INTRODUCTION

0.1 A discourse is generally considered as a unit of linguistic performance which stands complete in itself. The name gives no information about the size, style or quality (Chapman, 1973) of it. A discourse is in Halliday's description an 'effective' or 'operational' unit of language, (ibid 1973). We reach the fullness of a discourse, through the linear progression of sentences encountered in the order which the performer gives them.

Discourse Analysis is a vast and ambiguous field. This term has came to be used with a wide range of activities. It is also a far too slippery ground, where any new proposition becomes target of much dispute. This disputability arises mainly from the inadequacy of research work that has been done in it. Despite the works of many linguists, discourse analysis is yet a kind of labour dubious and/or unexampled. The vastness of the subject is another factor which makes discourse analysis a complex area.

Due to the dissimilar treatments of different linguists, Discourse Analysis has become an impossible mixture of heterogenous mass. Sociolinguists, for example concentrate on the surface of social interaction manifested in conversation. Their discriptions
emphasise features of social context which are particularly relevant to sociological classification. Psycholinguists are particularly concerned with semantic relationships between constructed pairs of sentences and with their syntactic relationships. The techniques employed by these disciplines also differ in conformity with their interests. Sociolinguists pay much attention to spoken texts, and strive to generalise it. Psycholinguists, on the other hand, employs a tight methodology derived from experimental psychology to short written texts and sequences of sentences. Because of this no one theory or account of discourse has had a wide or strong enough acceptance to have an imperalistic monopoly on it.

Even before the advent of linguistics as a science, discourse analysis was existed as a study and very important results were achieved by critics. But their studies were the products of their intuitive knowledge of the subjects rather than any scientific system of analysis. Since discourse in language itself, the structure and pattern of discourse should have something in common with language. The method of analysis of sentences may, therefore, be extended to the analysis of discourses.

The analysis of discourse is the analysis of language in use. some scholars do not restrict the term to the description of linguistic forms independent of the purposes or functions which
these forms are designed to serve in human affairs (Brown and Yule 1983). Though discourse can be analyzed as something more than language also, using the techniques of linguistic analysis one can analyse the discourse to a certain extend. Stubbs (1983) suggests the idea more clearly when he states that discourse analysis consists of attempts to study the organization of language above the clause or sentence, and therefore larger linguistic units, such as conversational exchanges or written texts. According to Harris the term discourse analysis means the breaking up of a discourse into its fundamental elements or component parts, by standard distributional methods.

Above all, discourse analysis is based on the assumption that we can identify and isolate, from the sentences that comprise a text or speech, the linguistic features which give coherence to the entire discourse. A discourse is coherent, because it has some of the features that contribute to the cohesion of a text or speech.

0.2 APPROACHES TO DISCOURSE ANALYSIS

The need for the study of discourses as an important part of language was recognised by some linguistics sociologists and psychologists. There are three phases of the study of discourse analysis which fell with in distant chronological boundaries. In the earliest phase which is lasting upto 1960, there were a few clear suggestions that the text or the discourse ought to be the main
object of linguistic study. (Ingarden 1931, Buhler 1934, Hjemslev 1943, Harris 1952, Pike 1954, Coseriu 1955, Uldall 1957, Karlsen 1959, Slama Cazacu 1961, Hartmann 1964, Weinrich 1966). These arguments did not affect the main stream of conventional linguistics, because the available methodologies to analyse minimal units or isolated sentences were alone used for the study of the whole texts. Around 1968 linguists working mostly independently of each other, conveyed on the notion of "linguistics beyond sentence" (e.g. Heidolph 1966; Pike 1967; Crymes 1968; Dijkstra 1968, Harweg 1968, Harson 1968, Isenberg 1971, Koch 1971). But later on this perspective allowed to see only a fraction of the interesting properties of the text (C.f. Van Dijk 1972, Kintsch 1974:79, Weinrich 1976:148). The unity of the text was left obscure. In the year 1968 there was disagreement among representatives of the then dominant "transformational paradigm". The counter movements of "Casegrammar" (Fillmore 1968) and "generative semantics" (Lakoff 1968a, 1968b, 1968c: Emmon Bach 1968, Macawley 1968a, 1968b) attested to the dissatisfaction with the usual treatment of meaning in grammar. However the basic tenets were retained and many key issues such as expansion of study from sentences to texts, were hardly raised. For this reason the situation has remained unsettled. However in the year 1972 ushered a new stage of research towards alternative theories of language rather than revisions of older theories was proposed. The new works were manifestly critical of the foundations of sentence grammars and were

In other fields such as sociology and psychology the eminent scholars sought for a new approach. Sociologists opposed the context free abstractness of older methods and pointed out the importance of social interactions in language groups (C.f. Gumperz and Hymes (eds) 1972, Labov 1972a, 1972b). Psychologists embarked on studies of memory. The interdisciplinary demand for theories and models has been a major impetus in the development of text linguistics. It is clear that these disciplines are concerned with the process by which language is utilized by human beings.

Different scholars affiliated to different schools have accordingly carried on research on the level of discourse organization in language. The category of discourse analysts belong to widely differing schools of thought. One such school is that of the structuralists such as Harris, Pike and the formal method of Halliday and Hasan (1976) as well as the European Structuralists in the area of poetics and stylistics. A short account of some of the approaches are mentioned below:
0.2.1 The American Structuralist tradition

0.2.1.1. Bloomfield

It was the Bloomfieldian era in the middle of the 20th century (from the thirties through the sixties) in America. Bloomfield advocated in favour of a behavioural but taxonomic approach: he aimed at formulating discovery procedures which would help in the grammatical description of texts. The Bloomfieldians made a study of exotic American languages which were dying out. Their handling of syntax was in terms of 'Form Class' and 'Constituent Structure'. According to them, such data based description would yield the grammar of a language. They rejected mentalism out of their analysis: i.e they did not give any reference to meaning.

Segmentation and classification of sentences and their parts were the main principles underlying their approach to syntactic analysis. Structures beyond sentences were ignored. They believed that each sentence is an independent form. Analysis above the sentence level did not attract their attention.

