A MODEL OF A DISCOURSE GRAMMAR
FOR THE ANALYSIS OF THAI

Peansiri E. Vongvibanond
Chulalongkorn University
Bangkok, Thailand

1. INTRODUCTION.

It is obvious that discourse is as natural a linguistic unit as sentence, phrase, or words. However, discourse had not been the focus of linguistic study until the 1960's, when Pike (1964) and Gleason (1968) proposed their work in discourse analysis. Discourse is defined as an extension of sentence, a construction at a higher level than a sentence in the structural hierarchy of linguistic units. It is clear to those working with discourse that though there are discernible patterns of regularity, these are not readily reducible into explicit rules by which speakers of the language have to abide as they construct their oral or written discourse. As a consequence, most linguists working on discourse, such as Dijk (1972 and 1973), Grimes (1975), Halliday and Hasan (1976), Longacre (1983), and Werth (1984) have come to content themselves with the definition of discourse as a semantic unit with a semantic structure, which is overtly manifested as a cohesive composite of sentences. It is not surprising then that a great deal of attention has been focused on the unifying relations of sentences which make up a discourse, variously referred to as cohesion, coherence or connectivity. However, attempts have been continuously made to formulate a discourse grammar or a text grammar, which is categorically different from a sentence grammar. This paper is another attempt in this direction. The model of discourse grammar to be outlined here has evolved from a practical framework for a systematic analysis and an evaluation of expository discourses in Thai. The grammar, however, is not exclusive to any particular types of discourse.

Two terms, "discourse" and "text", have been used by different groups of linguists working with linguistic units larger than sentences, who happen to be in geographically as well as culturally different regions of the world. "Discourse" is used by American linguists such as Pike (1964), Gleason (1968), Grimes (1975) and Longacre (1983), whose work concentrates on
oral narratives in lesser known languages without a writing tradition. The term is also used by British language educators, such as in Sinclair and Coulthard (1975) and Coulthard (1977), who focus on the analysis of oral discourses in classroom. The term "text" is used by European linguists, who do not usually recognize the distinction between grammar and language performance and prefer to look at language in social context. Halliday (1967 and 1968), van Dijk (1972 and 1973) and de Beaugrande (1985) are among this group. Their text grammars usually account for both the text and the context in which the text itself is situated. The author shares the same assumption about linguistic study as that of the European text linguists; however, the term "discourse" will be used in this paper, since it implies the communicative aspect rather than the physical aspect of this linguistic unit.

2. FILTERS AND CHOICES IN GRAMMAR

Since the proposal of the generative theory of language by Chomsky (1957,1965), grammar has come to be defined for many linguists as an enumeration of sets of rules which make explicit the linguistic competence of an "ideal speaker-hearer", which enables him to determine the grammaticality of an utterance. An utterance is taken to be equivalent to a sentence. The rules proposed in the grammar serve as "filters" to block ungrammatical sentences from being generated. This type of filter grammar cannot be easily adopted for discourse for two reasons. First, a discourse is not an absolute, autonomous syntactic construction like a sentence. It is possible to prescribe constituency of a sentence as consisting, at a minimum, of a subject noun phrase and a verb phrase which predicates it. It is also possible to determine the beginning and the ending of a sentence though theoretically a sentence can be of a definite length. This certainly is not the case with a discourse. Secondly, grammaticality and ungrammaticality are not attributes of a discourse. A discourse, spoken or written, can be well composed or badly composed. The judgement is usually subjective and it is not the concern of ordinary native speakers. They can manage to obtain the meaning of even a discourse which is considered badly composed. Only professional groups such as teachers, editors and psychiatrists concern themselves with quality of discourse. As a consequence, it is not possible to write a grammar of filters for discourse. Instead, one finds grammars of frequency of occurrence like those proposed by Grimes (1975) and Longacre (1983) or cognitive grammars of
text production and comprehension like those proposed by de Beaugrande and Dressler (1981) and van Dijk and Kintsch (1983) or a functional grammar like the one proposed by Halliday (1967 and 1968), which describes the various structures in language as manifestations of systems of meaning postulates, the choice of which is made by speakers.

