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.
Ekniyom, Peansiri

A STUDY OF INFORMATIONAL STRUCTURING IN THAI SENTENCES

University of Hawaii

University Microfilms
International 300 N. Zeeb Road, Ann Arbor, MI 48106

Ph.D. 1982
A STUDY OF INFORMATIONAL STRUCTURING
IN THAI SENTENCES

A DISSERTATION SUBMITTED TO THE GRADUATE DIVISION OF THE
UNIVERSITY OF HAWAI'I IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN LINGUISTICS
MAY 1982

By
Peansiri Ekniyom

Dissertation Committee:
Roderick A. Jacobs, Chairman
Byron W. Bender
Thomas W. Gething
Robert W. Hsu
Laurence C. Thompson
ACKNOWLEDGEMENTS

My deep gratitude goes to all the members of the committees of my comprehensive examination and my dissertation: Byron W. Bender, Robert W. Hsu, Laurence C. Thompson, Fang Kuei Li and the late Ruth Crymes. I thank Thomas W. Gething for sharing with me his insightful views on Thai. I also want to thank Jeannette K. Gundel for the stimulating discussions and ideas. My special thanks go to my chairman, Roderick A. Jacobs, for his rigorous criticisms, which always come with constant encouragement. I also wish to thank all my professors and friends in the Department of Linguistics, especially Eiko Uehara and Leok Har Chan, for making my four and a half years of linguistic apprenticeship a memorable one. Wendy Onishi deserves special thanks for her kindness to me.

I acknowledge the generosity of the Culture Learning Institute of the East-West Center for providing me with a study grant. I am grateful to all the members of the staff of the Institute for their assistance during my stay at the Center. I specially want to thank Lyn Anzai and June Sakaba. Larry Smith has been both a helpful professional advisor and a caring friend. To him I offer my special thanks.
I thank the Faculty of Arts, Chulalongkorn University, for granting me the study leave. I am indebted to my friends and colleagues in the Linguistics Department for taking on extra teaching loads during my absence.

I wish to express my deepest gratitude to Professor Philip N. Jenner for his kind help and endless encouragement.

To Annette Fukuda, Julia Kwan, Seng Seok Hoon, Steve and Bonnie Dong, Elaine Tsui and Hideo Fujita I am grateful for their friendship.

I would like to give special thanks to Pairoj Vongvapanond for his patience and love.
ABSTRACT

This is a study of the interrelationships between syntactic forms in Thai and their informational properties. We are concerned primarily with the forms having noun phrase status. First, we examine simple noun phrases functioning as topics and/or subjects. We find that when they occur in sentence-initial position they are required to represent given information; i.e., they must be either definite or generic. This is hardly surprising for topic noun phrases. However, not all languages have such a requirement for subjects in initial position. We propose a constraint for Thai that sentence-initial noun phrases be [+given], to be referred to as SINPC or Sentence-Initial Noun Phrase Constraint. An examination of sentences with [-given] subject noun phrases shows that these can only occur in existential constructions, which have verb-initial surface forms. If we assume that underlying subjects are all generated in the same position, we must posit rules shifting [-given] noun phrases from sentence-initial position. A slight revision of SINPC is necessitated by our finding that any and all noun phrases occurring in pre-predicate position must be [+given].

We next turn to embedded constructions which appear to function as noun phrases. Two types of clauses can occur in sentence-initial position: the thî clauses
and the subjectless hypothetical clauses. The former do not assert their propositional content but rather take it as given. These appear to correspond to simple definite noun phrases. The subjectless hypothetical clauses, having no specific time reference of their own, correspond to simple generic noun phrases. Clause positioning, then, also conforms to the proposed SINPC. These clauses clearly function as topics and subjects.

The embedded clauses in the subject position of presentative and existential sentences are shown to represent [-given] information. These cannot remain in sentence-initial position. They are found instead in the position after the verb, resulting in verb-initial surface forms for these types of sentences.

Finally, we examine one problematic construction: a focused construction in which the initial noun phrase constitutes the comment rather than the topic. We find, however, that it is still [+given] information, and therefore does not violate SINPC. What prevents these [+given] noun phrases from being interpreted as topics is an obligatory focus marker which shows that the noun phrase with which it occurs is informationally assertive.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>iii</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>v</td>
</tr>
<tr>
<td><strong>CHAPTER I. INTRODUCTION</strong></td>
<td></td>
</tr>
<tr>
<td>1.1. Defining Informational Structuring</td>
<td>1</td>
</tr>
<tr>
<td>1.1.1. Informational function</td>
<td>3</td>
</tr>
<tr>
<td>1.1.2. Informational status</td>
<td>4</td>
</tr>
<tr>
<td>1.2. Statement of Purpose</td>
<td>6</td>
</tr>
<tr>
<td>1.3. Theoretical Framework</td>
<td>9</td>
</tr>
<tr>
<td>1.4. Data</td>
<td>14</td>
</tr>
<tr>
<td>Notes to Chapter 1</td>
<td>16</td>
</tr>
<tr>
<td><strong>CHAPTER II. SENTENCE-INITIAL NOUN PHRASE CONSTRAINT</strong></td>
<td></td>
</tr>
<tr>
<td>2.0. Introduction</td>
<td>19</td>
</tr>
<tr>
<td>2.1. Types of Sentences Studies</td>
<td>20</td>
</tr>
<tr>
<td>2.2. Subject and Topic</td>
<td>22</td>
</tr>
<tr>
<td>2.2.1. Definiteness</td>
<td>22</td>
</tr>
<tr>
<td>2.2.2. Semantic role</td>
<td>22</td>
</tr>
<tr>
<td>2.2.3. Relations with verbs</td>
<td>22</td>
</tr>
<tr>
<td>2.2.4. Agreement with verbs</td>
<td>23</td>
</tr>
<tr>
<td>2.2.5. Position in a sentence</td>
<td>23</td>
</tr>
<tr>
<td>2.3. Noun Phrases in Sentence-Initial Position</td>
<td></td>
</tr>
<tr>
<td>2.4. Subject As Topic</td>
<td>36</td>
</tr>
<tr>
<td>2.5. Basicness of Informational Function</td>
<td>47</td>
</tr>
<tr>
<td>2.6. Postposed Topic</td>
<td>55</td>
</tr>
</tbody>
</table>
2.7. Some Conclusions .......................... 57
Notes to Chapter 2 .............................. 58

CHAPTER III. EMBEDDED CLAUSES: SYNTACTIC FORMS AND
INFORMATIONAL STATUS

3.0. Introduction ............................... 59
3.1. Terminology and Assumptions ............. 60
3.2. Previous Analyses of Embedded Clauses
   in Thai .................................. 63
3.3. Proposed Analysis of Embedded Clauses
   in Thai .................................. 67
   3.3.1. Distinctions in sentence mood ....... 68
   3.3.2. The assertive-nonassertive
distinction .................................. 73
   3.3.3. The generic-nongeneric
distinction .................................. 81
3.4. Some Conclusions ......................... 92
Notes to Chapter 3 ............................. 95

CHAPTER IV. PRESENTATIVE AND EXISTENTIAL SENTENCES

4.0. Introduction ............................... 90
4.1. Presentative Sentences .................... 99
   4.1.1. Presentative verbs .................. 113
   4.1.2. Presentative subject and SINPC ... 115
   4.1.3. Presentative sentences with
distinct topic ............................. 122
4.2. Existential Sentences ..................... 125
   4.2.1. Presentative and existential
sentences .................................. 125
   4.2.2. Existential verbs .................... 128
4.2.3. Underlying topiclessness .......................... 137
4.2.4. Functional use of existential sentences ............... 141
4.3. Some Conclusions .................................... 144
Notes to Chapter 4 ....................................... 145

CHAPTER V. FOCUSED CONSTRUCTIONS

5.0. Introduction ......................................... 150
5.1. Surface Forms of Focused Constructions ........ .......... 151
5.2. Focused Noun Phrase ................................ 153
5.3. Focus Markers ........................................ 157
5.4. Some Conclusions .................................... 160
Notes to Chapter 5 ....................................... 161

VI. CONCLUSION ........................................... 163

BIBLIOGRAPHY ............................................. 166
CHAPTER I
INTRODUCTION

1.1. Defining Informational Structuring

*Informational structuring* is the term we propose to use in this study to refer to the correlation between syntactic forms on the one hand and the informational function and informational status on the other hand. A study of informational structuring, which is the goal of this thesis, involves an investigation of the extent to which and the means by which the syntax of a language reflects or corresponds to the informational properties of its structure.

Languages vary in respect to informational structuring. Some languages are likely to have direct correspondences between syntactic forms and their informational properties. In other languages, such a correspondence may not be easily discernible. The typological classification of languages along a parameter stretching from topic prominency to subject prominency, as proposed in Li and Thompson (1976), is one attempt to capture this difference among languages. A topic-prominent language, according to Li and Thompson, exhibits a tendency to reflect in the structure of its sentences the distinction between those elements functioning as topic and those functioning as comment. A subject-prominent language, on the other hand, tends not to
highlight this informational distinction but distinguishes the major constituents of a sentence in terms of their syntactic functions as subject, predicate, direct object, and so forth.

Other terms have been used in the literature to refer to what we term informational structuring. These are terms such as thematic organization and informational organization (Halliday 1976), functional sentence perspective (Mathesius 1975 and Kuno 1972), functional syntax (Kuno 1980) and topic-comment structure (Gundel 1974). These terms not only reflect differences in terminology but also imply conceptual differences. We propose to use the term informational structuring on the basis of our assumption that the conveying of some information is the raison d'être of a sentence. We claim no special advantage for our term. The terms we use simply reflect how we conceptualize this aspect of the language and how, through this conceptualization, we have come to understand certain syntactic phenomena in the Thai language.

Most linguists recognize the distinction we make between what we refer to as informational function and informational status. Halliday (1967), for example, differentiates the informational organization from the thematic organization. The former involves the distinction between the given and the new information, while the latter involves the functions theme and rhyme of elements in a
sentence. Gundel (1977 and 1978) uses the cover term the given-new distinction and differentiates two different types of this given-new distinction. One is the topic-comment distinction and the other is the activated-unactivated distinction.

1.1.1. Informational Function

The distinction between the elements functioning as the topic and the comment of a sentence will be considered under the rubric informational function. We have adopted the definition of topic and comment as proposed in Gundel (1974), in which the following logical structure is proposed as the underlying structure for all sentences:

(1)

The definition for the two terms are given by Gundel (1975:108) as follows:

Topic may thus be formally defined as the relation NP₁:S and comment as the relation S':S, where topic and comment are assigned semantic values roughly as follows:
Definition. If T is the topic of S, then S asserts, asks, promises, etc., something about T depending on the type of speech act act that S is used to perform.

Definition. If C is the comment of S, then C is what S asserts, asks, promises, etc., about the topic of S, depending on the type of speech act that S is used to perform.

Functionally, Gundel defines topic as "a point of departure" for the conveying of a message or information between a speaker and his addressee. For the information to be felicitously communicated, the topic of a sentence is required to have certain informational status—which will be discussed in detail in 1.1.2. The informational status required of a topic is that it represent information which is given and generic or definite. As for the comment, there is no such requirement. In Thai, however, as we will show in Chapter 5, the so-called focused construction is an exception to this requirement.

1.1.2. Informational Status

Informational status concerns those informational properties that are distinct from the topic/comment distinction. Thus, information may be categorized according to the following dimensions:
1. given or new
2. definite or indefinite
3. generic or nongeneric
4. contrastive or noncontrastive

We will represent these properties as the binary features $[\pm \text{given}]$, $[\pm \text{definite}]$, $[\pm \text{generic}]$, and $[\pm \text{contrastive}]$. Given information is information assumed by the speaker to be part of the general knowledge of both the speaker and the addressee. Givenness is the crucial property of the topic of a sentence. To have the status of being given, a topic phrase has to have a referent which is assumed to be known to both the speaker and the addressee. This means that it has to refer to some individual, object, or entity assumed to be in existence in a real or imaginary world; i.e., it has to be existentially presupposed. It also has to have a referent which is generic or definite--that is, a phrase which has a nondefinite but specific referent cannot be a topic of a sentence. Newness or $[-\text{given}]$ is the property of comment. This does not mean that the information within the comment has to be unknown to the addressee. Rather, it is what the speaker wants to convey to the addressee. It can be completely unknown information or information which is already part of the addressee's knowledge. This definition of given and new is adopted from Gundel (1977 and 1978).
Definiteness is defined in Chafe (1976) as the status which speakers assign to what they think their addressees already know and can identify. The topic of a sentence has to have a referent which is either generic or definite.

Contrastiveness can be the informational status of both topic and comment. A noun phrase is contrastive if there is more than one candidate that could qualify as its referent and only one candidate is selected. Chafe (1976) provides a similar definition though the example given suggests that what he has in mind is only a contrastive comment, which in English is marked by emphatic stress as in Rónald made the hamburgers, or a so-called cleft sentence as in It was Rónald who made the hamburgers. In Thai, contrastiveness is shown with special lexical markers for both topics and comments.

1.2. Statement of Purpose

In this study, we want to examine the extent of informational structuring in the syntax of Thai. The following phenomena are those which we propose to investigate:

1. noun phrases in sentence-initial position
2. forms and functions of embedded clauses
3. sentences with verb-initial surface forms
4. focused sentences.
The study will be presented in the following manner. In Chapter 2, we will argue that the Thai language favors the positioning of the noun phrase representing given information or topic in sentence-initial position. Data will be presented to show that when there is no distinct topic, the subject noun phrase, occupying the initial position, will also be interpreted as topic. Furthermore, the sentences we study also reveal that the noun phrases which occupy the positions before the main predicate phrase of sentences are all required to have informational status [+given]. We then propose to account for this phenomenon in terms of a constraint, referred to as Sentence-Initial Noun Phrase Constraint, or SINPC, which can function like a surface filter to eliminate sentences with sentence-initial noun phrases which violate this constraint. In addition, we will also show in this chapter that Thai has a set of lexical markers which can be employed not only to distinguish the topic from the comment but also to differentiate different types of both topic and comment. In Chapter 3, we examine the effect of the SINPC on the occurrence of embedded clauses in the initial position of the sentence. Of the types of embedded clauses, only the thii and the subjectless generic clauses are found to occur in the topic position. These forms, we argue, are the only ones permitted because they have the status [+given], and since embedded clauses are also given noun phrase status, they
are also governed by the proposed SINPC. In Chapter 4, we examine two types of sentences with verb-initial surface form: presentative and existential sentences. We argue that these sentences have in their sentence-initial position a noun phrase, embedded as well as unembedded, which is [-given] and thus cannot remain in those positions without violating the SINPC. We propose to account for the tendency in the language to preclude such sentences in terms of the Extraposition Transformation, which moves any noun phrase or embedded clause with [-given] status from pre-predicate position to post-verbal position. We will also discuss the differences between presentative and existential sentences and propose that they may be due to the fact that topic is absent in the underlying structure of the latter but present in that of the former. In Chapter 5, we examine focused constructions, which are informationally equivalent to cleft sentences in English. We will argue that a focused sentence in Thai has the surface form of an equational sentence with the identifying copula khít, one in which the topic phrase has been moved to sentence-final position. We will point out that even in this type of sentence the focused noun phrase, which is the comment of the sentence, is also governed by the SINPC. This focused noun phrase must be either definite or generic; it cannot be a noun phrase with an indefinite but specific referent. Since the informational status of
the focused noun phrase and its sentence-initial position make it eligible to be interpreted as topic, a focus marker is an obligatory element in the surface form to prevent such an interpretation.

1.3. Theoretical Framework

The general framework for this study is that of Generative Transformational theory. Our theoretical assumptions are not dissimilar to those of the Extended Standard Theory, especially as described in Chomsky (1976) and Chomsky and Lasnik (1977).

In our framework, the grammar consists of a base component, a transformational component, a semantic component and a phonological component. The base component contains a set of context-free phrase structure rules and the lexicon. The phrase structure rules generate an abstract phrase marker in which the grammatical relations are defined by the structural configuration as described in Chomsky (1965).

To this underlying structure are applied the transformational rules "restricted only to movement, left or right adjunction, and substitution of a designated element." Each movement is constrained to leave a trace, t, at the original site to preserve the grammatical information of the underlying structure so that it will be available for surface structure interpretation. The
output of the transformational component is the input to both the phonological and the semantic components.