0.2.1.2 Harris

Though a descriptive linguist basically, Harris has suggested many new methods of analyses by linguistic structure at various levels. Discourse analysis, one of the very important method of analysis he proposed is the basis of the present study.
We will have a detailed description of this method in the next section 0.3.

0.2.1.3 Pike

Tagmemics developed by Pike has extensive work on discourse. Many Tagmemists have worked on several American, Indian, Philippine and African languages. The tagmemic approach of discourse analysis recognizes a text to be divided into gross units, roughly comparable to the divisions of an oration recognized in traditional rhetoric (Hendricks, 1976).

Pikes' most important contribution to discourse studies was his insistence that certain chunks of human behaviour can be taken as actually given (Grimes, 1975, 21). These chunks are recognizable to those who participate in them. The non-participants who have a knowledge of the cultural system involved also can recognize them. Pikes' method is a starting point of the analysis of both verbal and non-verbal behaviour. According to Pike "The tagmemic approach to linguistic theory claims that certain universal invariants underlie all human experience as characteristic of rationality itself" (Pike 1964, 129).

The tagmemic theory maintains that language can be studied adequately only as a part of man's total behaviour. Emphasizing the need for a study of discourse, Pike (1964, 130) says that "the
production of separate, well formed correct sentences in sequence by no means guarantees either intelligibility of discourse or correctness of its structure", and that in order to give the full definitions and descriptions of sentences, clauses and phrases, their "discourse potential" should be studied.

Linguists according to tagmemic theory seek to describe the patterns of language. Pattern and pattern point are properly the primitives of linguistic theory. (Longacre, 1964). Patterns are called 'Syntagmemes' (i.e., the construction) and pattern points are known as 'tagmemes' (the elements of a construction). The concepts are co-relative. Syntagmemes cannot exist without tagmemes and tagmemes exist only by virtue of their placement in one or more syntagmemes.

The tagmeme involves more than mere labelling of a node in a tree. The tagmeme is a functional point not necessarily fixed in a linear sequence in which a set of items and/or sequences occur. A tagmeme is a correlation of function and set which are dependents on each other. Thus, a tagmeme is a correlative concept that relates grammatical unit with its function slot-filler correlation. Tagmemes are of different kinds: (1) obligatory or optional tagmemes (2) nuclear or peripheral tagmemes and (3) tagmemes which occur in fixed or movable positions.
Language is the structure of three simultaneous and interlocking modes: Phonology, grammar and lexicon (Pike 1967, 515). The grammatical description of a language must therefore contain (i) a phonological statement (ii) a grammatical statement and (iii) a highly sophisticated dictionary.

0.2.1.4 Longacre

In addition to the work done by Pike, the contribution to Tagmemics by Longacre is noteworthy. He has made significant contributions to tagmemic analysis of discourse (1968). He limited his work to prose. Longacre (1968) found there are six major types and some minor ones of discourses in the Philippine languages. These six major types are:

1. **Narrative discourses**: is characterised by its presentation of sequence in time, its setting in past or accomplished time and its orientation to a first or third person point of view. In other words, it recounts a series of events usually ordered chronologically in the past tense.

2. **Procedural Discourse**: Shares the characteristic feature of sequence in time with narrative discourse but it is set in projected or future time and it occurs in the second person as well as in the first or third.

3. **Hortatory Discourse**: attempts to influence conduct. It is also set in projected or future time because it tries to bring about
a change in behaviour and attitudes of addressee. Obviously it is oriented towards the second person.

4. **Expository Discourse**: does not have a necessary sequence in time nor does it focus on a particular setting in time. It is oriented towards the subject matter it explains rather than to a first or second person point of view.

5. **Argumentative Discourse**: tries to prove something to a hearer and tend to exhibit frequent contrast between two opposing ideas.

6. **Conversational Discourse**: takes place in conversations between two or more people. Oddly, although Longacre expresses interest in this last discourse type, his analysis is generally limited to other five and relies mainly on edited text.

The other minor types include Dramatic, Activities and Expository Discourse.

0.2.2 **EUROPEAN STRUCTURALISM**

The European structuralism was developed in ways differing from American structuralism. The difference occurs (i) in the positing of the abstract relational structure (the sign) underlying an utterance (ii) its relation to other signs to which it bears a
paradigmatic relationship and the distinction between the syntagmatic and paradigmatic relation among signs. The European structuralism began in 1916 with the circulation of Ferdinand de Saussure's, 'Cours de Linguistique Generale'. There was no sharp demarcation of literary and linguistic studies during this era.

European structuralism appears under several distinct schools. The Prague school, The Russian formalists, the Copenhagen or Glossematic School and the London School.

0.2.2.1 The Prague School

The important personalities in this school are Trubetzkoy and Roman Jakobson. Others include Mathesius and his followers. They developed the Saussurean doctrine of polarity including that of syntagmatic and paradigmatic relations of linguistic sign. This group is well known for advances in phonology.

Roman Jakobson has distinguished studies in rhetoric and poetics. He discerns a pattern of metonymy (Curiously related to similarity and contiguity disorders found among studies of aphasia) running through both isolated linguistic expressions and larger stretches of discourse. Jakobson's formal analysis of poetry makes a distinction between poetic and other varieties of discourse and provides us insights which aid in the understanding of poetry. Jakobson presents the communicative aspect of the poetic process,
and poetics as a part of linguists unconstrained by principles such as sentence being the highest analyzable unit.

Jakobson's presentation of Schema of elements of speech events were later incorporated into process oriented linguistic analysis such as that of sociolinguists. The speech events, 'addresser' 'addressee,' 'context' 'message', 'contact,' and 'code', shape the utterances (Jakobson 1960). The Schema have greater methodological importance than Jakobson's other discoveries. (symmetries, antisymmetries, balanced structures etc.).

0.2.2.2 Russian Formalism

The Russian Formalists studied both poetic language and the structure of prose texts, seeking to arrive at the structural interrelationships of the various systems aimed at discovering the universal principles that are basic to creative writing. All the uses of language are considered as integral to the system of human culture. They studied both the poetic language and the structure of prose texts to find out the structural interrelationship of the various systems.