The theoretical framework which underlies a discourse grammar proposed in this paper is the same as Halliday's systemic grammar. As summed up in Halliday (1976), a grammar is a network of systems bridging the meaning, which is determined by the choices made within the systems, and the syntactic and phonological manifestation of the meaning. Three systems have been proposed by Halliday: the Mood and Modulation System, the Transitivity System and the Thematization System which respectively represent three major socio-cultural functions: the Interpersonal Function, the Ideational Function and the Textual Function.

The difference between Halliday's systemic grammar and the grammar to be proposed in this paper lies in the focus and the scope of the grammars. Though Halliday recognizes the effect of textuality on sentence structure and accounts for it as the thematicization system and he co-authors with Ruqaiya Hasan (1976) an extensive analysis of cohesive devices in English, his grammar is sentence-oriented. He comprehensively illustrates the surface manifestation of the network of these three systems in his admirable description of the English language (Kress 1976: 101-233).

The model of grammar to be proposed in this paper is a discourse grammar, which is part of an entire language system, consisting of a discourse grammar and a clause grammar. Discourse grammar accounts for the choices to be made in the organization of information to be communicated in a unified linear sequence known as discourse. Clauses are constituents of discourse. The organization of information at the clausal level is different from that at the discourse level (Ekonyo 1982) and clause manifestations are subject to syntactic, morphological and phonological constraints or filters.

The grammar is formulated on the basis of an analysis of a discourse corpus, which originally consisted of 26 academic papers in humanities. Ten papers, five in social science and five in science and technology were later added. The grammar was also used as a framework for an evaluative analysis of two sets of compositions by senior high school students. One set consists of 22 compositions on "Teenagers and Love". The other consists of 22 compositions
"Biology".

3. INFORMATION STRUCTURING

The concept of information structuring which underlies this model of grammar is an adaptation of the Prague School linguists' notion of Functional Sentence Perspective or FSP (Mathesius 1961, Firbas 1964), which can be defined as the structuring of information within a clause. The concept has been used by the author (Ekniyom 1982) in a study of how the structuring of information determines the surface structure of sentences in Thai. Implicit in the FSP and information structuring concepts is the fact that information is communicated orally or in writing in a linear sequence due to the physical condition of human speech production and perception mechanism. Speech and writing are both linear sequencing of information. A sentence, or rather a clause, is a representation of an information unit. In each unit, there are a number of blocks of information, represented by words and morphemes, or groups of words and morphemes, in the language. To facilitate comprehension, information within a clause is assigned different information functions. Information which serves as "point of departure" has the function of being the topic of the sentence. Information which is being imparted about the topic has the function of being the comment of the sentence. In Thai, these information functions determine the sequential positioning of sentence constituents. There are also morphological devices for marking emphatic topic and comment.

The author has extended the same principle of information structuring for the analysis of discourse. Discourse is defined as a linear sequence of a set of semantically related information units. It is conceivable to have a discourse which contains only one information unit, and thus only one single independent clause. However, this one-clause discourse is most likely part of a larger discourse. An example is the ritual discourse at a wedding. A bride and a groom may deliver only one line, "Yes, I do" as a response to the long description of marital responsibilities cited by the officiating clergyman. Their one-line discourse is actually embedded in the entire discourse prescribed for the occasion.

Therefore it is not inconceivable to claim that people usually talk or write in discourse and not sentence. With a thought or a developing succession of thoughts in his mind, a person is forced to structure the information he wishes to convey in such a way that
it can be transmitted in a linear sequence of human speech. There may be a difference in how a person structures information in an unplanned and a planned discourse. In an unplanned discourse, information structuring and information transmission occur almost simultaneously so the structuring strategy may be localized among a very small number of pieces of information. It is not possible to revise and change the strategy unless the person is willing to repeat or rephrase the same message. In a planned discourse, there can be an interval, the length of which depends on various human and situational factors, between the structuring process and the transmission of the discourse. Revision and change are possible so structuring can be both localized and global. Public speeches and academic papers are examples of planned discourses, the preparation of which takes time for most people.

It is, therefore, imperative that information structuring be one of the systems of a discourse grammar in order to account for the choices a speaker can make as he tries to felicitously structure the information he wishes to convey into a discourse.

