposition is relativized, the preposition is either pied-piped with the relative pronoun or else is stranded in its original position. But in Korean, the postposition (which we will call 'particle' to avoid terminological confusion) attached to the relativized NP is simply deleted along with the NP. A similar
phenomenon is observed in thematization,\textsuperscript{20} where the particle attached to the NP is normally deleted when the NP is thematized.

If relative clauses are underlyingly topic-comment structures, we can go one step further and hypothesize that only topical sentences allowing particle deletion can be relativized. This is because relativization involves obligatory deletion of the particle. By the same token, topical sentences not allowing particle deletion cannot be relativized. Following this kind of implicit reasoning, Kuno hypothesized that "what is relativized is the theme immediately followed by wa" (1973:259) [Emphasis added].

It is not clear whether Kuno is implying an "if and only if" relationship between non-particle thematization (a term borrowed from Muraki (1970)) and relativization, or simply suggesting a high correlation between the two. But what Kuno seems to have in mind is an 'if and only if' relationship (see Muraki 1970:227-228). Thus, Kuno's statement might be rephrased: "Relativization is possible if and only if non-particle thematization is possible." That is, the acceptability of a relative clause is predictable from the deletability of the particle in the corresponding thematized sentence. This claim seems logical and elegant. It has one further advantage in that it allows us to simplify the grammatical descriptions of the two distinct structures by deriving one from the other. We need not repeat the same constraints twice (Kuno 1973:260)
Unfortunately, this hypothesis presents some problems. Cases have been reported which contradict any material implicational relationship between the two structures. Muraki (1970:203-216) has observed that particles of "instrument", "reason", "location" and "goal" in Japanese cannot be deleted in thematization, but the corresponding relative clauses in which the same particles are deleted are all acceptable. In Korean, likewise, the particles of "instrument", "reason", "animate goal" and "source" cannot be deleted in thematization, but relativization involving the deletion of these particles is acceptable.

**Instrument**

(72) a  John-i i khal -lo namwu-lul peyessta. 
   SM this knife with tree OM cut
   'John cut the tree with this knife.'

   b  i khal-lo-n John-i namwu-lul peyessta. 
      TOP
   'With this knife, John cut the tree.'

   c  *i khal-un John-i namwu-lul peyessta. 
      'As for this knife, John cut the tree (with it).'

   d  (John-i namwu-lul pey-n) i khal 
      'this knife (with which) John cut the tree.'

**Reason**

(73) a  John-un ku iyu -lo casalhayssta. 
   TOP that reason with killed himself
   'John killed himself for that reason.'

   b  ku iyu-lo-n John-i casalhayssta. 
      TOP
   'For that reason, John killed himself.'

   c  *ku iyu-nun John-i casalhayssta. 
      'As for the reason, John killed himself (for it).'

   d  (John-i casalha-n) ku iyu 
      'the reason (for which) John killed himself'
The (c) sentences of (72-75) show that particle deletion is impossible for the topical sentences. Yet, relativization is acceptable, as in the (d) sentences. Kuno has no explanations for this discrepancy between thematization and relativization, as he has acknowledged (Kuno 1973:260).

The fact that thematization and relativization show
differences in particle deletion in such sentences as (72-75) leads us to question the validity of Kuno's claim that what determines the well-formedness of a relative clause is the deletability of the particle in its corresponding topical sentence. In this sub-section we will first show for Korean that Kuno's claim is not correct. The deletability of the particle in thematicization is not relevant to the acceptability of the corresponding relative clause. Then, we will argue that particle deletions in both structures are governed by the same condition, "recoverability." Later, we will present an explanation of why the two structures differ as to particle deletion.

Japanese examples corresponding to the following Korean examples (76-77) led Kuno to assume that the deletability of particles in thematicization is the crucial requirement for relativization. Consider

(76)a John-i Mary-hako kongpwuhayssta.
   SM with studied
   'John studied with Mary.'

    b Mary-hako-n John-i kongpwuhayssta.
       TOP
   'With Mary, John studied.'

    c *Mary-nun John-i kongpwuhayssta.
       'As for Mary, John studied (with her).' 

    d *(John-i kongpwuha-n) Mary
       'Mary (with whom) John studied'

The comitative particle hako 'with' cannot be deleted in thematicization, as in (76c), and its corresponding relativization (76d) is also unacceptable. In contrast to this,
if there is a reciprocal adverb such as **hamkkey** 'together'
in the same sentence, as in (77):

(77)a  John-i Mary-hako hamkkey kongpuhayssta.
       SM  with together studied
       'John studied with Mary together.'

   b  Mary-hako-n John-i hamkkey kongpuhayssta.
       'With Mary, John studied together.'

   c  Mary-nun John-i hamkkey kongpuhayssta.
       'As for Mary, John studied together (with her).'  

   d  (John-i hamkkey kongpuha-n) Mary
       'Mary (with whom) John studied together'

then deletion of the particle is allowed, and both thematization and relativization are acceptable, as in (77c-d). Since a correlation exists between the topics and the relativized NP's in particle deletability, Kuno posits particle deletability in the topic as a prerequisite for acceptable relativization.