Among the Russian Formalists the most prominent are Vladimir propp, Shklovsky, Eichenbaum and Tomashevsky. The first of these contributed to the most significant study on the Morphology of Folktale (1928, 1968) which is an analysis of one hundred
Russian Folk tales. He was concerned with the nature of narrative literature. He has given the master structure of the Folktale as a literary product.

0.2.2.3 Copenhagen Glossematics.

The Glossematics poetics also investigated the linguistic properties of literary texts like that of the Russian, Prague and French school of structuralism. Louis Hjemslev considers a set of texts as a linguistic object. Segmentation and inductive generalization reveals the regularities of grammar in the text. According to him from such regularities by deduction all the possible texts of a language may be predicted. But Hjemslev's grammar is an abstraction. Its algebraic notations and relational intricacies are difficult to utilise in any formal procedure for segmentation of texts. To him, as for the other European Formalists and structuralists poetics is an extension of linguistics, the theory of all manifestations of language. But we realise that his notion of text is ambiguous.

0.2.2.4 The London School

The London School dates from John Rupert Firth (1944). To him Language is a meaningful activity; Linguistic analysis is the analysis of various types of linguistic meaning. Analysis involved in various levels - phonetic, phonological, grammatical and situational or semantic level. Into the scheme of analysis Firth
built a more abstract context of situation. Its categories were the relevant features of participants (persons, personalities) including verbal and non-verbal action, the relevant objects and the effect of the verbal action. Meaning which was considered the aim of analysis in the extralinguistic context of situation or the intralinguistic ones of grammar, phonology or other levels. That is the extralinguistic context of situation involves the 'Process approach' the intralinguistic analysis includes syntagmatic elements which are parallel structures with paradigmatic elements at different points in the structure. Eg:- at the grammatical level in a structure of NP-V-NP; a noun, pronoun or nominal group may occupy the place of NP. It is to Firth that we owe the concept of 'Collocation' and 'colligation' that is the regular concomitance of words and relations between classes and categories in grammar.

0.2.2.5 French Structuralists

Levi-strauss believed in the universality of myth as expressed in varied tales. To him myths are the embodiment of ideas revealing thought process. He has discussed various codes such as culinary, olfactory, acoustic, sociological etc. Many of which might be relevant to a certain discourse. He claims that the various codes express the fundamental semantic oppositions. Levi Strauss's Mythologiques (4 volumes) are extensive and significant as a structural analysis. He brings together the mythology of South and North America in order to illustrate the holistic nature of the
human mind. Levi-Strass's is also a type of discourse analysis. But the method fails when myths are drawn from different cultures. No common motif seems to emerge. Yet his method like Jakobson's poetic analysis is an attempt at the interpretation of fictional discourse.

The French group includes Greimas, Bremond and Todarov besides critics such as Roland Barthes and Genette. These structuralists in their study of poetics aimed at isolating the basic structures of fiction and arriving at their laws of combination. These laws would identify a composition as narrative. They tried to relate the basic structure of fiction to those of logic and grammar.

From the study so far, we have understood the structuralist critics as well as those who apply structuralist standards to literary texts. Roland Barthes approaches literary texts from the rhetorical point of view. He focuses on mysteries of identity. The readers can identify textual features which is a part of a hermeneutre process with his model of various roles. He also shows the manner how a text is patterning. Barthes, Gerard Ginette, another rhetorician, believes that the interpretation of a text depends greatly on the activity of the reader.
0.2.3 SOME OTHER INDIVIDUAL CONTRIBUTIONS

0.2.3.1 Halliday

Halliday (1977) defines text as a semantic choice in a social context, clause, group, word and morpheme complexes and structures formed by paratactic or hypotactic combinations at the several ranks. He regards each type of unit in turn as 'structural composites' e.g. a clause is formed out of elements such as agent, process, extent (Transitivity system) as well as Modal and Propositional elements. There is also a third structure viz-the Theme-Rheme one which is part of the textual component.

In systematics, Halliday's notion of 'cohesion' is an important contribution to discourse studies. Halliday proposes that a text in English displays inter-sentential relationship in terms of textual presuppositions. He identified this as 'ties'. Five types of such cohesive devices are distinguished for English (Halliday and Hasan 1976). In this lucid and metaculous study of cohesion he defines 'text' as "a semantic unit related to a clause or sentence not by size but by realization, the coding of one symbolic system in another" (Halliday and Hasan 1976). Cohesion is part of the system of the language and operates through conjunction, reference, lexical cohesion and substitution.

Halliday's mode of analysis is a detailed one in terms of the various components of language and their grammatical
realizations. It is formal and considers the given text as a product. Yet we discern the functional orientation of Halliday's approach, for there is a speaker writer slant to the way in which the text is analyzed as for instance in the theme rheme and information pattering.

0.2.3.2 Van Dijk

Text grammar is proposed by Van Dijk (1972, 79, 77). He presents arguments in support of the need for a text approach to the grammar of a language. His work on textual structure is significant. The concept of 'macrostructures is Van Dijk's most important contribution to Textual studies. Macro structure, according to Van Dijk's, reflect the global semantic planes of a text at the deep structure level. He maintains that Chomsky's concept of competence needs revision on empirical, psychological and methodological grounds. The tradition of 'text linguistics' or text grammar has flourished in Europe for nearly three decades. According to Van Dijk (1972) the approach of linguists in this area may be summed up as this: "Text or discourse are seen as the transphrastic realization of the system, as a set (if closed, a corpus) of utterances defined by a certain number of superficial properties of recurrence, substitution and co-occurrence.

A distinction is made between 'text' and 'discourse'. A discourse is regarded as an 'empirical' cognitive and social verbal
unit manifested as 'verbal utterances', visual or auditory in nature (Van Dijk 1977). Discourses have underlying text structure. A text is an abstract theoretical construct embodying discourse structure. Such a concept of 'text' is not to be confused with the general usage of the term for any printed or written passage.