4. PROPOSED DISCOURSE GRAMMAR

As mentioned earlier, the discourse grammar to be outlined here has evolved from an effort to formulate a framework for analysing and describing discourse data in Thai. A study of these discourse data reveals that at least four systems are needed in the grammar to account for linguistic phenomena at the discourse level: the INFORMATION STRUCTURING SYSTEM system, the COHERENCE SYSTEM system, the ATTITUINAL system and the ILLCUTIONARY system. The information structuring system specifies the level of significance of each information unit. It serves the RANKING FUNCTION. An information unit is assigned a rank based on the degree of significance of the message intended by the speaker. The coherence system serves the COHERING FUNCTION by specifying to the listener how information units are related as constituents of a discourse. The attitudinal system serves the PERSPECTIVE FUNCTION of making the listeners aware of the speaker's opinion and attitude on what he is talking about. The illocutionary system serves the INTERACTIONARY FUNCTION and reflects how the speaker interacts with the listener as he delivers his discourse.

4.1 Information structuring system

Physically, a discourse can be very simply defined
as a group of sentences. Within a sentence, there can be one or more clauses. These clauses can be independent, dependent or downgraded into a phrase. An independent clause can occur by itself in a simple sentence or it can co-occur with one or more independent clauses in a compound sentence. The union is signalled with a conjunction. A dependent clause needs to be embedded as a constituent of another clause. The product of this embedding process is a complex sentence. When a clause is downgraded, it assumes a new syntactic status, that of a phrase, and it can occur as a phrasal constituent of another clause.

The information structuring system is motivated by the fact that it is possible for the speaker to make a choice of how to package an information unit, as an independent clause, a dependent clause or a downgraded clause. His choice of clause type reflects the level of significance, or the information rank, he intends for that particular information unit in the discourse. An independent clause marks the high rank, a dependent clause the middle rank and a downgraded clause the low rank. The underlined part in each of the following sentences illustrates this ranking difference of the same information unit.

(1) khaw nanlon. khaw tatsincai ca? saraphap.
   He sat down. He decided to confess.

(2) khaw nanlon khanathii khaw tatsincai ca? saraphap
   He sat down as he decided to confess.

(3) khaw nanlon tatsincai ca? saraphap
   He sat down, having decided to confess.

In Thai, an independent clause is changed into a dependent clause through addition of complementizers, which are in fact embedding markers. These are waa thii ?an sun phuu. However, these embedding markers can be omitted in cases where the embedded status of the clause is already evident in the surface syntactic structure. One such case is when the embedded clause functions as an adverbial clause. The presence of a subordinating conjunction adequately marks the rank shift, as seen in (2). Another case is when a dependent clause is a modifier of a head noun phrase (Sornhiran 1978). Downgrading is marked by the stripping off of the subject and the tense and aspect auxiliaries, which changes the status of the verb from finite to non-finite. Identification of downgraded clause can be problematic due to syntactic and morphological characteristics of the language. Being a topic prominent language (Li and Thompson 1976, Ekniyom
1982), the Thai language is rich in topicless sentences. Subjectless sentences due to the use of zero pronoun (Surintramont 1978) and deletion of subject (Bhandhumedha 1976) are also prevalent. Very often one cannot be too certain if it is a subjectless or a downgraded clause because of the absence of inflection to differentiate between finite and non-finite verb phrases. Sentence (3) above is an example.

Cognitively speaking, one can say that the speaker packages his thoughts into information units, assigns to each unit an information rank of his choice and reflects his choice accordingly through the various clausal types available.

4.2 Coherence system

Coherence distinguishes a discourse from a mere string of sentences. It is not surprising that this is an aspect of discourse which has been studied most extensively. The study of cohesion in English by Halliday and Hasan (1976) provides a practical framework for the study of cohesion in other languages. Chanawangsya (1986) uses this framework in her study of cohesion in Thai. Werth (1984) proposes a theory of discourse in terms of an information processing programme, which is based on connectivity of information contained in a discourse constituted by reference chain, logical connectors and collocation. Terminological discrepancies among linguists cannot conceal their consensus that coherence, variously referred to elsewhere as cohesion or connectivity, is a fundamental property of a discourse.

