

On the structure of scientific texts

John Hutchins

The Library, University of East Anglia, Norwich

1. In the study of text structure there are basically two approaches which we may adopt. On the one hand, we may seek to understand the mechanisms of textual cohesion, to discover what makes a sequence of sentences into a coherent text. Such a study involves the discussion of anaphora, of reference and substitution, of ellipsis, of the role of conjunctions and 'sentence adverbs', of lexical and semantic cohesion between sentences, and so forth. The alternative approach is to take a more global perspective, to seek to understand the overall organisation of texts, to understand how one episode of a narrative develops from another and how paragraphs and chapters are built into cohesive wholes. Following van Dijk (1972) we may call the first approach the study of 'micro-structure' and the second the study of 'macro-structure'. The latter has long been the province of the literary critic, the rhetorician, and the analyst of folk-tales (e.g. Kinneavy 1971, Propp 1968). It is an area which linguists have neglected. Indeed it is only relatively recently that they have turned their attention to text structures at all, and it is perhaps natural that they should have concentrated primarily on 'micro-structure' (e.g. Grimes 1975, Halliday and Hasan 1976) since it is here that techniques used in the analysis of sentences can be most easily applied.

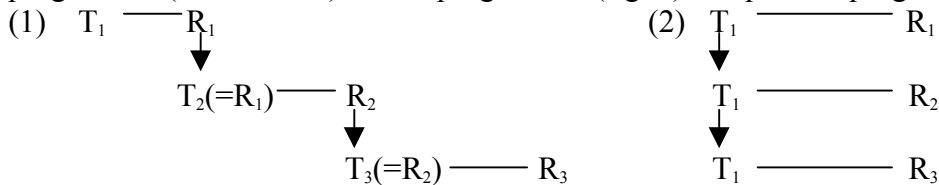
The two approaches to the study of text have thus proceeded with relatively little interaction. There has been scarcely any attempt to relate the two aspects of text structure within a comprehensive linguistic framework; the first real effort in this direction is represented by the work of Grimes (1975). Furthermore, the study of 'macro-structure' has been concentrated almost wholly on narrative texts, and primarily on fictional texts, short stories, novels and folk-tales. Narrative texts of a non-literary nature such as newspaper reports and historical writings have been generally neglected, and non-narrative texts have been almost completely ignored. In this paper it is my intention to examine the structure of an important type of non-narrative text. It is the expository text as represented by the article of a learned journal which argues for the revision of some accepted opinion in some area of academic study, i. e. the 'scientific paper'. I shall be concerned mainly with the macro-structure, of such texts, and I shall deal only with those aspects of micro-structure which are linked most closely to the global organisation of texts.

2. A characteristic of all texts at the level of micro-structure is that sentences are linked by some kind of 'thematic progression'. Every sentence adds some semantic content to what has preceded (unless, of course, it is a vacuous repetition), it builds upon what has gone before in order to convey something new. From the viewpoint of its communicational role a sentence has two parts, a theme and a rheme (Firbas 1966, 1974). The theme comprises those elements which are related in some way to what has been said or written earlier (in the same text or in preceding discourse) or to some feature of the extra-linguistic context, and which can be assumed by the speaker to be known about already – they can be taken as 'given'¹. The rheme is made up of those elements which convey information that is in some sense 'new' or unpredictable, and which therefore push forward the message being communicated.

Typically, thematic elements precede rhematic elements. It is more natural for speakers and writers to start from what is known before going on to what may be 'new' to their listeners or

¹ The linguistic devices for signalling 'given' elements are well known, the use of anaphoric pronouns, definite articles, repetition, relative clauses, generic expressions, etc.; they need not be elaborated here (cf. Halliday and Hasan 1976).

readers rather than vice versa; it is thus normal for the communicational or informational content of a sentence to show a gradual increase from a low point at or near the beginning to a peak at or near the end.² It is also more natural for thematic elements to refer back to some elements of the immediately preceding sentence. In this way the speaker or writer can convey his message by a natural progressive accumulation of ‘new’ information. In crude terms there are basically two ways a theme may be related to a preceding sentence: either it refers to elements of the foregoing rheme, or it repeats some or all of the preceding theme. We have thus two basic types of thematic progression (Daneš 1974): linear progression (fig. 1) and parallel progression (fig. 2).



(where ‘T-R’ represents the theme-rheme articulation of a sentence, and where the arrows indicate relationships of thematic cohesion.)