But, Kuno's argument totally misses the point. The coincidence in particle deletability between topical sentences and relative clauses in the (c-d) sentences of (76-77) may be just accidental. What determines the acceptability of the relative clauses (76d-77d) may have nothing to do with the particle deletability in the corresponding (c) sentences. That this may be the case is seen in the (c-d) sentences of (72-75), where the particle deletability of thematization does not correspond to the acceptability of relativization.

In Korean, particles differ from each other in deletability: some particles (e.g., **ka** 'SM', **lul** 'OM', **ey** 'locative
or time') are deletable both in thematization and relativization; some others (e.g., lo 'reason or instrument', sykey 'goal') are deletable in relativization, but not in thematization; 21 still others (e.g., wa 'comitative', pots 'comparative') are not deletable in either structure. Kuno's error was that he based his argument only on the examples where the deletability of a particle is identical in both structures, as in (7-8) and (76). This error led him to argue that deletion of the particle in thematization could be used to predict the well-formedness of relativization. The basic flaw in his argument is not revealed in data like (8) and (76), where particle deletion is the same in both structures. But if we include data such as (72-75), the incorrectness of Kuno's argument becomes clear.

The basic constraint on deletion is that the deleted information must be recoverable. The acceptability of the deletion of a particle depends on the recoverability of the relational meaning of the particle. We argue that the deletability of a particle in relativization and the acceptability of the relative clause have nothing to do with the particle deletability in the corresponding thematization; they are dependent upon whether the deleted particle is recoverable or not. By the same token, the deletability of the particle in thematization is also governed by the same recoverability condition. For example, (76d) is not acceptable because the meaning of the deleted particle is not recoverable, while (77d) is acceptable because it is recoverable. What plays a
crucial role in recovering the meaning of the deleted particle in (76-77) is the presence or absence of the reciprocal adverb hamkkey 'together'. (76d) is not acceptable because the deletion of the particle makes the sentence multiply ambiguous:

(78)a John-i Mary-lul wihayse kongpwuhayssta.
for 'John studied for Mary.' [purpose]
b John-i Mary-\text{ttaymwuney} kongpwuhayssta.
because-of 'John studied because of Mary.' [reason]
c John-i Mary-yepheyse kongpwuhayssta.
beside 'John studied beside Mary.' [location]

There is no way to determine uniquely the meaning of 'togetherness' in (76d). But in (77d), this is possible because of the reciprocal adverb hamkkey. Hence, the acceptability of (77d).

A further indication that our observation is correct comes from sentences like the following.

(79)a John-i Mary-hako kyelhonhayssta.
SM with married 'John married Mary.'
b Mary-hako-n John-i kyelhonhayssta.
TOP 'To Mary, John was married.'
c Mary-nun John-i kyelhonhayssta.
TOP 'As for Mary, John married (her).'
d (John-i kyelhonha-n) Mary
'Mary (to whom) John was married to'
(79d) is acceptable despite the deletion of the comitative particle and the absence of any reciprocal adverb. This is so because the verb kyelhonhata 'marry' is unambiguous as a reciprocal verb. Hence, (79d) is acceptable. (76-79) thus demonstrate that what determines the semantic well-formedness of a relative clause with a deleted particle is whether or not the particle is recoverable.

The remaining question is why thematization and relativization differ in particle deletion, i.e. why does the recoverability of the same particle differ between the two structures, as in (72-75)? We argue that this difference between thematization and relativization is related to the structural characteristics of the two structures.

Korean is an SOV language with pre-nominal relative clauses. And the topic always occupies S-initial position even before the subject. As an illustration, consider (72c-d), repeated here for easy reference:

(72)c *i khal -un John-i namwu-lul peyessta. this knife TOP SM tree OM cut 'As for this knife, John cut the tree.' [particle deleted]

\[ \] d (John-i namwu-lul pey-n) i khal 'this knife (with which) John cut the tree' [particle deleted]

The topic in (72c) occupies the farthest possible position from the verb, whereas the relative head NP in (72d) appears immediately after the verb. These two structural relationships hold true for all thematized and relativized sentences.
The relational meaning of a particle attached either to the topic or the relativized NP is always associated with the meaning of the verb. Therefore, when the particle is deleted, its realtional meaning is most easily recovered when the NP it is attached to appears immediately before or after the verb, with as few intervening elements as possible. The nearer the NP is to the verb and the fewer the elements that intervene between the NP and the verb, the greater the recoverability of its relational meaning. For example, the "instrumental" relationship associating the topic of (72c) with the three-place predicate peyessta 'cut' is hard to recover because the topic is too far away from the predicate and there are two intervening elements, John and namwu. In such a situation, it is difficult to understand the relation without the presence of an appropriate particle. Hence, the unacceptability of (72c).