The coherence system in the model of discourse grammar proposed here consists of two sub-systems: the ANAPHORIC sub-system and the COHESIVE sub-system. The former is a natural exploitation of "shared" or "known" information. Units of information are naturally related because they talk about the same topic or topics. The latter is the speaker's deliberate specification of the semantic connectivity of information units in a discourse as he sees it.

4.2.1 Anaphoric sub-system

Once information is presented to the listener, the speaker can assume that it is known or shared by both parties and reference can be made to it in an abbreviatory manner through anaphoric means available in the language. As a discourse develops, the amount of known information increases. Not all information is simultaneously called to attention or "activated" at a certain
point in the transmission of a discourse. Some information remains "unactivated" until it becomes relevant again in the course of discourse, when it will be activated. The anaphoric sub-system, therefore, has to allow two alternatives for representing known information: non-activating and activating.

In Thai, non-activated information is covertly represented as ellipsis or zero anaphora. This is commonly referred to by Thai grammarians as "left out as understood." It must be pointed out that zero anaphora represent not only non-activated information in a discourse but also non-activated information which is part of real world knowledge (Surintramont 1978). What distinguishes these two instances of zero pronoun is the presence of antecedent of the former within the same discourse.

Activated information is overtly represented as repetition, pronouns and semantically related lexical items.

Repetition can be whole or partial. It can also be made specific and contrastive with the addition of demonstrative words nii and nan.

Pronouns representing known information fall into 2 categories: demonstrative pronouns nii and nan and personal pronouns, of which the choice in the rich inventory of the language is determined by complex sociolinguistic conditions such as status, degree of intimacy, age, sex, and the speaker's attitude (Pala-kornkul 1972). It is not surprising that personal pronouns are used less frequently than other types of anaphoric items. Abstract topics such as those found in the corpus of expository discourses in this study are never represented as man, an equivalent of "it". When activated, they are simply repeated in whole or in part or by other semantically related lexical items as observed in Praphaphorn (1985).

Semantically related lexical items range from synonyms, antonyms, hyponyms and superordinates, meronyms or collocatives, and also words with generic meaning such as khon "person", sin "object", kaankra-tham "deed, action" karanii "case", etc.

Activation of known information constitutes DISCOURSE TOPIC CHAINS. Discourse topics can be defined as "what the discourse is all about." The discourse itself furnishes new information to the listener about these topics. There can be one or many topics per discourse. The choice is for the speaker to make. Unlike a sentence topic, discourse topic may occur either in the topic or the comment part of a clause contained within a discourse. Discourse topic chain emerges when a piece of information is introduced and
referred to again in the discourse.

As an overt cohesive device, discourse topic chains reveals a hierarchical internal structure of information in an entire discourse. A cluster of clauses on the same activated topics can be found embedded in, comjoined with and juxtaposed to another cluster (Vongvivanond 1991).

4.2.2 Cohesive sub-system

Cohesive sub-system accounts for the speaker's choice to specify the semantic relations which hold among information units within a discourse. There is a major difference between the anaphoric system and the cohesive system. Being a natural way of transmitting and processing information, anaphoric reference is used automatically so the system does not have a null option. In the cohesive sub-system, a null option is included to account for the case when the speaker may leave the relations among information units within his discourse unspecified, at the risk of being incoherent. This is observed in very young children's talk. As their language acquisition process develops, children exhibit the gradual expansion of their cohesive devices inventory.

Cohesion has been the focus of interest of many linguists who are also working on language teaching. It is perhaps one of the most visible criteria which differentiate well composed and poorly composed writings and speeches. A comprehensive study of cohesive devices in the English language is found in Gleason (1965), Quirk et al. (1972) and Halliday and Hasan (1976). Mahatdhanasini (1982) and Chanawangsa (1986) did similar studies for the Thai language.