The semantic component consists of various sets of semantic rules, which are applied to the surface structure. Examples of such semantic rules are the rules of construal, which coindex one element in the sentence as antecedent and another as anaphor, according to conditions on anaphora. These rules account for the interpretation of the moved elements as well as the anaphoric elements in the sentence. We assume as well that the interpretation of some anaphoric elements has to be derived pragmatically from nonlinguistic contexts or the real-world knowledge of the speaker/addressee. For this type of anaphoric element, the antecedent of which cannot be found in the linguistic context, we will use the symbol PRO, as an arbitrary symbol in the representation of sentences which we quote in our discussion. This is to differentiate it from the anaphoric form with an antecedent present in the linguistic context. For the latter, we use the symbol PRO; the symbol i marks the coindexing of such anaphor with its antecedent.

The output of the transformational component also becomes the input for the phonological component. Before the application of the phonological rules, the rules of deletion and filters may apply. Deletion rules eliminate, under strict conditions, certain elements in the abstract surface structure. Filters serve to rule out constructions
which are ill-formed. Then the string undergoes the phonological rules and becomes the phonetic representation of the sentence. Though we will not concern ourselves with the phonological aspect of the sentence in our discussion, we recognize that, for Thai sentences, informational structuring can have phonological manifestation in addition to syntactic manifestation. We will mention, however, only the occurrence of an internal juncture, or a pause, which can be employed to separate the topic of the sentence from its comment. We will use the symbol "//" to represent this pause in our sentence when it is needed for clarification.

There is another significant difference between our framework and the Extended Standard Theory as outlined in Chomsky (1976) and Chomsky and Lasnik (1977). This has to do with the informational structure of the sentence. In their work, topic and comment (presupposition and focus in their terminology) are interpreted from the surface structure. However, we shall assume, more or less as Gundel (1974) assumes, that topic and comment functions are configurations generated by the phrase structure rules of the base component. The initial phrase structure rule of the base component will be assumed to have a form roughly like (2) below:

(2) P.S. Rule 1: \[ S \rightarrow (NP) S' \]
in which the NP dominated by S has the relation "topic of S" while the S' dominated by S has the relation "comment of S." This rule can generate an abstract marker of a form very much like the one in (1), repeated here as (3) below, except that the variable x in the comment has been dropped in favor of the anaphoric form PRO, which is coindexed with the topic noun phrase.

(3)

\[
\begin{array}{c}
S \\
\text{NP}_i \\
\text{S'} \\
\ldots \text{PRO}_i \ldots \\
\end{array}
\]

In Gundel (1974), the variable we have shown as PRO\_i is generated as an empty variable in the comment. There is a rule which will copy the features of the topic onto this variable in the comment, which is bound by the topic. The rule of pronominalization will apply to pronominalize this set of copied features. Thus, the optional rule of deletion may be applied to delete this pronoun. We will assume, however, that this bound variable is generated in the base as PRO, which can be coindexed with the topic by the semantic interpretation rule. For the Thai language, this PRO can have the surface form of either a zero anaphor or a pronoun.\(^5\)
Though we have adopted a position similar to Gundel's concerning the specification of the informational functions *topic* and *comment* by the base rules, we are not convinced that this is the only possible theoretical assumption. As far as Thai is concerned, it seems to make little difference whether these informational functions are available in the base or interpreted in the semantic component, since the S-initial noun phrase in a Thai sentence always has the function of being topic, except in sentences where it has been postposed and separated from the comment by a pause.

We have also made a claim which is significantly different from Gundel (1974), in which all sentences are claimed to have a topic. We argue, on the basis of the Thai data discussed in Chapter 4, that there are sentences which present new information without making use of any given information (or topic) as a point of departure; i.e., these are sentences which are *topicless*. The topiclessness of these sentences, we will argue, is reflected in the obligatory absence of a sentence-initial noun phrase in these sentences, referred to as *existential sentences*. This is the reason for the parentheses in P.S. Rule 1, proposed in (2) and repeated here as (4):

(4) P.S. Rule 1: \[ S \rightarrow (NP) \ S' \]
1.4. Data

The sentences we study have been collected from both written and oral sources. However, in general, we have not concerned ourselves with phonological phenomena. Obviously, informational structuring can be manifested phonologically as well by means of sentence intonation and internal juncture. Such phenomena are complex and merit detailed study outside the scope set for this discussion.

Most of the sentences on which this study is based were collected from texts written by Thai authors, dialogs, and announcements on television and radio programs, and conversations among native speakers in which we were either an observer or a participant. The data represent the Bangkok or Standard Thai dialect. Occasionally, for ease of exposition, sentences used as examples were thought up by the author, although they almost always correspond to more elaborate sentences in the data collected.

When a native speaker's judgment on acceptability and grammaticality seems crucial to the analysis, the sentences were tested with members of our family and friends in Bangkok as well as with our Thai colleagues, who were participants at the East-West Center in Honolulu in 1980 and 1981.

The phonemic transcription used in representing the data is adopted from the phonological analysis of Thai
presented in Gething (1966). As for the morpheme-by-morpheme translation, the English equivalent provided for each word or morpheme is mainly for clarification of the semantic content of the sentence. Efforts have been made to preserve the equivalence in grammatical function between the Thai words and their English equivalents. However, this is not always possible. There are lexical items in Thai which do not have corresponding forms in English. Two examples of these are the classifiers and sentence particles. We will represent them simply as CL and PART, respectively, without making any attempt to identify their grammatical functions. At some points in our discussion information concerning grammatical categories becomes crucial to the discussion. We will, in such cases, use labels to identify those grammatical forms instead of translations. This will be true especially for discussions concerning complementizers and markers of tense and aspect modification. The label COMP will be used for the former. As for the latter, the label AUX^6 will be used, without implying any correspondence between these markers in Thai and what are usually referred to as auxiliaries in English.
NOTES TO CHAPTER I

1. In fact, this characterization, though convenient for the present exposition, does not reflect hierarchical relations among the properties; e.g., \([\text{generic}]\) is a property of \([-\text{definite}].\)


3. We assume along with Chomsky and Lasnik that this is in fact a shallow structure rather than a truly surface structure. However, since this is the structure to which the semantic interpretation rules are applied, the term does not seem inappropriate.

4. For Thai, this anaphor can be either a zero or a lexical form. The condition governing the choice between these two anaphoric forms is an issue which still needs further study. There are cases where the two forms are interchangeable, as in the sentence below:

   1. \text{khunpùu bòok wâa \{than\} pùathùa}  

      \text{Grandpa say COMP he headache}  

      'Grandpa said that he had a headache.'

   and cases where only the zero anaphoric form is allowed, as in the sentence below:
2. khunpùu mây yàak \{ \emptyset \} pay rooŋphayaabaan \{ *thân \}

Grandpa NEG want he go hospital

'Grandpa did not want to go to the hospital.'

It is possible that the obligatory zero anaphor is what is referred to in Chomsky and Lasnik (1977) as anaphor controlled by another NP in the same sentence. Readers are referred to Surintramont (1979) for discussion on this anaphoric representation.

5. N. Bandhumedha (1976) shows that there is no distinction in Thai between the Left-Dislocation Transformation and the Topicalization Transformation. Therefore, she considers them to be the same rule of Topicalization, since the copy left at the original site can be either a pronominal or zero anaphora.

6. Readers are referred to the works of Scoval (1970) and Kanchanawan (1978) for detailed analyses of tense and aspect modification in Thai, which makes use not only of lexical items but also of time phrases. The term AUX is used here only to refer to those lexical markers which are constituents of the verb in the predicate phrase. They occupy positions preceding as well as following the verb. The studies cited above do not provide a class label
to refer to these markers. The label AUX is used here with the same justification as that in Ekniov (1979).
CHAPTER II
SENTENCE-INITIAL NOUN PHRASE CONSTRAINT

2.0. Introduction

It is a well-accepted fact for Thai that the positioning in surface representation of constituents in a sentence is relatively fixed and plays a crucial role in the semantic interpretation of the sentence (Uppakitsillapasan 1937:9; Panupong 1970:11; Mahattanasin 1973:155; Rodman 1973:40; and B. Bandhumedha 1975:140). In this chapter, we will examine the positioning of noun phrases in a sentence. Our interest is focused especially on noun phrases which occupy the initial position of a sentence. We will argue that a noun phrase in this position always has the informational function of being the topic of the sentence. This, then, means that a noun phrase in this position must represent given information. We present data showing that when the subject happens also to be in sentence-initial position, it, too, functions as topic and is required to have the informational status required for topics. We will point out that the language excludes from initial position any subject which is not [+given]. This is clear from the behavior of existential constructions. We will propose for Thai syntax a Sentence-Initial Noun Phrase Constraint, or SINPC, to
account for the fact that only a [+given] noun phrase is allowed in sentence-initial position.

A study of the subject in sentences with distinct topics reveals that the proposed SINPC is applicable to the subject even when it is not in the sentence-initial position. We then propose that the SINPC be expanded to constrain not only the noun phrase in the topic position but any noun phrase in the position or positions preceding the predicate phrase. And by sentence-initial, we mean all such positions. The informational notions topic and comment are as basic to the syntax of Thai as such syntactic notions as subject, predicate, direct object, and so on. We will show that there are available in the language sets of special lexical markers which not only mark but also differentiate the different types of both topic and comment.

2.1. Types of Sentences Studied

The sentences included in the analysis in this chapter are those with only one main proposition. This rules out coordinate constructions, which contain more than one proposition joined by conjunctions such as lé? 'and', r̄̄ 'or' and t̄̄ 'but, as in (5) below:
The wind was strong and the waves were strong, too.'

However, we will consider as having one main proposition sentences with conjunctions linking noun phrase constituents as in (6) below.

(6) khawaamthùuk lé? khwaamphît baaqthii
right and wrong sometimes
mây tàqkan lêoy
NEG differ at-all

'Right and wrong do not differ at all sometimes.'

Our analysis includes embedded clauses such as (7) below,

(7) khraykhray màk khît wâa raw mii khwaamsùk
everybody tend think COMP we have happiness

'Everybody tends to think that we are happy.'

but these are not the main concern here. A more detailed analysis of embedded clauses will be presented in Chapter 3.
2.2. Subject and Topic

As a preliminary to our discussion of the informational function of the noun phrase in sentence-initial position, we will clarify our stand as to the notions topic and subject. We are in agreement with Li and Thompson (1976), who distinguish subject and topic in terms of the following criteria.

2.2.1. Definiteness

Subjects tend to be, but are not necessarily, definite in the sense of Chafe (1976); i.e., proper and generic nouns are also considered to be definite. Topics, however, are necessarily definite.

2.2.2. Semantic role

Subjects do not have a constant semantic role in the sentence (Li and Thompson 1976:464). It is even possible for some languages such as English to have a so-called "dummy subject." Topics have the constant semantic role of signaling the point of departure and telling what the sentence is to be about.

2.2.3. Relations with verbs

Subjects are always arguments of verbs. Topics, on the contrary, do not need to have any predicate-argument relation with verbs. Their case relations, in the sense of Fillmore (1968), are not constant. Unlike subjects,
topic choice is not predictable according to any hierarchical selection.

2.2.4. Agreement with verbs

If verbs are marked for agreement, they agree with subjects. There is no such agreement between the verb and the topic of a sentence.

2.2.5. Position in a sentence

Topics are associated with sentence-initial position in many languages. The position of subjects, however, is relatively free across languages. It can be found in various positions in relation to other constituents as in the formulae SVO, VSO and VOS, etc.

In Thai, the subject is the noun-phrase constituent usually occupying the pre-predicate position. For example, the phrase khunpùu 'grandpa' is subject in both (8) and (9).

(8) khunpùu mây chêɔp mamûaq

Grandpa NEG like mango

'Grandpa does not like mangoes.'

(9) mamûaq i khunpùu mây chêɔp ê i

mango Grandpa NEG like

'Mangoes, Grandpa does not like.'
Usually, the relation "subject" is associated with the case roles "agent" or "experiencer" as in (10) or (11) below.

(10) phûak nîṣît bìa ?aacaan wîlay
group student bored Prof. Wilai

'The students are bored with Professor Wilai.'

(11) ?accan bañ khon ?awcay dêk mâak pay
Prof. some CL pamper student much go

'Some professors pamper the students too much.'

As for topic, in Thai this function is usually associated with a noun phrase in initial position, preceding the subject, such as the phrase mamûaŋ 'mango' in (9) above.

Thai has been identified in Gsell (1979) as being both topic prominent and subject prominent; therefore, it is not surprising to find in the language such sentences as (12) below,

(12) kûaytîaw chán chûop sênlék tômyam
noodle I like small-noodle chili-soup

'Talking about noodles, I like the small kind in chili soup.'

in which the topic does not seem to have any direct semantic case relation with the verb. Rather, it has a part-whole relationship with an NP in the sentence.
2.3. Noun Phrase in Sentence-Initial Position

Earlier analyses of Thai syntax seem to assume or assert that a sentence can have either a basic or a variant ordering. The basic ordering can be represented as the linear ordering in (13) below:

(13) SUBJECT VERB DIRECT-OBJECT INDIRECT-OBJECT

phSomè hai khwaarmak kë leukluk
parents give love to children

'Parents give love to their children.'

or as a tree diagram showing the hierarchical grammatical relations of the constituents in (14) below.

(14)

In the basic ordering as represented in (13) and (14) above, each syntactic relation—subject, predicate, direct object, and indirect object—is defined configurationally.
A variant ordering occurs when a noun-phrase constituent is found in a different position. Panupong (1970) claims that this variation in basic ordering does not affect the meaning of the sentence, except for a slight difference in emphasis. She also claims that only direct objects can be moved or "anteposed" to initial position from its usual position after the verb. N. Bandhumedha (1976) claims that there is a transformational rule of Topicalization which moves the noun phrase from within the sentence to the S-initial position. Rodman (1973) considers this to be a copying rule rather than a movement rule, since it is possible for a pronominal copy to be left behind in the original site of the moved constituent. Gsell (1979) refers to this as a process of thematization.

A variant ordering of constituents in (13) and (14) can be either (15), in which the direct object is "moved" to the initial position:

(15) khwaamräk₁ phɔomɛ hây ə₁ kɛɛ lûukiûuk
love parent give it to children

'Love (is something) parents give to their children.'

or (16), in which what is "moved" is the indirect object. Notice that the dative preposition kɛɛ 'to' is absent from the phrase when it is in initial position:
There is a consensus among linguists who have studied this phenomenon that sentences such as (15) and (16) do not represent the basic ordering of sentence constituents and sentences such as (15) and (16) are considered Topicalized or Thematized sentences. The implication is that a sentence does not have a topic unless one of its noun-phrase constituents is copied or moved to initial position.

If topic is to be defined and interpreted as what a sentence is about, a sentence like (13) and (14) will then be a sentence without a topic, since it has not been topicalized.

A significant generalization is missing in this implication. The topic-creating rules of the earlier studies mentioned above uniformly move a noun phrase from within the sentence to the initial position. Therefore, the sentence-initial position can be associated with the function "topic." What has not been mentioned in these studies is the fact that the subject itself, usually occupying the initial position of the sentence, can also function as topic.

(16) $\text{kuk₁kuk₁}$ phawmɛ́e hây khwaamrák $\emptyset$

children parent give love it

'Children, their parents give love to them.'
A subject functions as topic if it is in sentence-initial position, just as any other noun phrase in the same position, as a comparison of the following two equational sentences shows. These sentences contain the equational copula khít which gives a definite identification (Gething 1979). Notice how the placement of different noun phrases can affect the semantic interpretation of the sentence. A corresponding focused sentence (to be discussed in Chapter 6) is provided to clarify the semantic interpretation of both equational sentences in (17a) and (18a).

(17)a. khon thîi khun khuan pay hâa khít
    person COMP you should go see be
    phûu?amnuaykaan sathâaban
director institute

'The person you should go to see is the director of the Institute.'

b. phûu?amnuaykaan sathâaban gjyla? khon
director institute FM person
thîi khun khuan pay hâa
COMP you should go see

'It is the director of the institute whom you should go to see.'
In (17a) the speaker is offering to his addressee some information concerning the issue of whom he should visit. Compare (17a) with (18a) below, in which the speaker is giving information about the director of the institute, namely, that he is the person the addressee should go to visit.

(18)a. phuu?amnuaykaan sathâaban khii khon
director institute be person
thii khun khuan pay haa
COMP you should go see

'The director of the institute is the person you should go to visit.'

b. khon thii khun khuan pay haa
person COMP you should go see
qayla phuu?amnuaykaan sathâaban
FM director institute

'He is the person you should go to see, the director of the institute.'