Petofi (1969, 1971, 1972) also has been engaged in the development of text grammar. He (1973) considers text grammars to be generalized and expanded sentence grammars. Van Dijk and Petofi explored the pragmatics and the coherence conditions or connexity of discourse. Van Dijk engages himself in interdisciplinary discoursal studies including both linguistics and poetics. Van Dijk (1972) explores the global text structures and thematic structures under the linguistic and poetic process of metaphorization in a poem.

The fundamental sentence perspective (FSP) studies of Firbas and Danes extended the domain to include textual structure. Danes in particular studies textual unity in terms of intersentential theme – rheme arrays in a connected text. He employs for this purpose a notion called 'thematic progression' (Firbas 1974, Danes 1974).

0.2.3.3 Grimes

In the book 'The Thread of Discourse' Joseph Grimes (1975) attempts to show the sorts of things a linguist could find
out by looking beyond sentences. According to him a discourse has six areas, which correspond to the six parts of his 'Papers on Discourse' (1978). First, there are studies on morphology. Certain morphological information is shown to tie in with the total structure of discourse. He says that other morphological categories give information about the specific lexical items to which they are attached while others indicate syntactic constructions and agreement. The second area of discourse study in Grimes (1978) deals with reference and focus mainly on pronominalization. The studies show that there appear to be two distinct strategies that language uses for establishing and maintaining reference. The third area of discourse analysis in these studies shows that some languages have a clear-cut distinction among kinds of discourses, such as discussed earlier between explanatory, horatatory and argumentative. A fourth area demonstrates how some discourses are full of particle words that mean nothing in themselves but which act as pointers to discourse structure when considered in a larger context. In the fifth area of discourse analysis according to Grimes, a systematic repetition pattern called 'linkage' is used to either join together two consecutive sentences within a paragraph or to show the boundary between paragraphs. Finally the sixth area is composed of a miscellany of other linguistic signals. It becomes simple to explain using discourse contexts and difficult to explain without them.
When we make a survey on some empirical models of discourse analysis, the process-product binary opposition is the view of discourse as an ongoing process or as a complete product. Thus the study of oral discourse in the classroom situation (c.f. Sinclair, J. Mcht and R.M. Coulthard 1975) takes into account, aspects of the evolving speaker-hearer interaction. Sinclair, Coulthard and Jones, present hierarchical models of 'discourse' consisting of the units, lesson, transactions, exchanges, moves and acts on the basis of the data collected from classroom situation (Sinclair and Coulthard 1975 and Jones 1976).

The discourse analysis influenced by the theoretical insights provided by speech philosophical notions as the "speech act" of Austin (1962) and Searle (1968, 69) and the sociolinguistic notion such as 'speech event'; 'speech functions and factors' of Hymes (1962, 1964 and 1972) and Labov (1972).

Widdowson and Mountford propose models for discourse analysis with an applied orientation towards language teaching. Widdowson (1973b) identifies some communicative acts as rhetorical devices in scientific prose in written English. Werlich presents an overall textual organization in English in terms of 'thematic base expansion' of sentence sequences, paragraph structures and text types (Werlich 1975).
0.3 MODERN TRANSFORMATIONAL APPROACH BY HARRIS

Harris was the first to recognize the need to extend modern linguistic investigation beyond the sentence level. The structural methods he proposed for the analysis of discourse is in terms of structural equivalence and distribution of sentences in combinations. As Harris (1960,8) puts it "The basic operation is the forming of the classes of morphemic sequences".

The procedure he suggested provides statements of occurrence of elements relative to other elements within limits of discourse. (Harris, 1956) In this procedure there is an implication that discourse is structural or formal level which is higher than that of the sentence and lower than that of the language organization. His primary concern is with the form all distribution of the sentence without reference to meaning "The formal features of discourses can be studied by distributional methods with in the text. The method does not give any new information about the meanings that are being communicated in the discourse under investigation. (Harris : 1952, 1970, 357).

Harris with his emphasis on distributional criteria represents the structuralist tradition. He posited a form of transformational analysis as early as in 1950s. As Harris (1952,1) states "This method is formal, depending only on the occurrence of morphemes as distinguishable elements; it does not depend upon the
analysts' knowledge of the particular meaning of each morpheme". The theory outlined by him keeps two aspects of discourse analysis separately: (1) extending descriptive linguistics beyond the sentence level and (2) the correlation of 'culture' and 'language'; though he never really explored the latter.

Discourse Analysis as proposed by Harris (1952 1-30) consists essentially of the following steps. From a given particular text the linguistic elements (morphemes or sequences of morphemes) which have identical environments will be first collected within a sentence, these will be equivalent to each other. Thus if we find the sentences AF and BF in the given text, it will be assured that A=B, i.e., A is equivalent to B or that both A and B belong to the same equivalence class. Thus collect all the linguistic elements which have equivalent environments. They will be equivalent to each other. This type of operation enables to group many, or all the, linguistic elements or sections of many particular text into few equivalence classes. The text can be represented in terms of these classes which can be put into double array the horizontal axis representing the equivalence classes contained in one sentence, and the vertical axis representing successive sentences. This tabular arrangement makes it possible to inspect some characteristic pattern of the text as a whole. Various kinds of information about the text can be derived and certain critiques of the text can be made.
The determination of equivalence classes is simply based on finding elements which have identical environments. The more frequently repeated words of a text are considered as the 'key words' of that text. The various occurrences of such words can be put into one column; i.e., one equivalent class and the neighbouring words can be put into another single equivalence class because they occur in identical environments. An illustration of the procedure in detail as it is applied by Harris (1952) to the following text is given below.

Millions Can't Be Wrong!

Millions of consumer bottles of X-have been sold since its introduction a few years ago. And four out of five people in a nationwide survey say they prefer X-to any hair tonic they've used. Four out of five people in a nationwide survey can't be wrong. You too and your whole family will prefer X-to any hair tonic you've used! Every year we sell more bottles of X-to satisfied customers. You too will be satisfied!