The study of cohesion in the past was usually sentence oriented. Cohesion was usually described as relations between sentences or among group of sentences. However, when cohesion is placed in a discourse perspective, a different pictures of cohesion emerges. There are two types of cohesion: macro cohesion or discourse cohesion, which organizes and arranges information units in a discourse, and micro cohesion or inter-sentential cohesion within a group or a pair of sentences. The cohesive subsystem is therefore, divided into two mini-systems: macro-cohesion and micro-cohesion.

a. Macro-cohesion

Macro-cohesion operates globally in an entire discourse. It tells of the boundary, the scope, and the arrangement of information within a discourse. It is the choice the speaker can make to help his listeners
keep track of the flow of information and facilitates the processing of information.

Macro-cohesion is manifested in terms of labels, phrases and clauses. "Label" is a cover term for titles, headings and subheadings. Labels mark the beginning of a discourse. The use of labels is typical of the planned speech style in expository writing. In place of labels, the speaker may devote a clause or a sentence for this purpose. This is typical of monologue oral discourses such as tales and comic stories.

The ending of a discourse is not marked as clearly as its beginning. The best illustration of this macro-cohesive device is the last sentence of a Buddhist monk's sermons and prayers.

\[
\text{ewaN ko mii duay prakaara chani} \\
\text{"Thus concludes this sermon."}
\]

Usually the speaker would end his discourse with one or more clauses containing expressions with the verb sarup "to conclude" or the verb cop "to end."

The scope of information, i.e. the main discourse topic, of the discourse can be overtly marked also by labels. They help provide a frame of reference for the listener's information processing. The heading tha-nakhaant "bank" enables the listener to interpret the meaning of sakhaa "branch" correctly and it alerts him that the discourse may have digressed or moved on to a different topic when he is receiving information which his real world knowledge tells him has nothing to do with banks or banking.

Macro-cohesion also includes arrangement of information. This is manifested as expressions which indicate the positioning of information in an organized series. Examples of these expressions are ordinal numbers, which may occur alone or as modifiers of words such as pra?kaan khoo ?andap?, all meaning "item." The series is usually introduced with a clause containing the expression dantoopainii "as follows".

b. micro-cohesion

Micro-cohesion operates locally between sentences. It is realized as connectors. There are two types of connectors: conjunctions and conjunctive adverbs.

In Thai, both conjunctions and conjunctive adverbs manifest their connecting function very visibly. They strategically occur at the connecting point, i.e. the initial position of the sentence which follows, with the only exceptions of the pre-predicate connectors koo and cun. These two conjunctive adverbs are multi-purpose connectors (Vongvikanond 1986). Their semantic interpretation varies with the context. In other
words, they simply shows the speaker's choice to emphasize that there is a semantic relation.

Micro-cohesion is classified into seven types: addition, contradiction, alternative, cause and effect, exemplification, restatement, and sequence. The following are examples of connectors representing micro-cohesion in Thai.

**Addition**
- nookcaaknii  "besides this"
- naithamnooNdiaawan  "in the same manner"
- yiNpaikwaanan  "more than that"

**Contradiction**
- mxxkrana  "even so"
- yaaNraikotaam  "no matter what"
- haaktxxwaa  "except that"

**Alternatives**
- mi?chanaan  "if not so"
- ryymaiyaANnan  "or if it is not that"
- ryy?iikyaANnyN  "or in another case"

**Cause and effect**
- phro?chanaan  "because of that"
- chandaichannan  "so it is"
- dooinaiyanii  "in this way"

**Exemplification**
- chen  "as example"
- pentonwaa  "as example"
- daNchen  "as an example"

**Restatement**
- khyy  "(that) is"
- klaawkhyy  "it is to say"
- nankokhyy  "that is to say"

**Sequence**
- laNcaakna  "after that"
- rxxkthidiaw  "in the beginning"
- toomaa  "later"

To sum up, the coherence system consists of two sub-systems: the anaphoric system with no null option and the cohesive system with a null option in addition to the macro-cohesive mini-system and the micro-cohesive system.

4.3 Attitudinal system
The attitudinal system is proposed to account for the fact that the speaker can choose to be objective or subjective as he presents information to his listener. That is, he can tell it as it is or he can also add comments on how he feels or what he thinks about it.

In Thai, a speaker's attitude can be manifested by a choice of three linguistic devices: presentative and evaluative matrix clause, appositive embedded clause with a sentential matrix clause, what might be termed "modal auxiliary" and adverbial, as in these examples:

Matrix clause
penpaidai waa khaw ca? lyym
   It is possible that he will forget.

duumyan waa khon klap pai mot lxxw
   It seems that people have all left.