This observation that the sentence-initial noun phrase always signals what the sentence is about, regardless of its syntactic function, has been recognized in Phraya Uppakitsillapasan (1935:28), in which sentences are classified according to what its sentence-initial constituent is. His classification includes four types of sentences:
a. Subject sentence with the subject phrase in initial position
b. Object sentence with the direct object phrase in initial position
c. Verb sentence with the verb phrase beginning the sentence
d. Indirect object sentence with the indirect object phrase in initial position.

We believe that what Phraya Uppakit means by what "the speaker wants to foreground" may be close to our functional definition of topic as "point of departure" in the sentence.

Therefore, the implication that "topicalized" and "nontopicalized" sentences differ as to the positioning of constituents in a sentence is not valid. This is because the informational phrases (the topic and the comment phrase) and the syntactic phrases (subject, verb, direct object or indirect object) can all be positionally defined.

We can represent a so-called "topicalized" sentence like (19a) below as (19b):
(19)a. khamthāam khɔø nii₁ nisìt tɔɔp
    question  CL  this  student  answer
    PRO₁  mâyḍāay
    it  cannot

'This question the students cannot answer.'

b.

```
S
  NP
    N            CL            DEM
    khamthāam  khɔø  nii₁  nisìt  tɔɔp
  S'
    NP
      PRED
      V
      NP
      AUX
      PRO₁  mâyḍāay

question  CL  this  student  answer  it  cannot
```

The PRO₁ is bound by the topic. It is the constituent which the semantic rule of construal will coindex when the semantic interpretation rules apply to the surface representation of the sentence. The PRO₁ is what N. Bandhumedha (1976) and Rodman (1973) call the copy of the moved noun phrase, and it can be either zero or pronominal anaphor, as in (19c).
(19)c. khamthām khɔɔ nii₁ nislt tɔɔp
question CL this student answer
\{
\text{man}_1\}
mâydaay
\{
\emptyset_1\} it cannot

'This question the students cannot answer.'

As for "nontopicalized" sentences, or sentences in which the subject is also the topic, we represent them as in (20) below:

(20)a. khamthām khɔɔ nii₁ PRO₁ klɛŋŋ
question CL this give-a-hard-time
nákrian chātchāt
student clearly

'This question is clearly meant to give the students a hard time.'

b.}

\begin{center}
\begin{tikzpicture}
  \node (S) {S} [level distance=1.5cm, sibling distance=2cm];
  \node (NP) [level distance=1.2cm, sibling distance=1.5cm] {NP} [text width=3cm, align=center];
  \node (S') [level distance=1.2cm, sibling distance=1.5cm] at (S -| NP) {S'} [text width=3cm, align=center];
  \node (NP) [level distance=1.2cm, sibling distance=1.5cm] at (S' -| NP) {NP} [text width=3cm, align=center];
  \node (CL) [text width=1cm, align=center] at (NP -| NP) {CL};
  \node (DEM) [text width=1cm, align=center] at (NP -| NP) {DEM};
  \node (V) [text width=1cm, align=center] at (NP -| NP) {V};
  \node (NP) [text width=1cm, align=center] at (NP -| NP) {NP};
  \node (ADV) [text width=1cm, align=center] at (NP -| NP) {ADV};

  \node (khamthāam) at (NP -| NP) {khamthāam};
  \node (khɔɔ) at (NP -| NP) {khɔɔ};
  \node (nii₁) at (NP -| NP) {nii₁};
  \node (PRO₁) at (NP -| NP) {PRO₁};
  \node (klɛŋŋ) at (NP -| NP) {klɛŋŋ};
  \node (nákrian) at (NP -| NP) {nákrian};
  \node (chātchāt) at (NP -| NP) {chātchāt};
  \node (question) at (NP -| NP) {question};
  \node (CL) at (NP -| NP) {CL};
  \node (this) at (NP -| NP) {this};
  \node (it) at (NP -| NP) {it};
  \node (student) at (NP -| NP) {student};
  \node (clearly) at (NP -| NP) {clearly};
  \node (give a hard time) at (NP -| NP) {give a hard time};
\end{tikzpicture}
\end{center}
In line with all other noun phrases in a sentence, the subject may be a variable bound by the topic. It is possible, also, for (20a,b) to have a surface form with the PRO which is either a pronominal or zero anaphor, as in (20c):

(20)c. \textit{khamthām khēc nīi} \begin{align*} \text{question} & \quad \text{CL} \quad \text{this} \quad \text{it} \quad \text{give-a-hard-time} \\ \text{nākri} & \quad \text{chātchāt} \\ \text{student} & \quad \text{clearly} \end{align*} \\
\textit{'This question is clearly meant to give students a hard time.'}

We have been discussing topic as if it were always a noun phrase with a bound variable functioning as subject, direct object or indirect object of the sentence. It is possible for a topic noun phrase to have a bound variable which has a genitive relationship with one of the noun phrase in the sentence, as in (21) below:

(21) \textit{chāa nīi nī a} \begin{align*} \text{elephant} & \quad \text{CL} \quad \text{this tusk} \quad \text{it} \quad \text{break both} \\ \text{hāk thāqkhū} \end{align*} \\
\textit{'This elephant, its tusks are both broken.'}
We have also been discussing topics as if they were always noun phrases. Locative and temporal phrases can also function as topic, though it is not obvious that they are noun phrase. However, we give these phrases the noun phrase status for two reasons. Firstly, the prepositions which usually introduce these clauses become optional when they function as topics. Secondly, they can be followed by topic markers just like topics which are noun phrases. What seems to differentiate locative and temporal topics from noun phrase topics is the fact that the bound variables of locative and temporal topic are always zero while the bound variables of noun phrase topics can be either zero or a pronoun, as in (20c) above. Sentence (22) below is an example of a sentence with locative topic:

(22) (nay) ráan níi (nâ) khonkhãaykh tôq
   in shop this TM salesclerk
   yàapkhaay kàp lûukkâa bôybôy ñ
   rude with customer often there

' (In) that shop, the salesclerks are often rude to the customers.'

Sentence (22) is an example of a sentence with a temporal topic.
(23) (nay) pií náŋ₂ (nå) khon taaŋ kan
    in year that TM people die together
yè? g₁
lot then

'(In) that year, a lot of people died.'

There are sentences in which the topics seem "alien," i.e., the relationship that they have with the rest of the sentence is not too clear. (24) is an example of such sentences:

(24) (nay) kaancàppiaanāmlík₁ chaawprà?monŋ
    in deep-sea fishing fisherman
tșoŋ cháy ria lam yàŋ g₁
    have-to use boat CL big

'(In) deep-sea fishing, fishermen have to use big boats.'

The topic phrases of such sentences as (24) above are not usually introduced by a preposition. Gsell (1979) observes that it is usually possible to supply a preposition to introduce these "alien" topics in the surface representation. There is a tendency to do so especially in formal writing. These prepositions enable one to infer the location of the bound variable of these topics in the sentence. This bound variable is usually an adverbial phrase, the usual position of which is sentence-final,
as shown in (24) above. What seems to comprise the "alien" characteristic of these topics is probably the lack of pronominal anaphora to identify their source.

We acknowledge that the topics of sentences such as (22)-(23) can be optionally preceded by prepositions. However, they may be considered oblique noun phrases, since the preposition becomes optional when they occur in topic position.

2.4. Subject As Topic

In 2.3 we argued that a subject occupying the initial position in a sentence also functions as topic. In this section, we will present data to show that the constraint we have proposed for a topic noun phrase that it must represent [+given] information (that it has to be either [+definite] or [+generic]) also applies to the subject. There are two main points in our argument. One is that this informational constraint on topic phrases is also applicable to subjects which occur in sentence-initial position, as in (20) above. What this predicts is that there will be no sentence in which a subject noun phrase lacking in the informational properties requisite for a topic occurs in initial position. Existential sentences will be presented as supporting evidence. We will show that such unqualified subjects can occur in a sentence, but never in the initial position of a main sentence.
The second point in our argument is that this constraint governs not only the noun phrase in the position which begins a sentence but also all noun phrases in any position preceding the predicate phrase. What this predicts is that even when a sentence has a distinct topic, the subject is also required to conform to the same constraint on informational properties of sentence-initial phrases. That is, the subject of such a sentence will function as secondary topic to the higher topic. Therefore, any noun phrase which violates this informational constraint cannot occur in the subject position of the main sentences either.

The informational constraint we will discuss is the one which requires that all topic phrases have [+given] informational status. As discussed in 1.1.2, a phrase is [+given] if it has a definite referent in the sense of Chafe (1976). That is, the addressee must be able to identify the referent of the topic phrase. This means that only noun phrases which are definite—i.e., having the feature [+definite]—or indefinite but nonspecific—i.e., having the feature [+generic]—can occur as topics.

We propose to call this constraint the Sentence Initial Noun Phrase Constraint or SINPC. It can be stated as follows:
(25)a. No noun phrase with [-given] informational status can occur in surface sentence-initial position.

It is also possible to consider this constraint a filter in the sense of Chomsky and Lasnik (1977), which may be represented as follows:

(25)b. *[
        NP
        X]
S <-given> S

This filter rules out surface strings with [-given] initial NP. For convenience of exposition we will treat this as a constraint.

The effect of the SINPC is that the sentence will be ill-formed, if its initial noun phrase does not represent [+given] information, i.e., information that the addressee can identify. A [-given] initial noun phrase cannot be a felicitous point of departure for the communication of the information of the sentence. This is the case with sentences in which the initial noun phrase represents new information; i.e., it has an indefinite but specific noun phrase in initial position. When we test these sentences with native speakers, the response ranges from a definite rejection to a hesitant acceptance, with suggestions that the sentence be paraphrased by an alternative in which the indefinite but specific noun phrase is no longer in initial position. The strategy
to avoid such occurrence of an indefinite but specific noun phrase is to place before it mii, which we argue is an existential verb which allows subjects which represent new information but also requires that its subjects be extraposed to the position after it. The argument on this issue will be given in detail in Chapter 4.

We presented sentences like (26) to native speakers to find out if they think these sentences are acceptable.

(26) ?*nuu nèg tua mii tâay tó? tua nîi
snake one CL exist under table CL this

'One snake is under this table.'

All respondents agree that the sentence would be well-formed if the verb mii were placed in front of it, as in (27) below:

(27) mii nuu nèg tua tâay tó? tua nîi
exist snake one CL under table CL this

'There is one snake under this table.'

Further examination of existential constructions in Thai reveals that for an initial noun phrase consisting of a noun, a numeral, and a classifier, such as lûukkhâa sâ̄̄gkhon to have an indefinite but specific reading, the sentence has to be embedded beneath the existential verb mii. Compare the occurrence of
lūukkhāā ʂɔq khon in (28), in which it is embedded under mii and has an indefinite but specific referent,

(28) mii lūukkhāā ʂɔq khon maa thǎam
exist customer two CL come ask
hãa thúrian
search durian

'There were two customers coming in to ask for durian.'

and (29), in which it has a definite referent.

(29) lūukkhāā ʂɔq _ khon maa thǎam hãa
customer two CL come ask search
thúrian
durian

'Two of (our) customers came to look for durians.'

When the noun phrase is indefinite but nonspecific, it can occur in initial position. Compare the noun phrase nōkkāŋkhēŋ n̥ŋ tua 'one magpie' in (30):
(30)a. *nókkäŋkʰeēn ɲɛŋ tua kɔ? yùu bon
magpie one CL perch stay on
kiŋ mamuaŋ
branch mango

'One magpie is perching on a mango branch.'

b. mii nókkäŋkʰeēn ɲɛŋ tua kɔ?
exist magpie one CL perch
yùu bon kiŋ mamuaŋ
stay on branch mango

'Perching on a mango branch is one magpie.'

and nókkäŋkʰeēn tua ɲɛŋ 'a magpie' in (31) below.

(31)a. nókkäŋkʰeēn tua ɲɛŋ kɔ? yùu bon
magpie
kiŋ mamuaŋ
branch mango

'A magpie is perching on a mango branch.'

b. mii nókkäŋkʰeēn tua ɲɛŋ kɔ?
exist magpie CL one perch
yùu bon kiŋ mamuaŋ
stay one branch mango

'There is a magpie perching on a mango branch.'
It is apparent that the embedding of a sentence with an indefinite but specific noun phrase within an existential sentence is a strategy available in the language to preclude such unqualified noun phrases from occurring in the surface-initial position of a main sentence.

Another strategy is evident when some of our respondents suggest that sentence (30a) be replaced by (32), in which the noun phrase in initial position is not quantified and thus has an indefinite and nonspecific reading. The quantifier is placed instead at the end of the sentence as in (32) below:

(32) nôkkâägkhëen kò? yùu bon kìq
magpie perch -ing on branch
mamûaŋ nìŋ tua
mango one CL

'There is one magpie perching on a mango branch.'

We are not certain what the grammatical function of such a quantifier phrase in final position is. Recognizing that the issue deserves more extensive study, we will simply assume that it may function as some type of predicative adverbial phrase.

The data presented above clearly show the effect of the proposed SINPC on subjects which occur in initial position and, in so doing, function as the topic of the
sentence. Further study of subjects in sentences with separate distinct topics, however, reveals that even when not in initial position, subjects are still governed by the SINPC, originally proposed to account for the informational status of noun phrase topics.

Sentence (33a) is ungrammatical because its subject is an indefinite but nonspecific noun phrase, representing new information. To derive the grammatical surface string the subject has to be extraposed to the position after the existential verb mii, as in (33b).

(33)a. *khàawliit lāay krasè mii klàwkàp
    rumor many CL exist concerning
    khadii níi
    case this

b. mii khàawliit lāay krasè klàwkàp
    exist rumor many CL concerning
    khadii níi
    case this

'There are many rumors concerning this case.'

One would expect that such a sentence as (34a) below would be grammatical since the sentence itself has a definite noun phrase occupying the topic position. That is, (34a) should be a grammatical sentence, if the SINPC governs only the noun phrase in the topic position. However, we
find that it is not grammatical and that the SINPC also has an effect on the subject phrase as well, even when it is not in the initial position of a sentence, as can be seen in a comparison of the a. and b. sentences in (34) below.

(34)a. *khadii níi khàawlití l̈ay krasĕ
*case this rumor many

mii
exist

b. khadii níi mii
*case this exist
khàawlití l̈ay krasĕ
rumor many CL

'(Concerning) this case, there are many rumors.'

Compare also the a. and b. sentences of (35), in which a sentence with an indefinite but specific subject cannot occur as the subject of the comment of the sentence.
This boy, some people have seen him stealing things.

Notice that the same strategy of embedding a sentence like the comment of (35), in which the subject is indefinite but specific, in an existential construction is also applied to the sentence which functions as S' or comment of the higher sentence.

Therefore, we propose to expand the proposed SINPC to apply not only to noun phrases in topic position but also any noun phrase in subject position. That is, we propose that in "sentence-initial positions," we include any position preceding the main predicate phrase of the sentence. The SINPC as stated in (25), will be revised as (36a) below:
(36)a. Sentence-Initial Noun Phrase Construct (SINPC)

No NP with [-given] informational status can occur in any position preceding the main predicate\(^2\) in the surface structure.

It can be represented as a filter:

(36)b. \[
S \leftarrow \text{given}
\]

It is possible to find more than two noun phrases occurring in "initial position" of our definition. There are three noun phrases in the "initial" or pre-predicate positions in (37).

(37) bânánníi 1ûuklûuk\(i\) gaankaan mây
family this children work NEG
?aw \(q\_i\) că? ?aw tèt thîawlên
take they will take only play
thâadiaw
only

'(In) this family, their children do not pay attention to work. They are interested only in having fun.'

Sentence (38) has four noun phrases in its "initial positions."
What this seems to suggest is that Thai allows sentences with multiple topics. Each of the topics that make up the composite topic is constrained by SINPC as well. The subject of a sentence in which there are other distinct topics is itself a minor topic of the sentence. This seems to be a possible explanation as to why the subject is also constrained by the SINPC.

2.5. Basicness of Informational Function

We have shown in 2.3 that the informational phrases topic and comment can be as well defined positionally as such syntactic phrases as subject, predicate, verb and direct object. However, while there are no specific markers to mark syntactic phrases such as subject or object, there are lexical markers available in the language to mark the informational phrases, topic and comment.