In contrast, the relational meaning of the relative head NP of (72d), which is also an argument of the predicate peyessta 'cut' with the same 'instrumental' relation, is not difficult to recover because it appears immediately after the verb; no element intervenes between them. Hence, the acceptability of (72d). Therefore, we maintain that thematization and relativization differ in the extent to which the deleted particles are recoverable. That is, particles are more easily recovered in relativization than in thematization.

Our claim makes the following prediction: if the relative head NP is far from the verb of the embedded sentence,
its relationship with the verb of the embedded sentence will normally not be clear, and so the acceptability of the relative clause will decrease. This prediction seems to be borne out by such sentences as (80-81):

(80)a *[(John-i Mary-lul cwukyessta)-nun somwun-i SM OM killed Quote rumor SM
iss-nun] i khal
exist-ing this knife
'this knife, which there is a rumor that John killed Mary (with it)'

b (John-i Mary-lul cwukin) i khal
killed
'this knife (with which) John killed Mary'

(81)a *[(John-i casalhayssta)-ko allyeci-n] iyu SM killed himself is-known-as reason
'the reason, for which it is known that John killed himself'

b (John-i casalha-n) iyu
'the reason (for which) John killed himself'

In the (a) sentences of (80-81), the head NP's are separated from the verbs of the embedded sentences; so it is difficult to associate the relative head NP with the meaning of the embedded verbs in the (a) sentences. But this is not the case in the (b) sentences. From the viewpoint of the strict topic-comment adjunction discussed in 2.2, the (a) sentences contain a topic-comment split due to the intervention of the higher clause (e.g. somwun-i issnun in (80a)), while the (b) sentences do not manifest such a split.

In 2.1 we mentioned that Korean relativization sometimes leaves a pronominal copy in relative clauses. We
believe that this phenomenon of leaving pronominal copies also arises from the recoverability condition. That is, when a relative clause is separated from its head by higher clauses so that the relationship between the verb of the embedded clause and the head NP is difficult to recover (or predict), a pronominal copy is often left. 22 For example, the (a) sentences of (80-81) become much more acceptable to some speakers if there is a pronominal copy in the relative clauses:

(80)c [(John-i kukes-ulo Mary-lul cwukyessta)-nun that with somwun-i iss-nun] i khal
'this knife, which there is a rumor that John killed Mary with it'

(81)c [(John-i ku-ttaymwun-ev casalhayssta)-ko allye- that because-of ci-n] (ku) iyu
'the reason, which it is known that John killed himself for it'

Note that, in (80c-81c), the deleted particles are restored with the pronominal copies.

To summarize, we have shown, regarding Kuno's argument on particle deletion in thematization and relativization, that what determines the well-formedness of a relative clause is not the deletability of the particle in its corresponding topical sentence, but the recoverability of the meaning of the particle in the relative clause. We conclude that the difference between the two structures with regard
to particle deletion arises directly from the structural differences between the two. These differences are crucial to the recoverability of the semantics of the deleted particle.
1 Note Yang's statement: 'in such cases [i.e., in most relative clauses such as (2c)] pro-deletion is assumed to become idiosyncratically obligatory' (1975:153) [Emphasis added]. Though deletion of pronominal copies is difficult to specify, the use of pronominal copies in relative clauses can be specified and justified in terms of the communicative function of 'recoverability' (see Section 3).

2 I am aware that simple Equi-NP Deletion or Pronominalization based on coreference marking is problematic, e.g., due to sentences like the following (see Partee 1972, 1978):

(i) Every candidate wants ø to win.
(ii) Every philosopher contradicts himself.
(iii) nwukwutunci ø iki-ki-lul wenhnata.
   everyone win to OM want 'Everybody wants to win.'

We have pointed out problems of marking coreference in underlying structure in Chap 1: 3.1.4.1. To avoid problems like these, the use of variables is much preferable.

3 Sentence (5a) is actually three ways ambiguous. One reading is that of a left-branching relative clause: 'The dog that was hit by the truck that Chelswu likes died.' Here, what Chelswu likes is 'his truck.'

Another reading is that of a stacked relative clause: 'The dog that the truck hit (and) that Chelswu loves died.' This reading is clearest if there is only one dog in this particular universe of discourse.