In the text given above "Millions is equivalent to "four out of five people in a nation wide survey" since they occur before "can't be wrong" and/or "say they prefer X-to any hair tonic you
have used" which also constitute another equivalence class. Utilizing the repetitions found in the text the analysis will go on to construct chains of equivalence class, i.e. the other environments occur for **Millions** and for out of five etc. For **Millions** there is another environments such as of Consumer, bottles etc. It must be certain that any two members of one equivalence class have at least one environment in common. After finding out some patterned distribution of the classes, it can be possible to find out some structural fact about the text in terms of these classes. Harris (1952) states that "if two words occur only in identical or equivalent environments then there is no difference in their distribution. After recognizing two equivalence classes, one can be marked as P which includes:

i. Million

ii. Four out of five people in a nation wide survey, and other

W which includes

i. Can't be wrong

ii. Say they prefer X-to any hair tonic they've used.

The text will then be segmented into successive intervals in such a way as to obtain, in each interval, occurrences of the same equivalence classes.

If the classes are P and W then there are a few successesions as PW. The title of the advertisement is represented as PW. (P = Millions; W = can't be wrong). The first sentence after the title
contains a P 'the word millions' but the rest of the sentence neither equals nor contains W. We may perhaps be able to classify this sentence when our analysis proceeds.

The second sentence of the text is clearly PW and so also the third. The text is segmented into successive stretches so that each stretch will contain just PW and no more. These stretches will then be the successive rows of the double array. They may sometimes be whole sentences, but not always necessarily. They may sometimes be separate sections of compound sentences. Each of these will have its own sentence structure but it may also be a stretch is taken out of a sentence. For example, the stretches, Millions of people – four out of five – can’t be wrong – when they say they prefer X – seems to consist of PPWW, but may be reduced to two PW intervals (1) Millions of people can’t be wrong (2) and four out of five say they prefer X–

Sets of Equivalence Classes

The attempt to divide a text into intervals containing the same equivalence classes (in the same relation to each other) generally succeeds throughout a whole text. There may be individual sentences here and there which simply do not contain these classes. These may turn out to be introductory sentences, or off shoots of some other set of equivalence classes. There may also be successive sections of the text, each of which contains its own
equivalence classes different from those of other sections. These may be paragraph like, or chapter like sub-texts within the main text.

In the cause of seeking intervals which contain same classes, the procedures will discover the limits of this sameness, i.e., the points at which text intervals containing different classes are obtained/exist. In the general case, a text will be reduced not to a single set of identical rows, but to a succession of sets of identical rows, with occasional individually different rows occurring at one point or another.

Having obtained this result, compare the various sets and individual rows to see what similarities and differences exist among them in the arrangement of their classes, whether the specific classes are different or not. Discover the patterns in the occurrence of such similarities among the successive sets and individually different rows. For example let a text came out to be AB TE TE A'B' EP EP, AB, KD, LM, K'D', MS, MS MS FBV Ms, then, using (TE) to indicate a set of TE intervals, and temporarily disregarding BV, can represent the text by AB (TE) A'B' (EP) AB KD (LM) K'D' (MS) . AB (TEX) A'B' (EP) and KD (LM) K'D' (MS) are structurally identical as both of them have the form W (ey, W' (yz). This form is a particular relation form of W, X, Y and Z. This text consists of two occurrences of this structure, with the
W of the first occurrence (that is, the AB) appearing again between the two structures (or before the second structure) and with a unique FBV before the end of the last structure. (Harris's 1952).

According to Harris (1952) this method gives information about how the grammar of a language is exemplified within it. As a result of this analysis of the occurrence of the elements in a text, the particular interrelations of the morphemes of the text as they occur in that one text can be discovered. Definite patterns can be discovered for particular texts or for particular persons, styles or subject matters. The formal conclusion is that from the particular pattern of morphemic distribution, it is possible to show consistent differences of structure between the discourses of different persons, or different styles or about different subject matters.

In addition to recognizing the members of equivalence discoverable in the text, the text is represented in terms of these classes also. That is to state in what way the members of these classes occur in the text, in what way the text can be said to be consisting of these classes. The difference between the first step and this one is like that between a word class list (dictionary) and a grammar.

Transformation and Discourse Analysis

Harris in another model developed in 1963 has applied some transformations which reduce some of the dissimilarity among
the various sentence structures in discourses. This is essential for setting up equivalence classes. "Since most of the equivalence are based on environment within a sentence structure, the dissimilarities among the various sentence structures of the discourse restrict the applicability of the equivalences. The method of linguistic transformations makes it possible to reduce some of these dissimilarities." (Harris, 1963.11).

Suppose, $S_1$ is a sentence which has a particular grammatical arrangement $G_1$ of particular morphemes (words) $M_1$ a transform of this sentence is $TS_1$ which has another grammatical arrangement $G_2$ satisfying certain conditions, of the same $M_1$. Grammatical arrangement of each sentence structure is subject to one or more transformations. Preliminary lists of English transformations were given by Harris in his earlier papers on discourse analysis and/or transformations (1952, 18-23 and 33, 1957 283-340).

If every sentence of a discourse is operated on by one or more of the transformation to which its structure is subject, the succession of sentences $S_1$, $S_2$, $S_3$........$S_n$ which constitutes the discourse will be obtained. This discourse contains a succession of transformed sentences $TS_1$ which is itself a succession of sentences. This succession of $TS_1$ will be called a transform (TD) of the original discourse. To a certain extent $TS_1$ paraphrases $S_1$, for each $S$ in the discourse, TD paraphrases the original discourse.
The availability of various transformations for each sentence of the discourse gives us a set of transforms TiD of the discourse, including the original as a succession of identity transformations on its sentences. Members of TiD differ from each other only by transformations. One of these TiD is the Kernal form of the discourse.