The following are lists of topic markers (TM) and comment markers (CM).
Below are examples illustrating where in the surface structure these markers occur:

(40) khwaampentham nay sāŋkhom nán campen justice in society TM necessary yīŋ very

'Social justice is very important.'

(41) taa khon nán nâ? tèŋqaan léew fellow CL that TM marry already ná? CM

'That fellow is a married man.'
These markers can be constituents of their respective structures. P.S. Rule 1, proposed in (4), may be revised to incorporate the optional occurrence of topic and comment markers, as in (42) below:

(42) P.S. Rule 1: \[ S \rightarrow (NP \quad (TM)) \rightarrow S'(CM) \]

The topic markers are not obligatory. When they occur, they add emphasis, calling special attention to the topic the speaker wishes to address. It should be emphasized that it is not the occurrence of a TM that makes a noun phrase a topic.

It is apparent that there is similarity, phonologically as well as semantically, between the first two topic markers in the list and the demonstrative form, \textit{nán} 'that', and between the last three and the demonstrative form, \textit{níi} 'this'. However, topic markers and demonstratives have distinct and separate functions, as evident in sentence (41), in which both occur but perform different functions. The demonstrative \textit{nán} 'that' modifies the classifier \textit{khon} and together they modify the head noun \textit{taa} 'fellow'. The topic marker \textit{nâ} co-constitutes with the topic phrase \textit{taa khon nán} 'that fellow'. It is noticeable that while the topic markers exhibit phonological variation, the demonstratives do not. The topic markers \textit{nán} and \textit{nâ} are variants, the choice of which is determined by speech style. The former is used for
written and very formal spoken language. The latter is used only in casual and rapid speech. The níi, nî? and nîâ? topic markers can also be differentiated in terms of the speech style in which they occur. níi is used for formal speech, while the other two are used for less formal speech.

The choice of topic marker is independent of the choice of demonstrative. A demonstrative and a topic marker of similar form can occur together in the same sentence, as in (41) above. Or, they may have different forms, as in (43) below:

\[(43) \text{camnuan } níi \text{ nán náp chaphó?}\]
\[\text{number this TM count only}\]
\[\text{phasãa sãmkhansãmkhan}\]
\[\text{language important}\]

'This number includes only important languages.'

A topic can also have markers not included in our list (39). These are topic markers of different type and include the prepositions suan 'as for' and sãmràp 'for'. They form a subgroup of markers which not only emphasize the topic but also indicate that the speaker has chosen to pick this particular topic from among other possible topics that he could talk about. That is, we think that this is a case of contrastive topic, as defined
in Chafe (1976). Below are two examples of these prepositional topic markers:

(44) สุน พหูผทพhayop ที่ฝิลก้า ทาย
TM refugee remaining Thailand
ตพ รํา รําว วาย คํ่า
have-to receive take keep self

'As for the remaining refugees, Thailand has to keep them herself.'

(45) สําmrํา ทําฟหอมคํ่า นํา ฟํม คํ่า
TM myself TM I probably
may ผียน คําย นี่สี่
NEG change mind certainly

'As for myself, most likely I will not change my mind.'

As for the comment markers, two of them, นา? and la?, mark comments in formal and casual speech style, respectively, as in (46) and (47) below:

(46) กษัณิ ผาย ที่stan นํา สำัน ตั้สงิน
case log illegal TM court sentence
pay ริ้าปปํ ลําว นา?
go fine already CM

'The court has already prosecuted the illegal logging case.'
(47) chaam bay níi chán láaŋ léew la?
dish CL this I wash already CM

'I have washed this dish.'

dòok, ròok, and lòok are variants in different speech styles. The first is a written form hardly ever heard in speech. The second and the third are for spoken language. The alternation between r and l is very common in Thai. It has been claimed that when this alternation occurs, the use of r differentiates the careful speech of educated speakers from the less educated speakers who use l. These comment markers are easily distinguished from the first two on the list, ná? or lá?, which are used in sentences that simply inform. This act of informing can be done in either a positive or a negative sentence, as in (48) and (49) below:

(48) kaafee mōt léew ná?
coffee end already CM

'We have run out of coffee.'

(49) naŋsíi lêm níi khâw khîan mây dii
book CL this he write NEG well CM

'ná?

'This book is not well written.'
Compare the use of nà? above with the use of dòok and lòok in (50) and (51) below:

(50) kháw tèŋqaan kan lëcw lòok
they marry together already CM
'They ARE3 married.'

(51) naŋší lêm níi khýan máy dii dòok
book CL this write NEG well CM
'This book is NOT well written.'

dòok and lòok are used when the speaker would like to contradict a certain idea he assumes the addressee may have. They can also be used for directly contradictory responses.

The nà? and dòok sets provide markers for emphatic comments, parallel to those of emphatic topics.

The last three comment markers in list (39) are contrastive comment markers. Their pattern of distribution is similar to that of emphatic comments. nay and naylâ? are used when the speaker wants to inform, while tàañhàak is used when he wants to contradict his addressee's idea, as can be seen in the following two sentences:
The contrastive markers are obligatory in focused constructions, in which the contrastive noun phrase occurs in sentence-initial position as in (54):

(54) bānn lă̂ng nîi ɲìay lâ? thîi khâw
    house CL this FM COMP they
cà? khâay
will sell

'It is this house which they are selling.'

We will use the symbol FM, focus marker, for this particular type of comment in focused construction, which will be discussed in detail in Chapter 5.
To sum up, we have presented data from the language which show that the informational functions topic and comment are as basic to the syntax of the language as the syntactic functions subject, object and predicate. The existence of sets of special markers which can be employed for each of these two informational functions, topic and comment, strengthens our thesis as to their saliency.

2.6. Postposed Topics

We have argued that sentence-initial position is a syntactic means of signaling the topic of a sentence. There is an important exception to this generalization. There are sentences in which topics occur in the surface representation as the S-final noun phrases. (55) below is an example of such a sentence.

(55) ɲəm məy ɣàak ɣʊŋduay (lòɔk)
I NEG want deal CM
// lʊuksɔaw khun
daughter you

'I don't want to have anything to do with her, your daughter.'

This sentence can be paraphrased as (56), in which the topic occurs in the initial position:
Notice that the topic lûuksāaw khun 'your daughter' in final position follows the comment marker lèok. It is usually separated by a pause, represented by the symbol // in (55). The pause reflects the afterthought manner in which topic is reinstated. We propose to account for this phenomenon in terms of optional transformation of Topic Postposing, which can be formulated as (57) below:

(57) SD: NP S'
     1 2
SC: Ø, 2 + 1

The existence of sentences with postposed topics is simply another available syntactic device for the manifestation of informational functions. Furthermore, the postposed topic is always recognizable by a special manifestation such as a pause or its positioning after the comment marker, which is otherwise the final element of a sentence.
2.7. Some Conclusions

In this chapter we have argued that the informational functions, topic and comment, are as basic to the syntax of Thai as the syntactic functions of subject, predicate, object, and indirect object. We have presented data to show that both of the informational and syntactic phrases in a Thai sentence can be positionally defined.

We have also argued that topic is marked structurally by being in sentence-initial position; therefore, there is no need to distinguish a topicalized from a non-topicalized sentence. A subject can function as topic when there is no other distinct topic and it occurs in surface initial position.

We have discussed a constraint on the informational status of the noun phrases which can occur in the topic position. We have shown that this constraint, which we refer to as the SINPC or Sentence-Initial Noun Phrase Constraint, governs not only the noun phrase in the first position of the sentence in the surface representation but all noun phrases preceding the predicate phrase of the main sentence.

In the next chapter, we will discuss the effect of the SINPC on the occurrence of embedded clauses in sentence-initial position.
NOTES TO CHAPTER II

1. Phraya Uppakitsillapasan (1937) is an exception to this.

2. By main predicate, we mean the predicate phrase of either a simplex or the matrix sentence. A [+specific] noun phrase representing [-given] information such as dek sãam khon 'three boys' cannot occur in the position preceding the predicate phrase of a simplex sentence, as in (la), but it can precede the predicate phrase of a clause embedded as subject of mii 'exist', as in (lb) below:

   la. *dek sãam khon deoën maa
        boy three CL walk come

   b. mii  
       dek sãam khon deoën maa
       exist boy three CL walk come

      'There are three boys walking this way.'

3. The upper-case letters and the underlining represent emphatic stress in the sentence.
CHAPTER III
EMBEDDED CLAUSES: SYNTACTIC FORM AND
INFORMATIONAL STATUS

3.0. Introduction

In this chapter we examine the interaction between embedded clauses and the Sentence Initial Noun Phrase Constraint or SINPC proposed in Chapter 2. We examine in particular the fact that, of all embedded clauses, only thi clauses and subjectless generic clauses, as in (58) and (59) respectively, can occur in sentence-initial position:

(58) thi khaw yakan day na? chan henggay
COMP they divorce can TM I approve

'I approve of the fact that they could settle for a divorce.'

(59) phutphisyphi phi pen baap kapa paka
talk carelessly be sin with mouth
tua?eeq plawplaw
oneself useless

'Careless talking is futile and a sin to one's own mouth.'

We argue that this is because only these two types of embedded clauses have the informational status [+given]
and conform to the SINPC. Other types of embedded clauses represent [-given] information and thus cannot occur in sentence-initial position.

3.1. Terminology and Assumptions

The term *embedding* is used to refer to the occurrence of a sentence as a constituent within another sentence, either as its topic, subject, or object, or as a modifying clause within a noun phrase. The term *clause* is used to refer to embedded sentences to distinguish them from main sentences. When the functions of embedded clauses are crucial to the discussion, terms of more specific nature such as *sentential topic*, *sentential subject*, *sentential object* and *relative clause* are used.

It is assumed that all embedded clauses discussed here are dominated by the node NP. That is, there is no distinction between verb phrase complements and noun phrase complements as suggested in Rosenbaum (1967). However, a distinction exists between embedded clauses which occur as relative clauses, as in (61a), and those which occur as major sentence constituents functioning as topic, subject, and object, as in (61b):
A relative clause always contains a gap or a pronoun which is construed as coreferential with the head noun phrase, as in example (62) below:

(62) khon₁ thîi ə₁ khâav prakanchiivít ʰāy
       kháw₁

man COMP he sell life-insurance give

khun klâp maa mây ʔìíklèsw
you return come anew again

'The man who sold the life insurance policy to you came back again.'
Note that in (61) we showed the head noun phrase as an optional constituent. This is to account for headless relative clauses like the one in (63) below. We will represent the absence of the head noun phrase in this example by $\emptyset$. The gap or zero anaphor within the relative clause itself will also be represented by $\emptyset$.

(63) \[ \emptyset_i \text{ thi } \emptyset_i \text{ yu } \text{ khâ qnâa} \]
\[ \text{COMP you see it stay front} \]
\[ \text{khun nân khîî chàak kaancalaacon} \]
\[ \text{you TM be scene riot} \]
\[ m\text{awaannîi} \]
\[ \text{yesterday} \]

'What you see in front of you is a scene from yesterday's riot.'

A more detailed discussion of headless relative clauses is given in Chapter 5.

Complement clauses as shown in (61b) differ from relative clauses in that there is no requirement of coreference between the head noun phrase and a referent understood but not expressed inside the embedded clause. Usually, this type of embedded clause does not have a head noun phrase, as in (64) below:
'That they came late was because of the traffic jam.'

However, there are abstract noun phrases which can function as head noun phrases for such constructions. The most frequently occurring ones are rīaŋ, kaaŋ, khwaam, all translatable roughly as 'matter, fact, state of affairs, state, situation'. With one of these head noun phrases, (63) will be like (64) below:

(64) thīi khāw maa sāay pen phrö? COMP they come late be because rōt tīt car stick 'That they came late was because of the traffic jam.'

3.2. Previous Analyses of Embedded Clauses in Thai

Previous analyses of embedded clauses in Thai deal primarily with the structural property of these clauses. Both relative clauses and complement clauses have been studied in detail. Ekniyom (1971) and Sornhiran (1978)
deal with relative clauses while Suwannamalik (1974) and N. Bandhumedha (1976) deal with complement clauses.

Ekniyom (1971) argued that relative clauses in Thai can be divided into three types: restrictive, appositive, and sentential relative clauses. The distinction is made on the basis of the head these clauses modify, the grammatical function of the relative clause in the sentence, the choice of relative pronouns (or complementizers in our terminology), and the underlying structure. Only the restrictive type of claimed to be generated in the underlying structure as an embedded clause. The other two are generated as conjoined sentences which are moved into the main sentence by transformational rules. We agree with this classification and claim about the underlying difference between the restrictive and the other two types of relative clauses. However, we now take the position that the relative clause is embedded as a constituent of only noun phrase, rather than of many different types of heads as suggested in Ekniyom (1971).

Sornhiran (1978) claims to refute the distinction made in Ekniyom (1971). However, her own classification is not very different from Ekniyom's. Sornhiran also distinguishes three types of relative clause: an anaphoric type, containing an anaphoric element coreferential with the head nominal antecedent; a nonanaphoric type, containing no such element; and a nonnominal type,
the head noun of which can be either a verb phrase, a sentence, or even a sequence of sentences instead of a single sentence as suggested in Ekniyom (1971).

There are also two works which deal with complement clauses quite extensively: Suwannamalik (1974) and N. Bandhumedha (1976). In the first, three types of complement clause (or complementized clause in Suwannamalik's terminology) are recognized and are used to subcategorize complement-taking verbs. N. Bandhumedha (1976) argues that of the three types of complement clause recognized by Suwannamalik, only the háy and wâa clauses are true complement clauses. The thîi clauses, she argues, can occur with the head noun phrase râaŋ, and should, therefore, be considered relative clauses, even if there is no gap or reference within the embedded clause understood to be a noun phrase coreferential with the head noun phrase. Apparently, N. Bandhumedha overlooks the nature of the general distinction between relative clauses and noun-phrase complements. She proposes, following Kiparsky and Kiparsky (1970), a transformation to delete the head noun phrase.

Little attention, however, has been paid to embedded clauses not introduced by a lexical complementizer. Such clauses can function both as sentential objects, as in (66) below:
(66) phôm ̀ yàak ̀ khuy ̀kap ̀khun
I want I talk with you
sák ̀ thíi
about once

'I want a chance to talk to you sometime.'

or as sentential topics, as in (67):

(67) thíăng ̀kap ̀khon ̀mây ̀míi ̀hèetphôn
argue with people NEG have reason
phôm ̀ b̀za
I tired

'Arguing with unreasonable people--I am tired of it.'

or even as relative clause, as in (68) below:

(68) khon ̀khàp ̀ròtpràcamthåaŋ ̀nay
person drive bus in
krugthëep ̀khîmoohôo ̀kè̄q ̀catalay
Bangkok irritable well extremely

'Drivers of buses in Bangkok tend to be very irritable.'

To facilitate the discussion of our proposed analysis, we will adopt the convention of representing these clauses as having a zero complementizer. To differentiate this use of zero as a complementizer from the zero marking a
nonlexical anaphor, we will underline the zero when used as complementizer and index the zero which is used as an anaphoric form. The following example illustrates this convention.

(69) dirêek̂ i bôck wâa ø̂ i yàak ø̂
    Dirake say COMP he want COMP
    pay ?ameerikaa
    go U.S.

'Dirake said that he wanted to go to the U.S.'