A third reading is like the first reading, but the outer relative clause modifies the head noun plus the inner relative clause: '[The dog that the truck hit, that Chelswu loves died.' This reading is clearer when there are three dogs, two of them are hit by the truck, and Chelswu loves one of the two. This reading has the following underlying structure:
Note that the circled NP includes the squared $S_2$. (i) can be converted to (5a) by Deletion rule (3) without complication.

4 Note that Term 2 is optional, i.e., can be null. This optionality is required to derive sentences like (i) in ft. 18.

5 These languages allow relativizing of the subject only, which is actually the topic in these languages.

6 This sentence is all right in a "goal" reading, but not in its intended "source" reading.

7 (12b) is acceptable when John-un is understood as a contrastive focus, but unacceptable as an 'unmarked topic'.

8 See Kuno (1973:260) and Muraki (1970:227-228).

9 There are some native speakers for whom (16) is unacceptable.

10 $T_1$ and $C_1$ stand for the topic and the comment of the main clause; $T_2$ and $C_2$, those of $S_2$; $T_3$ and $C_3$, those of $S_3$, and so on.

11 To assist easy identification, the following notations will be used to indicate the depth of embedding: a parenthesis ( ) means 'most deeply embedded'; an angled bracket $< >$, 'second most deeply embedded'; a curled bracket { }, 'third most deeply embedded'; a bracket [ ], 'fourth most deeply embedded'. If there are more embeddings, the same notations will be repeated in the same order.


13 Kwuke [Korean Language] Vol. 3 (1972), 39

14 Allerton (1978:137) argues that a sentence usually has only one comment: "It eases the listener's job of decoding if each sentence contains only one item of new information. It is therefore natural for a sentence to contain one new element--typically the rheme--and a number of given elements--typically the theme." Contrary to his claim, a sentence usually has only one topic but it often has several comments, as (5C-54) show.

15 The text (65) is taken and adapted from Kwanchonswuphil [Essays on Kwanchon] written by Moon-Koo Lee.

16 I owe this information to Frank Lichtenberk (personal communication), who is a native speaker of Czech.
A similar proposal has been made in phonology by I. Howard (1972), who argues that the direction of application of a phonological rule is determined by the surface form.

In Reordering rule (6), Term 2 can be null. Consider the following stacked relative clause:

(i) \([\text{kicha-ey than (Seoul-eyse o-nun)]salam train on rode from coming man 'the man who is from Seoul who is on the train'}\]

(i) has the following underlying structure:

```
NP1
  S
    NP2
    V
      NP1
      salam
      NP2
    kicha-ey thassta
    salam
  Seoul-eyse wassta
```

Deletion rule (3) applies to NP2 to produce an intermediate structure (ii):

(ii) \([S_2 (S_1 \text{Seoul-eyse onun}) \text{salam-i kicha-ey thassta}]\)

We will assume that there is a trace \(\emptyset\) before \(S_2\), as in (iii):

(iii) \([S_1 \emptyset (S_2 \text{Seoul-eyse onun}) \text{salam-i kicha-ey thassta}]\)

Deletion rule (3) will apply again to (iii) to produce (iv):

(iv) \([S_1 \emptyset (S_2 \text{Seoul-eyse onun}) \text{kicha-ey than}]\)

Since (i) is a stacked relative clause, Reordering will apply to put together the trace \(\emptyset\) and \(\text{kicha-ey than}\), as in (v):

(v) \([S_1 \text{kicha-ey than (S_2 \text{Seoul-eyse onun})}]\)

As H.M.Sohn (personal communication) points out, the so-called 'non-restrictive relative clauses' are a type of discourse conjunction in Korean (I.S.Yang 1972:225-226). Note that text, conjunction, non-restrictive relative clauses,
and restrictive relative clauses actually form a continuum in representing discourse information in a sentential form.

20. In this sub-section, Kuno's terms such as "theme," "thematized," and "thematization" will be used freely. They correspond roughly to our term "topic."

21. The deletability of the source particle is not clear. Sometimes it is deletable, as in (75d), and sometimes not, as in (7c).

22. This is one of the two motivations for using a pronominal copy in a relative clause. The other is to emphasize the role of the relativized NP as focal information within the relative clause, as in (i):

(i) (caki-ka swukcey-lul ha-n) Chelswu
    self SM homework OM did
    'Chelswu who did homework for himself'
BIBLIOGRAPHY


Church, A. 1941. The Calculi of Lambda Conversion. Princeton University Press.


Lakoff, R. 1971. "If's, And's and But's about Conjunction." In C. Fillmore and D.T. Langendoen eds. 1971. 115-149.


261-277.