Optimal Transforms

The kernel form may not be always the most suitable transform of the discourse for discourse analysis. There are one or more optimal transforms of a discourse for the applicability of the equivalences. Let D is the original discourse and S the sentences in it. Then D' indicates the optimal TD and S' the sentences in it. S' is characterised by the property that there are more cases of various Si' having the same equivalence classes in the same grammatical position within the Si' than among the sentences of any other TD. The grouping of all the morpheme sequences into equivalence classes is necessarily the same for each TD. The various TD will be quite similar in the equivalence classes. Whatever the recurrence classes, the recurrence of equivalence classes of D' within the succession Si' will be simpler than the recurrences of the equivalence classes of any other TD within the sentences of that TD. Then D' is the optimal TD or optimal transform of D and we call the S' which comprise the sentences of D', the periods of D and S' comprise the sentences D'.
The optimal transform of \( D' \) is obtained by selecting particular transformations for each sentence of \( D \), and carrying out the equivalences on the resulting sentences. The obtaining of the optimal transform is purely an operation of transformations, and the application of the equivalences in a separate (discourse) operation. The recurrence of equivalence classes of \( D' \) within the succession \( S'_i \) will be simpler than the recurrence of the equivalence classes of any other TD within the sentences of that TD. Then \( D' \) the optional TD or the optional transform of \( D_1 \) and the \( S' \) the periods of \( D \). \( S' \) comprise the sentences of \( D' \). The optimal transform of \( D \) is obtained by selecting particular transformations for each sentence of \( D \) and carrying out the equivalences above on the resulting sentences.

The transformations required to obtain the optimal transform are more or less same as those required to reduce the discourse into Kernels, though there may be some differences. Let us examine an example given by Harris. The following text is analysed by Harris (1963)

Truman will, he is smart and
he isn't smart. He is democratic.
But he's a politician without question"

In the usual analysis of this text He's smart is the Kernel, and He isn't smart a transform of it. But in the optimal transform the analysis is as follows:
Truman is and isn't smart
Truman is and isn't democratic
But Truman is without question a politician

The first two periods each combine two kernels which are usually separated. As a result, the optimal transform does not match is with isn't (smart) but rather matches is and isn't (smart) with is without question (a politician), and isn't and Without question are both adjuncts of is, the former correlating with smart and with democratic and the latter with a politician.

Procedure

The preliminary sketch of the actual procedure of Discourse analysis is as follows: One can work (downward) from the original text, or (upward) from a kernelization of the text. The often recurring words or the morphemic sequences (other than the, and, and other words, which are common in almost all discourse) are roted first and then found out word in each other's neighbourhood among them and the grammatical relations among them. i.e., their relative positions within a sentence or a kernel.

The analysis is started by the breaking down of the grammatical stretches, sentence structures or constituents, that contain combinations of the same recurrent words, and seek the
transformations that will recast all others. It is breaking into tentative periods, the sentence structures which have these recurrent words in the same grammatical positions. If we start with Kernels, find out all the connected sequences of Kernels that contain the same combination of two or more recurrent words, and the transformations that will combine each of these sequence into a period.

The grammatical stretches may sometimes contain only one of the recurrent words that had elsewhere occurred in recurrent combinations. Check then whether it is possible to show by grammatical or by textual parallelism that the missing members of the combination are present in the zero form. If it is possible, we can proceed further. If it is not possible, seek the transformations that will bring these stretches into closest grammatical alignment with the already established periods containing these words. Finally find out those recurrent words that do not occur in combination with others, and seek the transformations that will align the grammatical stretches containing them into periods. Note the similarities or relations between the remainders of these periods and the remainders (or recurrent words) of all other periods, to see if it is possible to describe any part or another period.

When there are periods whose corresponding parts are equivalent, arrange them in a double array, each period being
assigned a row, and each relevant morpheme sequence a column. The rows show the relation of equivalence classes to their periods, and the columns show the successive members of each equivalence class in successive of the discourse.

Interpretation

Discourse Analysis permits useful interpretations in respect to the particular discourse. The types of discourse structure, and their interpretations are discussed by Harris in his later papers. As an example the anlysis of the following text is shown below.

"The optical rotatory power of proteins is very sensitive to the experimental conditions under which it is measured, particularly the wave length of the light which is used"

The starting point is to divide the sentence into sections. The sections are separated by slant bars and are marked by their grammatical classification (N = nounphrase, V=verbphrase, A=adjective phrase, P=preposition, C=conjunction, S=sentence-structure) promorphemes (pronouns etc.) are marked the class of their antecedents. If any further division of a section will not contribute to obtaining the desired result the section is not divided further.
1. The optical rotatory power of proteins is very sensitive to the experimental conditions under which it is measured, particularly the wavelength of the light which is used.

Now combine the sections into sentences filling in all indicated zero recurrences and combine the embedded sentences into the original sentence.

\[
S_1 = N_1 V_1 \text{ to } N_2 \\
S_2 = N_1 V_2 \text{ under } N_2 \\
S_3 = N_1 V_1 \text{ to } N_3
\]

The following equivalences are established from the above analyses: 'is very sensitive to' = 'is measured under' the experimental conditions = the wavelength of the light which is used.
The sentences of the original were transformed into periods that would be most similar to other periods obtainable from other sentences. When the successive sentences have been transformed into a sequence of periods there is the optimal transform of the discourse, which can be read through as a roughly equivalent paraphrase of the original text. Each successive line (row) in table 1 is a period. Each column is an equivalence class, every member of a column is related to every other member of the column by one of the equivalences listed above. The column arrangement makes it easy to exhibit the similarity among periods. The column C contains introducers of a period and connectors between periods P.

Discussion

The method used in Harris (1963) was this transformations which gave the optimal transform are given in a table. The periods are compared and some of them are grouped together. The meta-discourse periods could be made into C. The main shifts in each column can be given in another table. The differences in the column entries and all other details are shown in yet another table. The knowledge of the less specialised word meanings which was used in checking that the re-ordering of some periods in a table did not introduce meaning changes, in deciding the relevant synonymity of some words for coalescing of periods and in summerizing the covariation of columns. (Harris, 1963, 56).
Proteins have optical rotary very sensitive to the experimental power conditions.