3.3. Proposed Analysis of Embedded Clauses in Thai

Except for our recognition of the zero complementizer, and our assumption of the underlying structure of embedded clauses as presented in (61), we agree generally with the analyses of the structural properties of both relative and complement clauses in Thai proposed by the authors cited in 3.2 above. In our study we are interested in how the differences observed in syntactic properties and distribution reflect difference in informational status. The distinctions we make for embedded clauses in Thai are those relevant in showing how the structural differences reflect informational differences.
3.3.1. Distinctions in sentence mood

Sentence mood is usually associated with the categorization of sentence types in traditional grammatical description. Grammatical mood is usually marked on the verb form. For English, such labels as Indicative and Subjunctive have been used. Jacobs (1981) shows that certain mood contrasts are central to the syntax of English. He posits a major contrast between indicative and hypothetical moods. The basis for this distinction in English is the presence or absence of verb inflection in a sentence. He cites evidence to show that this distinction is motivated both semantically and syntactically and that it applies to both main and embedded clauses. Indicative sentences are defined as having predicates with tense modification and referring to states of affairs that are realis, i.e., that have real time reference of their own. Hypothetical sentences do not allow such tense modification. As a main sentence, a hypothetical sentence has a time reference which is dependent on the time of utterance, the time typically being concurrent or subsequent to the utterance time. As an embedded clause, a hypothetical sentence has a time reference which is dependent on the time of the predicate of the matrix sentence. The situation or event referred to in a hypothetical clause is irrealis, i.e., not asserted to be a real situation.
We believe that this mood contrast as posited for English by Jacobs also holds for Thai. It is true that verbs in Thai do not change form and that tense modification is manifested by the use of lexical markers occurring in preverbal as well as postverbal positions. (Readers are referred to Scovel (1970), Kanchanavan (1978) and Ekniyom (1979) for details on the use of lexical markers for tense modification in Thai.) Tense can also be marked by the use of time adverbials. Moreover, time reference may not always have an overt marking and may be inferred from the context. In the face of such structural facts about tense modification in Thai, we propose a criterion for the distinguishing of indicative and hypothetical sentences in Thai in terms of free and predictable tense modification. Indicative sentences, having independent time reference, allow tense modification freely. It is possible for any lexical marker for tense to occur in the predicate phrase of an indicative sentence. (70) is an example of an indicative sentence as main clause in Thai. Notice that the sentence can take free tense modification. (70a) has a marker for past tense (PAST) khaay . . . liew while (70b) has a marker for future tense (FUT) ca?. 
The same free tense modification is possible for an indicative sentence as embedded clause, too:

(71)a. kháw ṭèat wâa kháw ḫeey pay míaqnsôk lezw
he boast COMP he PAST go abroad PAST

'He boasted that he had been abroad.'

b. kháw ṭèat wâa kháw cà? pay míaqnsôk
he boast COMP he FUT go abroad

'He boasted that he was going abroad.'

Hypothetical sentences, having dependent time reference, do not allow any overt independent tense modification, especially as main clause:
(72)a. khâw maa khâaqnay bân kôn khrâp
enter come inside house first PART

'Come in the house first, please.'

b. *cà? khâw maa khâaqnay bân kôn
FUT enter come inside house first
khrâp
PART

c. *khêey khâw maa khâaqnay bân
PAST enter come inside house
kôn khrâp
first PART

As an embedded clause, a hypothetical sentence does not allow any tense modification except the future tense marker cà?, when the time reference is subsequent to the time reference of the matrix clause, as in (73) below:

(73)a. raw khêey yâak ʊ pay mîaânSok
we PAST want COMP go abroad
mîankan'
too

'We once wanted to go abroad, too.'
The distinction between indicative and hypothetical is informationally significant. With free tense modification an embedded indicative clause is capable of presenting information with definite time reference. An embedded hypothetical clause, having predictable and dependent time reference, cannot give any independent time reference freely. This explains why an indicative relative clause, as in (74a), has a definite, restrictive meaning, while a hypothetical relative clause, as in (74b), has a generic meaning. 2
3.3.2. The assertive-nonassertive distinction

The distinction between assertive and nonassertive clauses applies only to indicative embedded clauses. The terms assertive and nonassertive are borrowed from Hooper and Thompson (1973) and Hooper (1975), where they refer to two types of predicate in English. Assertive predicates allow the proposition of the embedded clause to be the main proposition of the sentence and mark the proposition of the matrix sentence as merely a parenthetical proposition.

Indicative clauses in Thai fall into two classes. Assertive clauses are clauses which represent [-given] information and constitute independent propositions in addition to the proposition of the matrix sentence. (75) is an example of an assertive embedded clause:

(74)a. náam thîi khêŋ
   water COMP solid
   'the water that has frozen up'

(74)b. náam ū khêŋ
   water COMP solid
   'ice'
That woman believes in the notion that her husband never cheats behind her back.'

Nonassertive clauses are embedded clauses which represent [+given] information. That is, the proposition they represent is assumed to be an established fact and assumed to be information possibly known and shared by both speaker and addressee. Compare the assertive clause in (75) above with the nonassertive clause in (76) below:

'That woman believes the fact that her husband never cheats behind her back.'

The informational distinction between assertive and nonassertive indicative clauses is not formally unmotivated. The clauses themselves exhibit the distinction in their structure and distribution, too.

First of all, they are introduced by different complementizers. Nonassertive clauses are introduced by the complementizer thī while assertive clauses are
introduced by the complementizer wâa or by a zero complementizer.

Secondly, they are distinguished distributionally by the class of verbs they can occur with. The verbs which take nonassertive clauses are of two distinct types. The first type comprises verbs which denote emotional or psychological reactions. It will be convenient to refer to clauses embedded beneath these verbs as reactive clauses. The following are examples of verbs taking reactive clauses.

(77) Verbs taking reactive thîi clauses:

bìa 'bored, tired' phòócay 'pleased'
chîa 'believe' plèìkcay 'surprised'
chôop 'like, enjoy' ramkhaan 'annoyed'
chaqôn 'puzzled' raŋklât 'hate, despise'
dîicay 'glad' sîadaay 'regret'
khâwçay 'understand' sîacay 'sad'
kîlât 'dislike' tôkcay 'frightened'

Verbs of a second "mental" type also takes thîi clauses. These verbs are essentially forward-looking and thîi clauses occurring with them are always marked for future; i.e., they have the preverbal tense marker cà?. These clauses are obligatorily subjectless but are understood to have a subject which is controlled, in the sense of Chomsky and Lasnik (1977), by the subject
of the matrix sentence. We will refer to these as *anticipatory clauses*. (78) is an example of such clauses.

We will mark the missing subject with a \( \emptyset \).

\[
(78) \text{kháw}_i \ sáňyaa \ \text{thî}_i \ \emptyset_i \ cà\ ? \ \text{pen} \\
\text{he} \ \text{promise} \ \text{COMP} \ \text{he will be} \\
\text{nákkaanmiąq} \ \text{bèsp} \ \text{mày} \\
\text{politician} \ \text{type} \ \text{new}
\]

'He promised to be a politician of the new type.'

These clauses are in contrast with the reactive *thî* clauses, which need not be subjectless and which may be marked for any time reference. 4

Examples of verbs which take anticipatory *thî* clauses are:

\[
(79) \text{Verbs taking anticipatory *thî* clauses:} \\
\text{khít} \ 'think' \ \text{bža} \ 'bored, tired' \\
\text{patisèet} \ 'refuse' \ \text{chšop} \ 'like' \\
\text{plľancay} \ 'change one's mind' \ \text{diicay} \ 'glad' \\
\text{ráp} \ 'accept' \ \text{phoocay} \ 'pleased' \\
\text{sáňyaa} \ 'promise' \ \text{raŋklat} \ 'hate' \\
\text{tàtsîncay} \ 'decide' \ \text{sỳadaay} \ 'regret' \\
\text{wąq} \ 'hope' \ \text{sỳacay} \ 'sad' \\
\text{waŋphęen} \ 'plan' \ \text{sâwcay} \ 'grieved' \\
\text{yàak} \ 'want' \\
\text{yoområp} \ 'admit, accept'
\]
Notice that some of the verbs in the right-hand column also take reactive thīi clauses.

The verbs which take assertive clauses are also divided into two types: those which take only clauses introduced by wāa, and those which take either clauses introduced by wāa or clauses with the zero complementizer. The following are examples of verbs which take assertive clauses introduced by wāa.

(80) Verbs taking assertive wāa clauses:

<table>
<thead>
<tr>
<th>Verb</th>
<th>Meaning</th>
<th>Verb</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>banyaay</td>
<td>'describe'</td>
<td>law</td>
<td>'tell'</td>
</tr>
<tr>
<td>baqkháp</td>
<td>'require'</td>
<td>ne?</td>
<td>'suggest'</td>
</tr>
<tr>
<td>chìfa</td>
<td>'believe'</td>
<td>phísùut</td>
<td>'prove'</td>
</tr>
<tr>
<td>chíicêsq</td>
<td>'explain'</td>
<td>phûut</td>
<td>'say'</td>
</tr>
<tr>
<td>fän</td>
<td>'dream'</td>
<td>róqq</td>
<td>'cry'</td>
</tr>
<tr>
<td>hâam</td>
<td>'forbid'</td>
<td>sóon</td>
<td>'teach'</td>
</tr>
<tr>
<td>kháan</td>
<td>'object'</td>
<td>sàq</td>
<td>'command'</td>
</tr>
<tr>
<td>klàawhāa</td>
<td>'accuse'</td>
<td>thāam</td>
<td>'ask'</td>
</tr>
<tr>
<td>koohòk</td>
<td>'lie'</td>
<td>takoon</td>
<td>'shout'</td>
</tr>
</tbody>
</table>

Additionally, the verbs taking reactive thīi clauses listed in (79) can also take wāa clauses. The choice of complementizer affects the semantic interpretation of the sentence. When the verbs in (79) take a thīi complement clause, they represent the emotional reaction of an individual to a situation or event, as in (81) below:
When the same verb takes a wåa complement clause, it represents an elaboration or an explication of why the individual should have such an emotional reaction. That is, there are two assertions being made in a sentence like (82) below. The proposition of the matrix sentence makes one assertion about the emotional reaction. The embedded clause, being assertive, adds another assertion as to the reason for the reaction.

There are two verbs which can take either wåa clauses or clauses introduced by the zero complementizer: wåa 'say, be of opinion, suggest' and bòok 'tell, inform'. The homophony of the verb wåa and the complementizer wåa is obvious; and they may be related. The act of assertion inherent in the semantic meaning of both the verbs wåa and bòok can be assumed to be a factor which explains why one can have an assertive clause without any lexical complementizer, namely a zero clause, for complement.
It is possible to consider this zero complementizer a case of optional complementizer after these two verbs. However, we propose to treat wâa clauses and zero complementizer clauses as two separate types of clause for two reasons. One is that they are not interchangeable in relative clauses such as (83), even though they are interchangeable in complement clauses after wâa and bòk, as in (84).

\[(83) \text{khon}_i \emptyset \emptyset_i \text{khöey} \text{òkhàk}_i^6\]

\[
\begin{array}{llll}
\text{person} & \text{COMP} & \text{he} & \text{PAST} \\
\*\text{wâa} & \text{khâw}_i & \text{broken-hearted} \\
\end{array}
\]

'a broken-hearted person'

\[(84) \text{khâw} \ wâa \ wâa \ \text{raw} \ \text{caydam} \ \emptyset\]

\[
\begin{array}{llll}
\text{they} & \text{say} & \text{COMP} & \text{we} \\
\text{unkind} & \emptyset \\
\end{array}
\]

'They said that we are unkind.'

The second reason is that they are not interchangeable in existential sentences, as in (85) below:

\[(85) \text{mii} \ *\text{wâa} \ \text{khon} \ \text{khâw} \ \text{maa} \ \text{lësw}\]

\[
\begin{array}{llll}
\text{exist} & \text{COMP} & \text{people} & \text{enter} \\
\emptyset & \text{come} & \text{already} \\
\end{array}
\]

'There are people coming in already.'
A third difference between nonassertive thîi clauses and assertive wâa or zero complementizer clauses is topic eligibility. Only thîi clauses can occur in sentence-initial position, being nonassertive and representing information assumed to be established fact. (86) is an example of a thîi clause in sentence-initial position:

(86)a. thîi khâw mây maa nâ? dii lësw
COMP they NEG come TM good already

'It was good that they did not come.'

b. dèk khon nîi thîi këe maa
c. thîi ke:.
dii lësw
child CL this COMP he come
good already

'This boy, it was good that he came.'

wâa clauses, asserting additional proposition, cannot occur in sentence-initial position, either as topic or subject:

(87)a. *wâa khâw sabaaydiì dii lësw
he well good already

b. *dèk khon nîi wâa khâw sabaaydiì
boy CL this COMP he well
dii lësw
good already
To sum up, we have distinguished two types of indicative clauses in terms of their informational status: nonassertive clauses, representing [+given] information, and assertive clauses, representing [-given] information. We have shown that this distinction is reflected in the form and distribution of the clauses. Nonassertive clauses are introduced by thi, while assertive clauses are introduced by either wâa or the zero complementizer. Both types can be used as a basis for the subcategorization of complement-taking verbs. We have also shown that wâa and zero-complementizer clauses are not variant forms of one another, since they behave differently in relative clauses and existential sentences. Therefore, assertive clauses can be subdivided into two subtypes: the assertive wâa clause and the zero complementizer assertive clause.

3.3.3. The generic-nongeneric distinction

The assertive-nonassertive distinction has been posited for indicative embedded clauses. For hypothetical embedded clauses, a different type of distinction needs to be made. We propose that these can be divided into two: generic and nongeneric.

A generic hypothetical clause has a nonspecific referent, as in example (88) below:
A nongeneric hypothetical clause has a specific referent. Compare (88) above with (89) below:

(88)  ninthaa  ?aacaan  láplăŋ  bäap  ná?
gossip  teacher  behind-back  sinful  CM

'Gossipping about one's teachers behind their backs is sinful.'

(89)  nákrian  hūarunræŋ  yú?  phzan
student  radical  stir  friend
hây  phzan  ninthaa  ?aacaan  láplăŋ
COMP  friend  gossip  teacher  behind-back

'Radical students got their friends to gossip about the teachers.'

This generic-nongeneric distinction in hypothetical clauses can be compared to the generic-nongeneric distinction between a noun phrase like dèk in (90) and dèk  nìŋ  khon in (91):

(90)  phôm  hën  dèk  nay  hʂŋ  nán  dùay
I  see  child  in  room  that  too

'I saw children (or a child) in that room too.'

(91)  phôm  hën  dèk  nìŋ  khon  nay  hʂŋ  nán  dùay
I  see  child  one  CL  in  room  that  too

'I saw one child in that room too.'
This informational distinction is also evident in the structural difference between the two types of clause.

First, generic clauses are not introduced by a lexical complementizer; i.e., they have the zero complementizer:

(92) \text{yùu} \text{ chèyèchèyè} \text{ phôm thon} \text{ Ø} \text{ mâyðày}

\text{stay idle} \quad \text{I stand it cannot}

'Being idle . . . I can't stand it.'

But nongeneric clauses can be introduced by \text{hày}:

(93) \text{mèè} \text{ sôon lûuk \text{hày} pen khon dìi}

\text{mother teach child COMP be person good}

'Mothers teach their children to be good.'

Whether \text{hày} is a true complementizer or a verb is debatable. The fact that it can be negated as \text{mây \text{hày}}:

(94)a. \text{mèè} \text{ sôon lûuk \text{hày} pen khon dìi}

\text{mother teach child COMP be person good}

'The mother teaches her children to be good.'

b. \text{mèè} \text{ sôon lûuk \text{mày \text{hày} koohèk}}

\text{mother teach child NEG COMP lie}

'The mother teaches her children not to lie.'
suggests that it may be the verb of a hypothetical embedded clause embedded directly below the matrix sentence containing the verb *sön* 'teach' in (94) above. If so, this kind of nongeneric clause with *hây* is a special case of embedded clause which can occur as complement of only one verb, *hây* 'allow, let, give, assign, order'. Since it is not crucial to our discussion whether *hây* is a verb or a complementizer, we will assume here that it is a complementizer.

Nongeneric clause can also occur without a lexical complementizer. Two types occur. In one, the zero complementizer can alternate with *hây*, as in (95):

(95) phôm câaq dèk \{hây\} maa tât yâa

I hire boy COMP come cut grass

'I hired the boy to come and cut the grass.'

In the other, only the zero complementizer is allowed, as in (96) below:

(96) phôm, yàak \{\} \{\} pay khondiaw

I want I go alone

'I want to go by myself.'
Nongeneric clauses can be either *causative* or *noncausative*. Causative clauses may take either the *hây* complementizer or the *zero* complementizer. Such constructions require the subject of an embedded clause, represented in (97) below as $\emptyset_1$, to be coreferential with the object of the matrix verb:

(97) ?aacaan sàq nísît$_1$ hây $\emptyset_1$

   teacher tell student COMP they
   tham kaanbaan
   do homework

'The teacher told the students to do their homework.'

There is, however, another surface form of the same sentence in which the object of the matrix verb has a zero anaphoric form, which will be interpreted as coreferential with the subject of the embedded clause, as in (98) below:

(98) ?aacaan sàq $\emptyset_1$ hây nísît$_1$

   teacher tell them COMP student
   tham kaanbaan
   do homework

'The teacher told the students to do their homework.'
It is also possible for both the matrix object and the embedded subject to be present in the surface representation in lexical form, as in (99) below:

(99) ?aacaan sàq nisìtì hây nisìtì
teacher tell student COMP they
tham kaanbaan
do homework

'The teacher told the students to do their homework.'