Appositive embedded clause
   Z
phuuyiN mak ca caiyen syN pen siN thii dii
   Women are usually patient, which is good.

khaw yaak ca? pen naayok syN penpaimaidai
   He wishes to become the Prime Minister, which
   is impossible.

Modal auxiliary
khaw ?aat ca? maa khondiau
   He may come by himself.

phuak naaNnai khuan pati?bat taam naN
   Those ladies in waiting should follow her.

Adverbial
khaw phuut thuuk lxxw thiiciN
   He said the right thing, in fact.

fon khoN fok taNwan nxxnxx
   It will rain all day for certain.

Study of the meaning of lexical items making up linguistic manifestations of the speaker's attitude in Thai reveals three distinct types of comment: confidence, hesitancy and evaluation. The choice of lexical items also indicates justification of each type of attitude. The speaker expresses his confidence of the accuracy of his information by using words which mean that it is either "obvious", "factual" or "natural".
Words used in expressions to signal the speaker's hesitancy indicate that what he is saying is either a conjecture or a conclusion based on probability.

Evaluative expressions indicate that the speaker thinks that it is either regrettable, unusual, ordinary, advisable, and important.

To sum up, the attitudinal system consists of two choices: objectivity or subjectivity. If the speaker chooses to be subjective and he is speaking Thai, he can express either his confidence, his hesitancy or his evaluation of the information he is imparting to his listener.

4.4 Illocutionary system

The illocutionary system serves the interactionary function. Two choices of illocutionary mode are made available to the speaker: non-interactive or interactive. In the non-interactive mode, the listener is considered "given" and he is not engaged in any interactive communication. In the interactive mode, he is involved, referred to and represented in the discourse.

This distinction is not exactly the same as the difference between a monologue and a dialogue, which is based on the number of participants engaged in the delivery of a discourse. A dialogue is naturally interactive. An exception may be a case of a party bore who turns a conversation into a sermon. A monologue, however, is not usually non-interactive. In order to reach his listener better, the speaker can engage his listener into a pseudo-conversation by using interrogative and imperative sentences or give a response to imaginary questions or comments of the listener. The choice also accounts for the use of raw "we", which tactically allies the writer with his reader. Study of Thai expository discourse shows that authors of papers in humanities exhibit a greater tendency to be interactive than those in the fields of social science and science and technology.

The illocutionary system as formulated here makes it possible to account for the existence of sentence types. Statements, positive or negative are non-interactive. Imperative and interrogative sentences are interactive. It also accounts for stylistic variation and types of discourse.

5. APPLICATION

A grammar, either sentence-oriented or discourse oriented, is formulated with different objectives. Structural grammarians aim at an empirical description
of a language as exhibited in the corpus of data. Generative grammarians aim at making explicit competence of a native speaker of the language in terms of filtering rules. Functional grammarians aim at an explication of linguistic devices in terms of how they serve the various socio-cultural functions. All types of grammar offer insights into natural language. These insights are utilized in various practical and intellectual fields such as language teaching, translation, communications, psychology and cognitive science.

The systemic discourse grammar as outlined in this paper can be applied in language teaching, stylistic study and language acquisition research. It provides language teachers with a concrete framework to teach, analyse and evaluate students' written or oral discourses. As a grammar of choice, it makes it possible to explain stylistic variation in terms of choices made by speakers or writers. It can be applied in designing tools for a study of language acquisition of native as well as foreign language. Acquisition of what might be called discourse competence may be a process which is gradual enough to be closely monitored and studied among young children and foreign language learners.

REFERENCES


--------- 1985. HANDBOOK OF DISCOURSE ANALYSIS.  


--------- 1968. Contrastive Analysis in Discourse Structure. Georgetown University Monograph Series on Languages and Linguistics 21. 39-64


Mathesius, V. 1961. A FUNCTIONAL SENTENCE PERSPECTIVE OF CONTEMPORARY ENGLISH ON A GENERAL LINGUISTIC BASIS. Prague: Czechoslovakian Academy of Sciences


--------