Table I

<table>
<thead>
<tr>
<th>S</th>
<th>CHV</th>
<th>R</th>
<th>W</th>
<th>K</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Proteins</td>
<td>have</td>
<td>optical rotary power</td>
<td>very sensitive to</td>
</tr>
<tr>
<td>2</td>
<td>proteins</td>
<td>have</td>
<td>optical rotary power</td>
<td>measured under</td>
</tr>
<tr>
<td>3</td>
<td>particularly proteins</td>
<td>have</td>
<td>optical rotary power</td>
<td>very sensitive to</td>
</tr>
</tbody>
</table>

Table - I
The collecting (or recording) similar periods make the structure of the discourse clearer. The different sections of the discourse differ either in the classes which the periods contained, or in the major entries in the classes.

Analysis of many discourses of various types would extend and simplify the procedure for obtaining compact and reduced tables out of the table of periods of one discourse and may yield an inspectable method of scientific statement.

0.4 Discourse Analysis and Stylistics

One of the important goals of discourse analysis is the study of the literary style of written texts. Discourse Analysis gives enough information about how sentences are built up into larger units. These larger units become the corpus for the study of style which is a term referring to the way in which language is used in a given context, by a given person, for a given purpose and so on. Style can be applied to both spoken and written: both literary and non-literary varieties of language.

The study of style is known as stylistics. According to A.A. Hill (1958, 406) "stylistics concerns all those relations among linguistic entities which are statable, or may be statable, in terms of wider spans than those which fall within the limits of a sentence". Hockett (1958) considers it as a term of optional
structure, which language, as a potential offer, and an individual or a social group, makes a choice in its use for expressing the same thought.

Stylistics is an exercise or analysis in describing what use is made of language. It seeks to understand the way a writer exploits the possibilities of expression. It studies how the writer moulds the language to his own purpose. The analysis of style makes it known what the writer is doing with language.

To the linguists the investigation of style is essentially a scientific description of certain types and sets of linguistic structure that occur in a given text. When the types of linguistic units are identified a statistical approach for searching how the units are distributed in a particular text can describe the style of that text, it is assumed here.

Some stylisticians, notably Herdan (1964) (1969) and Winter (1961) have tried to formulate a statistical theory of style. Doležel sums up the theoretical assumptions of his school in the statement "style is a probabilistic concept". The probabilistic theory derives from the information theory and modern mathematics. Each linguistic unit in the text is treated as a sign and on the basis of a mathematical count based on a sample text one can predict the probability of occurrence of a linguistic item in the whole text.
Statistical theory of style has not found favour with critics of style. Ullmann rejects it because it does not take context into account, and says that it may fail to handle the data which is too fluid. It is not adequate to capture the subtle nuances of style, emotive overtones, evocative resonance complex and rhythematic effects and the like.

However the statisticians have made interesting contributions in stylistics. The counting technique has been extensively used. Josephine Miles (1967) was the first to use the counting technique on a significant scale to determine stylistic features. In style and proportion, she used concept of relative proportions of nouns, verbs and its style. She considers proportion an important concept "for the analytical study of an art". She finds "proportionings of materials" illuminating for understanding similarities and contrasts between different literary modes, and between different temporal manifestations of the same mode. Curtis Hayes (1966) in his studies of the "styles of Johnson, Gibbon and Hemingway", and Ohmann (1962) in his book "Shaw: The style and the Man", have made effective use of the counting technique to collect data on which to base their statements.

Problems of the stylistics of fiction:

The stylistics of fiction faces certain problems peculiar to this genre which make its stylistic criticism a particularly
difficult task. The very size of the novel creates problems for the students. Because of its big size, the book can't be read at one stretch, and by the time one reaches the last page one's impression of its earlier parts may get blurred. As also all kinds of linguistic patterns may occur in a novel and its is by no means an easy task to single out those which may be described as constituting the writer's style. Because of this it may be difficult to analyse a novel in its entirety.

In such cases sampling is an economic processes which relies upon a short passage (leaving conversations) which we get a stable frequency of stylistic variants. The frequencies of the linguistic items of a passage can be compared with the corresponding features in another text, which is regarded as a norm (John spencer, 1964)

0.5 The Present Model

The model of Discourse Analysis adopted in the present study is based on Harris. It has been considerably revised in the Department of Linguistics, University of Kerala, by professor E.V.N. Namboodiri. The theory thus revised aims mainly at (1) stylistics and (2) content analysis. The present work is restricted to stylistics only.

The frame work of Transformational Generative Grammar has been used for the study of the style (and of the content also) of
the discourse. This choice is determined mainly by the fact that stylistic analysis is focussed on formal features and a transformational model of grammar is the best available device for bringing out the syntactic characteristics and complexities of it. It permits the restoration of deleted portions of sentence and breaking down of complex sentences into simple sentences, there by giving a better idea of the 'literary build' of a work. The theory is generally concerned with the structure of sentences in isolation. But, we can expand the domain upto discourse level. By this expansion we are not ruling out the methodology of Transformational generative theory.

However, some of the details of sentence structures discussed in transformational literature are disregarded here because of their insignificance in the study of stylistics or content. The transformations shown here are 'optimal' in the sense Harris has used this term. The differences between optimal and ordinary transformations have been minimised as far as possible in this study.

The present model of discourse analysis has got a wider aim than that of Harris. First it finds out the total grammatical structure of a discourse. Secondly, it defines stylistic variations of discourses. It also shows how different writers and speakers express their ideas in effective styles. This will help the new
writers to improve their way of writings. It also segments the discourse into minimum units of content thereby making a systematic analysis of content possible.

A discourse D has a base structure (BD) and a surface structure (SD) derived out of it. The actual linguistic form of the discourse itself is SD. It is subject to stylistic variations. That is, two different SDs are possible to be derived from one and the same BD. If two different SDs have the same base structure (BD), then the differences between these SDs are nothing but stylistic and these SDs does not differ in their content.

The base structure of the discourse contains a set of base sentences with zero or more primitive adjuncts and combiners. The transformation theory assures that any discourse can be reduced into its base sentences which are particular types of elementary sentences having a fixed form and so not subject for stylistic variations. The transformations are structural changes without changing the principal meaning of the original base sentences. Hence the base and derived sentences will have the same principal meaning.