It is possible also for the embedded nongeneric clause to have an independent subject which has no anaphoric relation with any noun phrase in the matrix sentence, as in (100) and (101) below:

(100) khâw bòok mìi hây raw deen
he wave hand COMP we walk
tòçpay continue

'He waved his hand for us to walk on.'

(101) phòò kamchàp mëi hây lûuklûuk pra?phít
Dad emphasize Mom COMP child behave
tua hây PROi pen khon dii
self COMP they be person good

'Dad emphasized to Mom (to see to it) that the children behave themselves as good people.'
As for nongeneric clauses with the zero complementizer (noncausative nongeneric clauses), the underlying subject is usually controlled by the subject of the matrix sentence; i.e., the subject of the matrix and the embedded clauses are coreferential:

(102) \textit{l̪u-klu-k} ᵃaak ə ə pay mi̯aŋnšok
children want COMP they go abroad

'The children want to go abroad.'

In contrast to both causative and noncausative generic clauses, as in (97)-(102) above, generic clauses must always have zero subjects and they are always interpreted as having a generic subject, as in (103) below:

(103) t̩eŋtua taam ʃeːchən tələtweedəa
dress according fashion all-the-time
pliŋ sataaq taay lesəy
waste money very certainly

'Dressing after the current fashion all the time must be exceedingly costly.'

To facilitate our discussion on two other structural differences between generic and nongeneric clauses, we present the following diagrammatic representation for each type and subtype of hypothetical clause:
HYPOTHETICAL

GENERIC

Subject: [+generic]

Surfase form of subject: 

{ *

form }
A third difference between generic and nongeneric hypothetical clauses lies in the possibility of generic hypothetical clauses being introduced by the head noun phrase kaan and khwaam,\(^7\) abstract nouns meaning 'action, deed, state, state of affairs'. Only generic hypothetical clauses can be embedded with such abstract head nouns, as in the following example:

\[(105)\]
\[
\begin{array}{cccc}
\text{a.} & \text{kaan} & \varnothing & \varnothing & \text{chùaylža} & \text{phûuʔžin} \\
\text{b.} & \varnothing & \text{*hây} & \varnothing & \\
\text{c.} & \text{hây} & \text{khon} & \text{khon} & \\
\text{d.} & \varnothing & \text{khon} & \\
\text{act} & \text{COMP} & \text{person} & \text{help} & \text{other} \\
\end{array}
\]

'The helping of others'

These abstract head noun phrases occur with one other type of embedded clause, the nonassertive thîi clause discussed in 3.3.2 above. Compare the two types in (106) and (107) below:

\[(106)\]
\[
\begin{array}{cccc}
\text{khwaam} & \text{thîi} & \text{kháw} & \text{pen} & \text{khonkhîinĩaw} \\
\text{state} & \text{COMP} & \text{he} & \text{be} & \text{miser} \\
\end{array}
\]

'His being a miser' or 'the fact that he is a miser'
One final structural difference between generic and nongeneric hypothetical clauses is their topic eligibility. Only generic clauses can occur as topics. Compare the causative nongeneric clause in (108) and the noncausative nongeneric in (109) with the generic clause in (110):

(108)a. khunphgreso khōo hāy khunmēs tham
   Dad ask COMP Mom make khanōmkhéek
   cake

   'Dad asked Mom to bake a cake.'

b. *hāy khunmēs tham khanōmkhéek
   COMP Mom make cake khanphgreso khōo
   Dad ask

(109)a. khunmēs yāak ∅ tham khanōmkhéek
   Mom want COMP make cake

   'Mom wants to bake a cake.'

b. *∅ tham khanōmkhéek khunmēs yāak
   COMP make cake Mom want
(110) a. ráthhabaan hàam ə sàq
government forbid COMP import
sǐnhhaa càak yìpun khàw pràthâet
goods from Japan enter country

'The government prohibits the importing of merchandise from Japan into the country.'

b. ə sàq sǐnhhaa yìpun khàw
PRO import goods Japan enter
pràthâet toon níi phìt kòtmǎay
country time this violate law

'The importing of Japanese merchandise into the country at this time is illegal.'

To sum up, generic and nongeneric hypothetical clauses differ in four structural respects. First, generic clauses are never introduced by lexical complementizers. Nongeneric clauses are of two types: causative and noncausative. The former can have the lexical complementizer hày, while the latter cannot have a lexical complementizer. Second, generic clauses have a zero subject in the surface form, one which has generic interpretation. Nongeneric noncausative clauses also have zero surface subjects, but these subjects are interpreted as coreferential with the subject noun phrase of the matrix sentence. As for causative nongeneric clauses, they take either a lexical or a zero anaphor as their subject.
The lexical form can be either an independent noun phrase or one coreferential with the object of the matrix object. When its surface form is a zero anaphor, it is always coreferential with the matrix object. The third difference is the possibility of being embedded in a noun phrase in which the head noun phrase is the abstract noun \textit{kaan} or \textit{khwaam}. Only generic clauses allow this. The final structural difference discussed is topic eligibility. Here, too, only generic clauses have this quality.

Clearly the topic eligibility of the generic clause can be explained in terms of its structural characteristics. Being hypothetical, it has no definite, independent time reference. Having a subject with generic interpretation, it does not predicate something of a specific individual; thus, it has a general, indefinite meaning. Its [+generic] informational status enables it to represent [+given] information and it is eligible for topichood without violating the SINPC.

3.4. Some Conclusions

We have shown that embedded clauses can be distinguished both informationally and structurally. The structural characteristics of each type of embedded clause correlate with its informational properties. Of all the types of embedded clauses, only nonassertive \textit{thi}i clauses and generic clauses are allowed in sentence-initial
position--i.e., topic as well as subject positions. This is predicted by the SINPC.

In the following diagram, we summarize the distinctions made with respect to the SINPC.
EMBEDDED CLAUSES

INDICATIVE

ASSERTIVE [-given]

wâa

-SINP

NONASSERTIVE [+given]

∅

-SINP

HYPOTHETICAL

GINERIC [+given]

thî

+SINP

NONGENERIC [-given]

∅

CAUSATIVE

+SINP

NONCAUSATIVE

-SINP

(+SINP: occurrence in sentence-initial positions)
1. It is possible to consider caa not as a marker for future tense but a clitic attached to a verb taking anticipatory hypothetical embedded clause. However, since the time referred to in the embedded clause is always interpreted as being subsequent to the time of the matrix sentence, we will represent it here as a future marker.

2. Kuno and Wongkhomthong (1980) interpreted the distinction between relative clauses introduced by thi and those without any lexical complementizer (thi-less in their terminology) in terms of the difference between the speaker's personal judgment, expressed by the thi relatives, and the general public's evaluation, expressed by the thi-less relatives.

3. In relative clauses, thi can alternate with san and ?an. Readers are referred to Ekniyom (1971) and Sornhiran (1978) for details of these complementizers. Interestingly, san is also found alternating with thi in complement clause in an archaic style of language.

4. Note that Needleman (1973) claims that thi ca is a single complementizer.
5. It would be reasonable to propose an alternative analysis in which \textit{w\^{a}}a is considered the verb of the clause embedded beneath the matrix verb and having the subject coreferential with the subject of the matrix verb. According to this hypothesis, what we consider a complement clause introduced by \textit{w\^{a}}a will be treated as a complement clause introduced by the zero complementizer embedded beneath one special verb \textit{w\^{a}}a.

6. This is in contrast to the \textit{thī} relative clause below:

\begin{verbatim}
1. khonı̄ thī ı̄ khəəy  gökhək
   person COMP he past broken-hearted
\end{verbatim}

'\textit{the person who has been broken-hearted}'

7. \textit{kaan} is used for the case when the noun phrase refers to physical actions or events while \textit{khwaam} refers to an abstract state of being. Compare the following pairs of noun phrases:

\begin{verbatim}
1a. kaan phuut koohək
    *khwaam phuut koohək
    act speak lie
\end{verbatim}

'lying'
2a. *kaan* mēettaa
act kind

'being kind'

b. *khwaam* mēettaa
state kind

'kindness'

8. This string is grammatical if *hay* is the verb meaning 'assign, allow, make possible'. The string then means 'the assignment of people to help others'.

9. This noun phrase can be modified further by a genitive phrase, as in the example below.

1. *khwaam* Ø *pen* khonkhùinśaw
state PRO be miser

khďq kháw
of he

'his being a miser'

Such modification is not possible for a phrase like (106), as can be seen in 2. below:
2. *khwaam thiŋ kháŋ pen khonkhininaw
   state COMP he be miser
   khɔŋŋ kháŋ
   of he
4.0. Introduction

In this chapter, we will examine the informational structure of a set of sentences in Thai which often have verb-initial surface forms. The data studied include two types of such sentences: presentative and existential sentences. We will argue that the verb-initial surface forms of these sentences are due to the fact that their subjects violate the SINPC and thus cannot remain in the regular subject position. We propose to account for the tendency in the language to disallow such subjects in sentence-initial position in terms of Extrapolation Transformations, which move an assertive embedded clause or an indefinite but specific noun phrase to the position following the predicate phrase.

4.1. Presentative Sentences

We define a presentative sentence as a construction in which there is a presentative verb followed by an assertive embedded clause introduced by \( \text{w\aa} \) or the zero complementizer. (112) is an example of a presentative sentence.
Unlike existential sentences, where only the assertive clauses introduced by zero complementizer are allowed, both the wâa and the zero-complementizer clauses are allowed in presentative sentences.

As for the presentative verbs, they can be divided into two types: regular presentative and evaluative presentative verbs. Semantically, the former introduce the embedded clauses as new information while the latter express an evaluation of a situation or event which may be expressed either as given or new information.

The following is a list of regular presentative verbs:

(113)³ baŋʔewen 'chance' kòertpen 'turn out'
duumʔan 'seem' klaaypen 'turn out'
kòet 'happen' praakot 'appear'
kòet...khîn 'happen' sôgsây 'likely'

The following is a list of evaluative presentative verbs:
These two types of verb differ not only semantically but also structurally. Regular presentative verbs occur only with wâa or zero complementizer clauses and are always found in the position preceding the embedded clause, as in (115) below:

(115)a. duumžan wâa khon khèp rôt khan
    seem    COMP person drive car CL

    nîi cà? pen khon phît
    this will be person guilty

'It seems that the driver of this car is the guilty party.'
b. *wâa khon khâp rôt khan nî
  COMP person drive car CL this
cà? pen khon phît duumžan
  will be person guilty seem

Evaluative presentative verbs, however, can occur with both assertive clauses and nonassertive clauses, as can be seen in (116):

(116)a. nâasįadaay thîi khun maa mâyďâay
  unfortunate COMP you come cannot

' That you cannot come is unfortunate.'

b. nâasįadaay wâa khâw maa mâyďâay
  unfortunate COMP he come cannot

' It is unfortunate that he cannot come.'

When an evaluative verb occurs with a thîi clause, it can either precede or follow the embedded clause, as in (117) below. Compare it to (116a) above.

(117)a. nâasįadaay thîi khâw maa mâyďâay
  unfortunate COMP he come cannot

' That he cannot come is unfortunate.'
(117)b. thįi kháw maa mâyďây nãasįadaay
COMP he come cannot unfortunate

'That he cannot come is unfortunate.'

It is apparent that the thįi clause is the subject of the evaluative presentative verb nãasįadaay in (117b). It is also the topic of the sentence. The difference between (117b) and (117a) is that the former has its topic in the regular sentence-initial position while the latter has its topic in the postposed position, as discussed in 2.6.

We will consider as presentatives only those sentences in which the evaluative verbs take assertive complement clauses, since these sentences always have the verb preceding the embedded clause.

4.1.1. Presentative verbs

We are in agreement with linguists like Rosenbaum (1967), Newmeyer (1969) and Gundel and Jacobs (1981) that presentative verbs are subject-embedding verbs. In fact, what we refer to as presentative verbs are not unlike the happen and likely- classes of English aspectual verbs in Newmeyer (1969). Newmeyer gives the following definition:

Aspectual verbs, essentially, are lexical items whose semantic role is to function as the predicate of a proposition rather
than to modify or refer specifically to
one item in that proposition. (3)
In deep structure they are all intransitive
and subject embedding. (5)

We decide to use the term *presentative verb* rather than *aspectual verb* as proposed by Newmeyer because the Thai equivalents of other classes of English aspectual verbs do not behave like aspectual verbs in Thai. These are such verbs as tāntēn 'begin', tōc 'continue', and lāek 'stop'.

The fact that presentative verbs in Thai are intransitive verbs which take embedded subjects is clear when we look at evaluative verbs taking thīi clauses. Since thīi clauses have the informational status [+given], they can appear in sentence-initial position as in (113b) above. But such thīi clauses can also be moved to the final position of the sentence, as in (113a). In this regard, evaluative verbs should be expected to have the same structural property, whether the embedded clauses they occur with are nonassertive thīi clauses as in (113) or assertive wāa clauses as in (116). Like evaluative presentative verbs, the regular presentative verbs should be expected to be similar structurally. Our claim that presentative verbs in Thai are subject-embedding, like their equivalents in English, is also supported by the
fact that when the embedded clauses contain a head noun phrase, this head noun phrase remains in the subject position. Compare the embedded clause without a head noun phrase in (118a) with the embedded clause with a head noun phrase in (118b) below:

(118)a. **duumšan wâa kháw cà? liîm raw ná?**
    
    seem COMP they will forget we CM

    'It seems that they may have forgotten us.'

    b. **rûupkaan duumšan wâa kháw cà?**
    
    circumstance seem COMP they will

    liîm raw ná?

    forget we CM

    'It seems that they may have forgotten us.'

4.1.2. Presentative subject and the SINPC

Why, then, is the embedded subject barred from occurring in its regular pre-predicate position?

As stated in (36), the SINPC requires that noun phrases which occur in any of the positions preceding the predicate phrase have [+given] informational status. Since **wâa** clauses and zero-complementizer clauses are embedded clauses of the assertive types--i.e., they represent new asserted information and they cannot remain in subject position.
We propose therefore a movement rule to be referred to as Assertive Clause Extraposition. This rule moves an assertive embedded clause out of the subject position to the position after the presentative verb. The following sentence, in which the predicate phrase contains not only a presentative verb but also other verbs and tense/aspect markers shows that the rule moves the assertive clause to the position after the verb rather than after the entire predicate phrase:

(119) kàet wâa khâw tokløŋ kan dâay
      happen COMP they agree together can

      khîn maa léw 1a?
      up come already CM

'It has happened that they have reached an agreement.'

Before we formulate the rule, we must first consider presentative sentences with a noun phrase in initial position. Thus, in (118b) above, we find the noun phrase rûupkaan 'situation' in initial position. Only a limited set of nouns occur as initial noun phrases in presentative sentences. These include:

(120) kaan 'matter' rûupkaan 'circumstance'
      hëetkaan 'event' sathâanâkaan 'situation'
      rɔaŋ 'story' thâathaŋ 'appearance'
In sentences with such a head noun phrase in initial position, what is extraposed is only the assertive clause, not the entire noun phrase:

(121)a. riŋ duumžan wâa kháw mòt yâyay
    story seem COMP he lose love
    kàp phûuyîŋ khon nîi ?aw ciŋciŋ
    for woman CL this take really

    'The story seems to be that he has really lost his love for this woman.'

b. *duumžan riŋ wâa kháw mòt yâyay
    seem story COMP he lose love
    kàp phûuyîŋ khon nîi ?aw ciŋciŋ
    for woman CL this take really

Notice that the head noun phrase left in initial position can have two possible semantic interpretations: a definite referent equivalent to riŋ nîi 'this matter' and an indefinite and nonspecific referent. When the head noun phrase really has a definite reference, it can also contain a demonstrative word, as in (122) below:
Another fact about presentative sentences needs to be taken into account. It is also possible for sentences to have the pronoun man occurring in the position preceding the presentative verb. This pronoun man 'it' occurs both in presentative sentences with a head noun phrase like (123) below:

(123) ṭǎːq nǐ man duumżan wâa khâw mòt
story this it seem COMP he
mòt yǎyay phûuyíg khon nǐi ṭaw
lose love woman CL this take
cińciń
really

'This seems to be that he has really lost his love for this woman.'
On the basis of all these facts about presentative sentences in Thai, we propose that they have the following underlying structure, one in which the embedded clause is generated in initial position as topic of the sentence. The fact that it is possible to have the pronoun man 'it' in the pre-predicate position shows that the bound variable of this topic is the subject of the sentence. Notice that the topic phrase is an embedded clause which may or may not have a head noun phrase.

(125)
The underlying structure for (124) will look like (126):

(126)a. man duumžan ə kháw cá? liím
       it seem COMP they will forget
       raw lësw ná?
       we already CM

'It seems that they may have forgotten us.'

b.  

To derive the surface form (126a), Assertive Clause Extraposition is applied to the underlying structure (126b) to move the S, consisting of the complementizer and the embedded clause, to the position after the verb, as shown in example (119). Assertive Clause Extraposition may be formulated as follows:
(127) **Assertive Clause Extraposition**

\[
\text{SD: } X \quad \overline{S} \quad \text{(PRO)} \quad V \quad Y \\
\text{[\text{-given}]}
\]

\[
\begin{array}{cccc}
1 & 2 & 3 & 4 \\
\end{array}
\]

\[
\text{SC: } 1,\emptyset,3+2,4
\]

If the topic clause has a head noun phrase, the head noun phrase remains in initial position. (128b) below is the underlying structure of (121a), repeated here as (128a):

(128)a. rāq duumžan wâa khâw môt

\begin{tabular}{l}
story & seem & COMP & he & lose \\
yəgay & kàp & phûuyĩq & khon & níi \\
love & for & woman & CL & this \\
\end{tabular}

\begin{tabular}{l}
?aw & cîncîq \\
take & really \\
\end{tabular}

'The story seems to be that he has really lost his love for this woman.'
When there is no noun phrase in underlying structure and the bound variable in the subject position also happens to be a zero anaphor, then the surface string will be verb-initial because the first lexical item in the string will be the presentative verb.

This clearly supports our proposed SINPC. The moved embedded clauses are assertive clauses, having [-given] informational status. If they remain in topic position, the sentence violates the SINPC and thus an ill-formed string results.

4.1.3. Presentative sentences with a distinct topic

Even in sentences in which there is a distinct topic like (129) below:
The topic is dèk khon níi 'this boy'. The subject is the assertive clause wâa k̥ːk̥ pùay 'that he is sick', containing the pronominal k̥ːk̥ 'he', which is the bound variable of the topic.
It is possible for subject clauses in presentative sentences like (130) to have a head noun phrase as well. (131) is an example of such sentences.

(131) phûulîiphay làw níi jí sathãanákaan
      refugees group this situation

  duumžan wâa râtthabaan wîatnaam cà?
  seem COMP government Vietnam may

  phlakdan hây \{kháw ñ\}_i \ ?õok maa ?eeq
  push COMP they leave come self

'These refugees--it seems that the Vietnamese government, itself, may have pushed them out.'

We have shown that the subjects of presentative sentences cannot remain in topic or initial subject position because this would violate the constraint on the informational status of noun phrases in initial position, the SINPC.

This phenomenon is not unique to Thai. Gundel and Jacobs (1981) discuss the same phenomenon for English and posit a constraint of similar nature for the informational property of the embedded clause in the subject position. This constraint requires that an embedded clause which remains in subject position must be the topic of the sentence. Those which cannot be topic are extraposed from the subject position.
4.2. Existential Sentences

4.2.1. Presentative and existential sentences

In spite of the fact that both presentative and existential sentences are often verb-initial in their surface forms, they differ in the following respects:

First, presentative verbs allow only embedded sentences as subjects. Both simplex noun phrase and embedded sentences are found in existential sentences. (132) is an example of a *simplex existential sentence*:

(132) mii sàetphông bon phôm khun nê?
exist dirt on hair you CM

'There is a piece of dirt on your hair.'

These usually occur with a locative phrase, as in (132) above or a temporal phrase, as in (133) below:

(133) mii kaanpàtì?wát sãam khrán nay
exist coup d'état three CL in
rìl pat sàp piì nìi
cycle ten year this

'There were three coups d'état in this decade.'

(134) is an example of a *clausal existential sentence*:
(134) mii khon sëɑg khon kamlan exist people two CL -ing thiąŋkən yùu argue -ing

'There are two people arguing.'

A second difference is that embedded clauses in presentative constructions can be introduced either by wåa or zero complementizer but existential sentences allow only zero complementizer clauses. Compare the presentative sentence in (135) below:

(135) duumţiŋ \{wåa\} khunphəo khəw sëa lësw {∅}

seem COMP father he die already

'It seems that his father has passed away.'

with the existential sentence like (136) below:

(136) mii \{*wåa\} khon kamlan chaay {∅}

exist COMP person -ing use
thoorasəp yùu
telephone -ing

'There is somebody using the telephone.'
A third difference is that embedded clauses in existential sentences do not have head noun phrases. Compare the presentative sentence in (137)

(137) ɹ̱an duum̃an khonduu čà? ďěn
story seem audience may walk
?ɔ̊ɔk mɔt
leave all

'It seems that the audience all walked out.'

with existential sentence in (138) below:

(138) {*[ɹ̱an] mii khonduu ľåy khon
{ g }
story exist audience many CL
ďěn ?ɔ̊ɔk
walk leave

'There were many in the audience that walked out.'

A fourth difference is that the pronoun man never occurs in existential sentences though it is allowed in all presentative sentences. Compare the presentative sentence in (139) below:
It seems that he will never forgive her.'

with an existential sentence in (140) below:

(140) *man mii phûuyî khon nèq maa

'There was a woman who came to ask for you this morning.'

4.2.2. Existential verbs

The class of existential verbs is rather small. It consists of two verbs: mii 'exist' and kàat 'come into existence, happen'. However, these verbs can occur with other verbs to form serial verbs such as mii khîn 'happen, occur' and kàat khîn 'happen, occur' as in (141) below:
The noun phrases occupy the position following the existential verb, not the entire predicate.

Not all noun phrases, however, fill this position. Compare the following pair of sentences in (142):

(142)a. mii sŏgkhraam khên lês\w
exist war up already

'A war has broken out.'

b. sŏgkhraam mii khên lês\w
war exist up already
(condaay naythísût)
inevitably finally

'The war has (inevitably) broken out.'
Note that the noun phrase sōŋkhraam 'war' has different readings in (142a) and (142b). In the (a) sentence, it refers to a particular war, of which the addressee may be unaware. In the (b) sentence, in which the noun phrase occupies initial position, sōŋkhraam 'war' refers to the war that both parties may be anticipating. The knowledge of the referent of this noun in (142b) is assumed to be shared by both the speaker and his addressee.

Unless this contrast of NP position is utilized, there is no way to derive the indefinite specific reading for sōŋkhraam 'war' in (142a) and the definite reading for (142b). It is also possible for noun phrases in topic position to have generic, i.e., nonspecific readings.

What this shows is that the noun phrase which occurs with mii is its subject. This subject can remain in pre-predicate position, a regular surface position for subject, if it is either [+definite] or [+generic]. When it is [-generic], it cannot occupy initial subject position in the surface representation. This is just what our proposed SINPC predicts. Only noun phrases representing [+given] information can occupy the subject or topic position in the surface form of a sentence.

We propose a transformational rule of Existential Subject Extraposition, approximately as in (143) below:
(143) **Existential Subject Extraposition:**

\[
\begin{array}{cccccc}
SD: & X & NP & Y & V & Z \\
\text{[-given]} & 1 & 2 & 2 & 4 & 5 \\
\text{[-exist]} & & & & & \\
SC: & 1,0,3,4+2,5 \\
\end{array}
\]

This rule applies only to the [-given] subjects of existential verbs. But it does not move all [-given] subjects of all verbs, as can be seen in (144):

(144)a. \*dèk n̂ng khon wîŋ m̄yûâay  
boy one CL run cannot

'There is a boy unable to run.

b. \*wîŋ m̄yûâay dèk n̂ng khon  
run cannot boy one CL

c. \*wîŋ dèk n̂ng khon m̄yûâay  
run child one CL cannot

For the sentence to be well-formed, it needs to be an existential sentence like (145b) below:

(145)a. \*dèk n̂ng khon wîŋ m̄yûâay  
boy one CL run cannot  

b. mii dèk n̂ng khon wîŋ m̄yûâay  
exist boy one CL run cannot

'There is a boy unable to run.'
The data presented above show that existential verbs in simplex existential sentences like (132) and (142) are intransitive verbs. The subjects are extraposed only when they violate the SINPC.

We maintain that they are also intransitive verbs taking sentential subjects when they occur in clausal existential sentences like (130) or (141b) above. It is possible for existential verbs to take both assertive and nonassertive clauses as their subject. A nonassertive clause, being [+given], can remain in subject position, as in (146a) below, or it can be postponed by the Topic Postposing Transformation, discussed in 2.6, to the sentence-final position, as in (146b):

(146)a. thî khon dœn khâw maa duu
    COMP people walk enter come look
    chêy chêy dooy mây sît ?aray
    only by not buy anything
    kâ? mî yûu mîankan
    also exist -ing too

' That people just walk in to take a look without buying anything does happen, too.'
When an assertive clause occurs in an existential sentence, it can occupy only the position after the existential verb in the surface representation, as in (147) below.

(147)a. *khon deen khaw maa kan
   people walk enter come together
   le:w mii
   already exist

b. mii khon deen khaw maa
   exist people walk enter come
   kan le:w
together already

'There are people coming in already.'

A comparison of sentences like (146) with those like (147) clearly shows that clausal existential sentences are also intransitive.
The rule of Existential Subject Extraposition proposed in (143) can be applied to derive the surface form (147b) from the underlying structure (147a). This rule, unlike the Assertive Clause Extraposition Transformation proposed for the derivation of presentative sentences, moves the entire noun phrase. This is possible because embedded clauses in existential sentences never have lexical complementizers, as shown in a comparison between presentative and existential sentences in (135) and (136), repeated here as (148) and (149) below:

(148) ɗuumţan \{\frac{wâa}{\emptyset}\} khunph9o khâw sìa léww

seem father he die already

'It seems that his father has passed away.'

(149) mii \{\frac{*wâa}{\emptyset}\} khon kamlq cháay

exist COMP people -ing use
thoorasâp yùu
telephone -ing

'There is somebody using the telephone.'

Another fact which makes it possible to formulate the rule in such a way that the entire noun phrase is moved is that embedded clauses in existential constructions
never occur with head noun phrases, as is clear if we compare the presentative and existential sentences (150) and (151):

(150) \{r\} duum\-yan khon\-duu ca\? de\-\-en
story seem audience may walk
\?\-\-ok mot
leave all

'It seems that the audience all walked out.'

(151) \{\} mii khon\-duu la\-\-ay khon
story exist audience many CL
de\-\-en \?\-\-ok
walk leave

'There are many in the audience that walked out.'

In our discussion of presentative sentences we saw that assertive clauses are always extraposed whether they occur in topic or subject position; likewise a [-given] noun phrase in existential sentences is always extraposed, whether it is in topic or subject position. This can be seen in a simplex existential sentence with a distinct topic like (152) below:
(152)a. *(nay) bān lăng nīi khon cèt
     in house CL this people 7
khon mīi
CL exist

b. (nay) bān lăng nīi mīi khon
     in house CL this exist people
cèt khon
7 CL

'In this house, there are seven people.'

as well as a clausal existential sentence like (153) below:

(153)a. *dèk khon ṇ̇ị mīi khon khṣey hën
     boy this people have see
∅₁ mīi
he exist

b. dèk khon ṇ̇ị mīi khon
     boy CL this exist people
khṣey hën ∅₁
have see he

'This boy, there are people who have seen him before.'
4.2.3. Underlying topiclessness

In this section we argue that the surface differences between presentative and existential sentences reflect significant differences in their underlying structures.

We have noted earlier sentences like (139) and (140) in which the pronoun man is found in the initial position of a presentative sentence but not in an existential sentence. We repeat the two sentences here to facilitate our discussion.

(154) \( \text{man} \) duumżan wâa khāw câ? māy
\[
\{ \emptyset \}
\]
it seem COMP he will NEG
hây?aphay thee dètkhàat
forgive she definitely

'It seem that he will never forgive her.'

(155) *\( \text{man} \) mii phuuyîŋ khon nîŋ maa
\[
\{ \emptyset \}
\]
it exist woman CL one come
thāam hāa khun mâacháaw
ask seek you morning

'There was a woman who came to ask for you this morning.'

We propose that this pronoun man 'it' in a presentative sentence is in fact a pronominal copy or bound variable of
the assertive noun phrase which is the subject of the presentative verb. We proposed earlier that a presentative sentence may have an underlying structure like (125), repeated here as (156):

\[
(156)
\]

\[
S \\
\leftarrow S' \\
\leftarrow NP_i \\
\leftarrow (NP) \\
\leftarrow S \\
\leftarrow COMP \\
\leftarrow \{\emptyset\} \\
\leftarrow \{\text{wâa}\} \\
\leftarrow S \\
\leftarrow NP_i \\
\leftarrow \{\emptyset\} \\
\leftarrow \{\text{man}\} \\
\leftarrow \text{PRED} \\
\leftarrow \text{[+present]}
\]

In this underlying structure, the topic is an assertive embedded clause with a bound variable which can have the surface form of either the pronoun man 'it' or a zero anaphor. The fact that a pronominal bound variable like man 'it' does not occur in an existential sentence is indicative of the fact that such sentences do not have underlying topics; i.e., they are topicless in underlying structure, as in (157) below:
The underlying topiclessness, as proposed in (157), can account for the fact that existential sentences, unlike presentative sentences, always have verb-initial surface form. Since there is no topic, there is no variable within the sentence which is bound by the topic. This explains why the pronominal copy man 'it' can be found in presentative sentences such as (154) but is never found in existential sentences, as can be seen in example (155). The underlying structures for both presentative and existential sentences as represented in (156) and (157) seem to provide an explanation for the fact that while the surface verb-initial form is optional for presentative sentences, it is obligatory for existential sentences.

As for sentences like (152) and (153), in which there are distinct topics, the underlying structure will
not be like the one in (157). Rather, their underlying structure will contain a topic noun phrase, as can be seen in (158a), which is an underlying structure for (153), repeated here as (158b):

(158)a.

\\begin{Tree}
  . S
    . NP_i
      . NP
        . S
          . COMP
            . S
              . AUX
                . V
                  . NP
                    . PRED
                      . mii

  . S'
    . NP
      . N
        . CL
          . DEM

  . dek
  . khon
  . nii
  . \emptyset

  . khon
  . kh\text{\`e}y
  . h\text{\`e}n
  . \emptyset

  . mii

\end{Tree}

'\text{This boy, there are people who have seen him before.}'

b. dek khon nii \ mii khon kh\text{\`e}y

\begin{itemize}
  \item boy
  \item CL
  \item this
  \item COMP
  \item people
  \item have
  \item see
  \item he
  \item exist
\end{itemize}
The rule of Existential Subject Extraposition will apply to move the subject embedded clause out of the subject position to avoid violating SINPC and will derive the surface form in (158b).

4.2.4. Functional use of existential sentences

The topiclessness of existential sentences allows them to be used to present information which is completely new—i.e., new information which is not a comment being made about any particular given topic. Compare a regular sentence like (159) below:

(159) khonkhây khon nán yaŋ roo yùú
patient CL that still wait -ing
khâ?
PART

'That patient is still waiting.'

with an existential construction (160) below:

(160) mii khonkhây khon nán yaŋ
exist patient CL that still
roo yùú khâ?
wait -ing PART

'There is that patient still waiting.'

Sentence (159) reports about khonkhây khon nán 'that patient' that he is still waiting, i.e., khonkhây khon nán
is the topic of the sentence. However, khonkhây khon nán is not what sentence (160) is about. The sentence simply supplies information that a particular patient is still waiting. Sentence (160) is an appropriate response for a nurse to give to a doctor who asks if there are any more patients for the day. It is, however, not an appropriate response to a question like khonkhây khon nán 1â? 'What about that patient?' Sentence (159) is an appropriate response to this question.

Another interesting use of the existential constructions is as "markers" for indefinite specific noun phrases in the initial position of sentences which are embedded as subjects. Compare a noun phrase with a quantifier in a sentence like (161) below:

(161) dêk sãam khon maa hâa khun
student 3 CL come see you
mîacháaw
morning

'The three students came to see you this morning.' OR 'Three of the students came to see you this morning.'

with the same phrase in the embedded clause of an existential sentence in (162) below:
When a quantified noun phrase occurs in topic or subject position, it is always interpreted as [+given]; i.e., it is taken to be either a definite noun phrase of a generic noun phrase, as can be seen in (161). The only time that such a phrase can be interpreted as indefinite but specific is when it occurs in initial position of an embedded clause after mii.