It is said that stylistic variations are not possible in BD and that SD is highly stylistic because many SDs are possible to be derived from a BD. An SD is a transform of a BD suppose, a
set, of transformational rules, Ti, derives SDi, from BDi. Another set, say Tj, derives SDj from the same BDi. Then, SDi and SDj are equivalent except for stylistic variations. The difference between the styles of SDi and that of SDj can be defined in terms of the comparison between Ti and Tj. In other words, the differences between the styles of two discourses can be defined in terms of two different sets of transformations which derive them from their base structures. This theory therefore claims that, leaving the features concerned with vocabulary, the style of a literary text can be defined to a large extent in terms of the types and frequency of the base sentences and transformations.

The present analysis consists of the following steps: First each sentence of the discourse is reduced into a set of base sentences and the transformations by which the original sentences are derived from the base sentences are shown. The base sentences and transformations are classified structurally and the number of occurrences and frequency of each class are computed. Finally the structural and/or stylistic pattern of whole discourse is inferred from the table of frequency count.

0.6 The texts under study

The texts analysed in the present study are 'Indulekha' and 'Marthandavarma' two popular Malayalam novels written by O.Chandu Menon and C.V. Ramanpillai respectively. The former was
published in 1889 and the latter in 1891. The authors were contemporary and members of Orthodox Nayar families. They had similar occupation also. Chandu Menon was a sub-judge and Ramanpillai a head clerk in the High Court.

However, the former belonged to North Kerala and the latter South Kerala. Their literary styles as well as the themes of their novels differ considerably. Chandu Menon was interested in social themes and Ramanpillai in historical themes. Marthandavarma, Dharmaraja and Ramaraja Bahadur are the three historical novels by Raman pillai (Preemaamṛtam is his social novel). But the social novel has not reached upto the level of the historical novels.

Indulekha and 'Sarada' are the two novels by 'Chandu Menon'. The first one is throwing light on the social structure of Kerala. The author observed a deplorable degeneration in the communities of Nayars and Namboodiris. He wanted the Nayar community to have equal status and previlages. He had given emphasis on the need and importance of English and higher education. It is a revolt against the joint family system. About seventy years ago, the social set up in Kerala was such that as seen in the novel.

Marthandavarma is more a historical romance than epic history. The author praised and glorified and recreated history
emphasizing the brighter side of life. Though these authors are different in the choice of theme they are equal in attitude and approach towards life. Both the novels have a romance note having a win over evils.

Though the styles are different in some way the novels, Indulekha and Marthandavarma, have some similarities also. Both these texts have some general features of style peculiar to the period in which they belong to.

In Modern Malayalam literature most of the writers use simple and short sentences, when compared to the earlier writings. The sentences are now with active verbs and with simple structures. They should be more clearer and more concise. But in these two selected novels we can see an elaborated presentation which is a feature of the 19th century Malayalam.

In the narrative discourse of these novels we can find clues in the choice of words and of syntactic constructions. In this the creators and the creator stand in relation to one another. In Indulekha the characters are common people. So the author uses a language that has spoken at home. It is not a formal one. There are many English and other loan words in plenty in it. The words and structure are not so complex as those of the Historical novel Marthandavarma in which a more complex style suitable to the
The sentences are long when compared to the former one. Subordination or embedding transformations are more than that found in Indualekha. The style of C.V. Ramanpillai is described as grand style by some critics.

Both these novels are popular in Kerala. They have translated in to English also. Indulekha is translated by W. Dumergue, C.S, published by Mathrubhumi, 1890, 1965. Marthandavarma is translated by B.K. Menon, published by Kamalalaya Book depot, Trivandrum, 1936.

0.7 The general organisation of the present thesis

The present work is an attempt to extend the transformational generative theory to the level of discourse. This study deals with the question how the transformational generative theory is useful in the interpretation of style of discourses. A detailed transformational discourse analysis of Malayalam sentences is beyond the scope of the present study. The base structure and Transformations presented in this thesis are not the out come of a detailed analysis of Malayalam in the model of Chomsky or McCauwley. However, they are found to be feasible in discourse analysis of the model developed in the Department of Linguistics, Kerala University.

Here the analysis is done in successive steps through four chapters. The first chapter 'Base structure' gives a detailed
account of the base structure of thirty sentences in each text. There are three major types of base sentences in both the texts. Each major type contains many subtypes. The frequency and percentage of each major type and sub-type are calculated. Three examples of each sub-type are given, wherever available, with their English meanings.

The second chapter describes the types of transformations that are found in both the texts. For each transformation separate transformational rules are used. There are many different types of transformations applied in the base sentences to generate the original sentences. They are grouped under separate heads, and their frequency and percentage are calculated. The examples of each transformation are given.

The third chapter 'Structural Analysis' analyses the original sentences into base sentences. The actual sentences are given showing how they are derived from the base sentences.

The fourth chapter is devoted to the list of peculiar vocabulary items that are seen in the text. The connectives that are described in the 'Structural analysis' (chapter III) is classified and grouped separately with their frequency. Vocabulary includes also the pronouns, caste names, honourific and polite terms, Sanskrit words, superlative words and long words which are relevant to stylistic analysis.
In the last chapter, conclusion, the important findings and observations regarding the application of Transformational generative theory on discourse analysis on literary texts are briefed.

The texts analysed in this study are given with their English translation as appendix I. This follows a list of select books and papers which were much useful for the present study.

The transcription used here is a kind of morphophnemic representation which is convenient and generally used by modern linguists who work on Dravidian languages. The retroflex phonemes are marked with a dot below them ( t, d, n, s). The palatal fricative and nasal are indicated by a bar above them ( n ̃ s) . The trill sound (R) is distinguished from flapped (r) by using capital letter for the former. The retroflex continuant which is found only Malayalam and some dialects of Tamil has been transcribed as 1

The symbols, notations and abbreviations are given in separate lists.