The verb mii also occurs in sentences in which the main verb is not an existential verb. The existential verb mii forms a verb series with the main verb of such sentences as (163)

(163) duumñan mii khon nìŋ khon kamlag
  seem exist person one CL -ing
deen maa thaaŋ raw
  walk come way we

'A person seems to be walking towards us.'
4.3. Some Conclusions

In this chapter, we have presented evidence from representative and existential sentences to show that subject noun phrases which represent new or \([-\text{given}\)] information cannot occupy initial topic or subject position, as predicted by our proposed SINPC. We have proposed two extraposition rules: Assertive Clause Extraposition and Existential Subject Extraposition. The former moves out of the subject position an assertive embedded clause, leaving behind the head noun phrase. The latter moves an entire noun phrase subject of an existential verb from the pre-predicate position to the position following the verb itself.

Unless there is a distinct topic with its bound variable within the embedded clause, an existential sentence is always verb-initial in its surface form. This means that it is topicless in underlying structure.
1. There are other sentences which may also have surface verb-initial forms. One is a sentence in which the subject, which also functions as topic, is not present in the surface structure but can be understood as being coreferential with another noun phrase in the same or the previous sentence, as in the (b) sentence of the following example:

   la. kwaanị pen sàt thị i lịaŋ lọuk
       deer be animal COMP feed young
     dụay nom
     with milk

   b. ọ i mii khwaamsuuŋ challa pràmaan
       it has height average about
     hàa fút
     five foot

   'Deer are mammals. They have an average height of about five feet.'

The reference of this understood topic or subject can be inferred from the linguistic context. Notice, however, that the (c) sentence in example (2) below has a verb-initial surface form and the reference of its Ø topic can be inferred from the physical presence of a child in the vicinity.
Another type of verb-initial surface form results from the postposing of the topic, which is also the subject phrase, to sentence-final position, as discussed in 1.6. The following is an example of such a sentence.

2. These terms are borrowed from Gundel and Jacobs (1981).

3. It is apparent that some of these verbs are compound verbs. For example, duu 'look' and mian 'look alike' form the compound verb duumian, which is equivalent to seem in English, and klaay 'turn' and pen 'be' form a compound klaaypen 'turn to be'.

The following is an example of such a sentence.

3. ?arɔy caŋ khà? khanőm khɔɔq
     delicious very PART cake of
  khusmêu
  Mom

'Your cake is really delicious, Mom.'
4. The situation is the same as in Note 3 above. These examples are not single verbs. Some of them are nominal predicates. For example, čhōokdiį literally means 'good luck', čhōok 'luck' and diį 'good'. Many of these have the prefix morpheme naa 'deserving of'. While sōncay means 'be interested in', naasōncay means 'interesting, deserving of one's interest'.

5. thāathaaŋ seems to be used both as a head noun phrase meaning 'appearance', as in (4) below:

4. thāathaaŋ nāklua wāa fōn cā? tōk
   appearance likely COMP rain will fall
   'It looks as if it might be raining.'

and a presentative verb as in (5) below:

5. đēk khon nīi thāathaaŋ keŋ cā?
   boy CL this appear he may
   keŋ
   intelligent

   'This boy appears to be intelligent.'

6. We consider mii in sentences like (6):
6. fi̋n man mii yùu wâa khâw
   story it exist -ing COMP she
   khâwcay phûm phît thâwnàn?eeŋ
   understand I wrong only

   'It is only that she misunderstood me.'

to be a presentative verb, since it takes a wâa
clause, allows a head noun phrase, and can have the
pronominal copy man. Therefore, mii can function
as both presentative and existential verbs.

7. Like mii in Note 6 above, khât can be either a
   presentative or existential verb. As a presentative
verb, it often means 'happen unexpectedly or contrary
to expectation'.

8. mii may seem to function as transitive verb, meaning
   'have, possess, contain', as in (7) below:

   7. phûm mii fĕeŋ lĕ̂eŋ
      I have friend already

   'I have got a girlfriend already.'

   It is possible, however, to consider mii in sentences
like (7) to be the same as the existential verb mii.
phûm functions as the "alien" topic of the sentences.
fĕeŋ functions as the subject and is moved to the
position after the verb because it violates the SINPC. That is, the sentence may be interpreted as 'Talking about me, a girlfriend has already been in existence.'

9. A verb series may involve either embedding or conjoining. This phenomenon merits more detailed analysis. We will assume that mii forms a serial verb with other verbs and reserve our judgment as to what process is involved in the formation of these serial verbs.
5.0. Introduction

In this chapter, we examine the initial noun phrases of a set of sentences in the language, to be referred to as focused constructions. These sentences are more or less the equivalents of cleft sentences in English. Constituting the comment rather than the topic of the sentence, these noun phrases appear to be counterevidence to our proposed SINPC, which requires that a noun phrase in initial position represent [+given] information. Note, however, that these initial noun phrases are always followed by focus markers, discussed in 2.5. We will argue that these noun phrases are a special type of comment. They are comments which need to be either [+definite] or [+generic], the usual requirement for topic phrases. As a special type of comment, these noun phrases support rather than contradict our proposed SINPC. The requirement that they represent [+given] information allows them to occupy the initial position without violating SINPC. It is the obligatory presence of focus markers in surface representation which prevents them from being interpreted as topics.
5.1. Surface Forms of Focused Constructions

A comparison of the surface form of a focused construction such as (164) below with its negative counterpart in (165) reveals that a focused construction must be a type of equational sentence.

(164) ผู้หญิง เกิดนี้ อย่างไร ที่ฉันช่วย
lady CL this FM COMP help
พยาบาล ดู
take care I

'It is this lady who helped take care of me.'

(165) ไม่ช่วย ผู้หญิง เกิดนี้ นี่บอก
NEG right lady CL this CM
ที่ช่วย พยาบาล ดู
COMP help take care I

'It is not this woman who helped take care of me.'

The focused sentence in (164) has no copula while its negative counterpart, (165), has ไม่ช่วย, the negative counterpart of the copula ช่วย. Native speakers react differently to the sentence in (166) in which the copula ช่วย occurs in the same position as ไม่ช่วย in (165):
Native speakers on whom this type of sentence was tested
did not consider this unacceptable but most preferred to
use sentence (164).

However, this shows that focused constructions are
equational sentences in which the predicate phrase precedes
the subject phrase in the surface structure. As can be
expected, sentences (164) and (165) can be paraphrased
respectively by (167) and (168) below:

(167) thîî chûay phayaabaan phôm khîî
COMP help take care I be
phûuyîŋ khon nîi qaylâ?
lady CL this FM

'(The person) who helped take care of me
is this lady.'
The Topic Postposing Transformation proposed in 1.6 can be applied to (167) and (168) to derive (164) and (165) respectively. As for the absence of the copula khiit in (164), a deletion rule of Initial Identificational Copula can be proposed. If we adopt the model in Chomsky and Lasnik (1977), this will not be a rule in the transformational component but one of the deletion rules that apply to the surface structure, which is the output of the transformational component. This deletion rule may be obligatory for some speakers but optional for others. After it has been applied, the surface form will have the nominal predicate in initial position, as in (164).

5.2. Focused Noun Phrases

In 5.1, we showed that the initial noun phrase of a focused construction like (164) is a nominal predicate which occupies surface initial position because the topic or subject phrase has been postposed and the copula
has been deleted. In this section we will examine the informational properties of this type of noun phrase.

A study of focused constructions reveals that not all noun phrases can occur in initial position. Only noun phrases which are either [+definite] or [+generic] are allowed. Compare the three sentences in the following example.

(169)a. đèk khon níi gaylâ? thîi kháw
boy CL this FM COMP he
phaa maa
bring come

'It is this boy that he brought with him.'

b. kaafe: gaylâ? thîi cà? tham hay
coffee FM COMP will make COMP
khun máy ʔaqu ʔâa kɔon
you NEG sleepy lose before

'It is coffee which can help you from being sleepy before it is time.'

c. *ʔaaʔan nə̀ caan gaylâ? thîi
food one plate FM COMP
kháw khɔo
he ask
These noun phrases must have a definite referent capable of being identified by the speaker. This is a requisite for a noun phrase occurring with the copula khít. As discussed in Gething (1979), khít is a "definite" copula. It is identificational in its semantic function. This differentiates it from the other copula in the language, pen, which is descriptive in its semantic function. Compare khít and pen in the following examples.

(170) phûuchaay khon níí \{ pen \} mño
       \{ *khít \}
man  CL  this  be  doctor

'This man is a doctor.'

(171) khunmño thîi raksaa khunphño \{ khít \}
       \{ *pen \}
doctor  COMP  treat  father  be
phûuyî kwon nán ñay
lady  CL  that  FM

'The doctor who is treating Dad is that lady.'

It is possible for the Topic Postposing Transformation to apply to (170) to derive a sentence like (172):

(172) pen mño (//) phûuchaay khon níí¹
be  doctor  man  CL  this

'He is doctor, this man.'
and to (171) to derive (173):

(173) \text{khī} \text{ phūuyīq} \text{ khōn nān ṇay khunmō} \\
\text{be} \text{ lady CL that FM doctor} \\
\text{ṭīi ráksāa khunphsō} \\
\text{COMP treat father} \\

'It is that lady, the doctor who is treating Dad.'

After the application of Initial Identificational Copula Deletion, (173) will have the surface form of a focused construction, as in (174):

(174) phuuyīq \text{ khōn nīī ṇay khunmō} \text{ thīi} \\
\text{lady CL this FM doctor COMP} \\
\text{ráksāa khunphsō} \\
\text{treat father} \\

'It is this lady, the doctor who is treating Dad.'

The noun phrase in the initial position of a focused construction corresponds to the noun phrase following the identificational copula \text{khī} in the corresponding non-focused construction as in (173). It always has a definite referent, in the sense of Chafe (1976); i.e., it has to have [+given] informational status. Therefore, it constitutes special type of comment, a comment which contains [+given] information.
Chafe (1976) considered this type of noun phrase to be contrastive. The speaker singles out a particular object or entities out of a number of candidates which are all assumed to be known or familiar to the addressee.

When the Initial Identificational Copula Deletion Rule has been applied to delete *khii*, this contrastive noun phrase is left as the initial noun phrase of the sentence. Due to its status of being [+contrastive], and thus also [+given], it can occupy this initial position without violating the Proposed SINPC even with its informational function as comment.

Therefore, rather than being counterevidence to our proposed SINPC, the initial noun phrase of a focused construction provides additional support.

5.3. Focus Markers

In 5.2 we have shown that the initial noun phrase of a focused construction conforms to the SINPC in that it represents [+given] information. Being a [+given] noun phrase occupying the initial position of a sentence, this initial noun phrase qualifies to be interpreted as topic of the sentence. However, this interpretation is blocked by the presence of the focus marker after these initial [+given] noun phrase.

Without these focus markers, the surface string will be only a series of one noun phrase followed by
another,\(^1\) as in (175b). Compare it to (174) above repeated here as (175a):

\[
(175)\text{a. } \text{phûuyûŋ} \text{ k hon níi } \text{pay} \text{ khunmño thîi} \\
\text{lady CL this FM doctor COMP} \\
ráksåa k hunph‰c \\
treat father
\]

'It is this lady, the doctor who is treating Dad.'

\[
(175)\text{b. } ?\text{phûuyûŋ} \text{ k hon níi khunmño thîi} \\
\text{lady CL this doctor COMP} \\
ráksåa k hunph‰c \\
treat father
\]

The presence of a focus marker is absolutely necessary when the topic of a focused sentence happens also to be a headless relative clause\(^2\) as in (176a). This is because a focused noun phrase not marked by a focus marker will be interpreted as the head noun phrase of the relative clause that follows it, as in (176b).

\[
(176)\text{a. } \text{phûuyûŋ} \text{ k hon níi } \text{paylâ? thîi} \\
\text{lady CL this FM COMP} \\
chûay phayaabaan phôm \\
help take care I
\]

'It is this lady who helped take care of me.'
There are three fairly common focus markers: ɲay or ɲaylâ? and tàaŋ hàak. The first two are interchangeable, as can be seen in (177) below:

(177) mìit lèm níi {ɲay} thîi bèat
      {ɲayla?}
kîfie CL this FM COMP cut
mìì nₚu
hand I

'It is this knife which cut my hand.'

tàaŋ hàak has added meaning, equivalent to 'instead' in English. It is a contradiction or a correction to an idea the speaker assumes his addressee has. (178) is an example illustrating the use of this focus marker.

(178) sìa tua níi tàaŋ hàak thîi pen
dress CL this instead COMP be
khɔɔŋ nₚu
of you

'It is this dress instead which belongs to you.'
These focus markers are found not only with nominal complements of *khiː* but with other types of comment as well, as can be seen in (179):

\[
\text{jen kh정책 khun ph柬埔 khiːn hây}
\]

money of you I return to

khun lɛːw ɲaylâ?
you already FM

'Your money, I have returned it to you already.'

5.4. Some Conclusions

We have shown that the initial noun phrase of a focused construction is a special type of comment. It is required to represent [+given] information since it is the complement of the copula *khiː*. Its occurrence in initial position provides added evidence in support of SINPC, since it has [+given] informational status. Its obligatory focus marker is what prevents it from being interpreted as topic.
NOTES TO CHAPTER V

1. It is possible that with the help of distinct sentence intonation and internal pause, (175b) can be interpreted as a focused sentence. However, a vestige of the focus marker is most of the time clearly detectable.

2. The topic of a focused construction can but need not be a headless relative clause. Here is where focused constructions in Thai differ from so-called cleft sentences in English. The following are some examples of focused constructions with different types of topics in Thai. (1) has a headless relative:

1. น้ำสีล้มนี่ไถ่ที่ผมเขียน
   book CL this FM COMP I write
   'It is this book that I wrote.'

(2) has a relative clause with a head noun phrase:

2. นี่ไถ่น้ำสีที่ผมเขียน
   this FM book COMP I write
   'THIS is the book that I wrote.'

(3) has a regular noun phrase as the topic of the sentence:
3. นี้ แย่ น้ำสีา ข้อง คุณ

this FM book of you

'THIS is your book.'
Our investigation of sentences in Thai has shown that there is a clear distinction between the constituents functioning as topic and those functioning as comment of a sentence. Our data have revealed that topics usually occupy sentence-initial position, i.e., any position preceding the predicate phrase. Forms having the status of noun phrase occupying this position are required to represent given information. We have proposed to account for this requirement in terms of a constraint on informational status of sentence-initial noun phrase, referred to as the SINPC.

We have investigated the interaction of the SINPC and embedded clauses. Our study shows that only the embedded clauses which have the [+given] informational status can occur in either the topic or subject positions. These are the nonassertive thii clauses and the subjectless hypothetical clauses. The former have informational status equivalent to a definite noun phrase. The latter have informational status equivalent to that of a generic noun phrase.

Our study has also revealed that the verb-initial surface forms of presentative and existential sentences can be accounted for in terms of how their subjects,
representing new information, violate SINPC and have to be extraposed to the position after the verb. There are many differences in the structural properties of the surface forms of presentative and existential sentences. We argue that these surface differences are reflections of the differences in the underlying structure of these two types of sentence. We propose that presentative sentences have an underlying structure with a [-given] topic while existential sentences do not have any topic at all. This seems to account for the fact that existential sentences have the function of introducing completely new information which is not about any given or known topic.

We have also studied a special set of sentences in which the initial noun phrases are a part of the comment of a sentence instead of the topic. We have found that even these comment noun phrases conform to the SINPC. They have to be [+given] noun phrases. What prevents them from being interpreted as topic is their obligatory focus markers.

It is apparent that there are direct correlations between the syntactic forms and the informational functions and the informational status of these forms in Thai. However, this is hardly unique to Thai. Hope (1973) discusses the effect of informational factors on the positioning of noun phrases in Lisu sentences. Contreras (1976) presents a similar study of Spanish. What this type
of study implies is that the inclusion of nonsyntactic factors in studies of syntactic phenomena can enable us to gain insights in syntax itself. After all, linguistic forms are available in a language to serve certain well-specified functions in the communication of messages between speakers and addressees. However, there is no denying that the correlations between forms and informational functions vary in degree across languages. Thai, as our study has revealed, seems to exhibit a rather direct correlation between certain linguistic forms and their informational functions.
BIBLIOGRAPHY