Linear progression may be illustrated by the simple ‘text’:

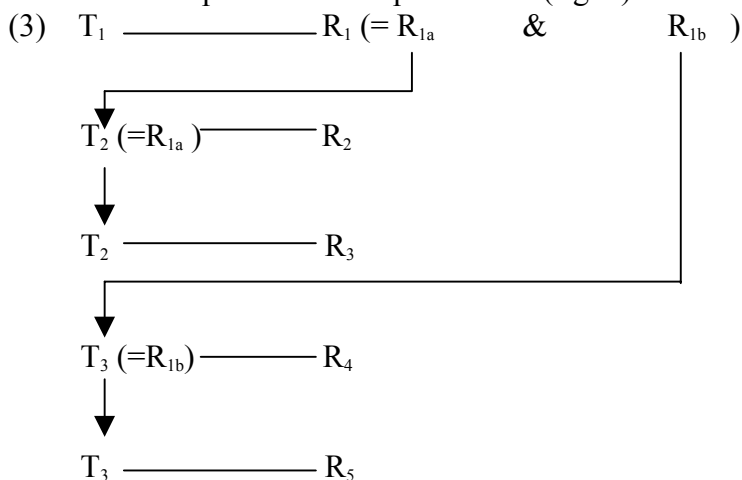
The boy was reading a book. It was about elephants. These animals live in Africa and India.
 $(T_1) \quad (R_1) \quad (T_2) \quad (R_2) \quad (T_3) \quad (R_3)$

and parallel progression by:

The boy came home from school. First he had something to eat.
 $(T_1) \quad (R_1) \quad (T_1) \quad (R_2)$

Then he went off to play football in the park.
 $(T_1) \quad (R_3)$

In most cases thematic progressions are mixtures of these two basic types. A common example involves the exposition of a ‘split rheme’ (fig. 3).



This may be illustrated by a paragraph such as:

All substances are divided into two classes: elementary substances and compounds. An elementary substance is a substance which consists of atoms of only one kind... A compound is a substance which consists of atoms of two or more different kinds...

² In this paper I shall be using ‘theme’ in Firbas’ sense as the element(s) with the lowest degree(s) of ‘communicative dynamism’ in the sentence and which are normally ‘given’ elements. I shall not follow Halliday’s usage of ‘theme’ as the starting point or ‘base’ of the sentence, i.e. as the first word or phrase in the linear utterance whether ‘given’ or not. In other words, I assume, for the sake of simplicity, the normal and most common case where the initial part of a sentence expresses ‘given’ elements. I have no doubt that the following account of text organisation could be elaborated to accommodate more subtle and complex thematic relationships.

In each of these thematic progressions we see that the first sentence provides the starting point or foundation for the following sentences. In this sense it functions as the theme for the paragraph as a whole. Such an observation is not of course new. Christensen (1967), for example, refers to the initial sentence of a paragraph as the ‘topic sentence’, to which the following sentences are related either coordinatively or subordinatively. Coordinative relations correspond to parallel progressions, and subordinative relations to linear progressions. Analogous remarks can be found in many writings on rhetoric from the earliest times to the present day.

3. A second basic characteristic of all texts at the level of micro-structure is that sentences are linked by what may be called ‘semantic progression’. This term embraces the numerous devices by which temporal, logical, causal and many other types of relationships between sentences or clauses are signalled. The most obvious of these ‘connectors’ are the conjunctions. Two clauses may be ordered in a temporal sequence, for example, by the subordinate conjunctions *before* and *after*:

After Bill arrived, Jim left
Bill arrived before Jim left
Before Jim left, Bill arrived

Alternatively, temporal sequence may be expressed by ‘sentence adverbs’ as connectors:

Bill arrived. Then Jim left
Bill arrived. Afterwards Jim left.

These and other temporal connectors such as *while*, *until*, *since*, *meanwhile*, *at the same time*, *subsequently*, are typical features of the semantic progression of narrative texts. But there are many kinds of non-temporal progressions. A few examples will suffice.

The typical connectors of descriptive passages are those expressing spatial relationships (*behind*, *in front of*, *beside*, *next to*, etc.), e.g.:

Opposite the town hall on the other side of the square stood the medieval church. Beside it was the old grammar school, behind it . . . and through the churchyard . . . [etc.]

Another common group of non-temporal connectors is formed by those expressing teleological relations, such as *because*, *therefore* (cause):

Because Mary had forgotten to buy some bread, Harry went to the shops. Mary had forgotten to buy some bread. Therefore Harry went to the shops.

since (circumstance):

Since the concert had been cancelled we went to the cinema.

so (result):

The soup was too hot, so we could not eat it.

Both circumstance and result may be regarded as weaker forms of the causative, thus the replacement of *since* and *so* by *because* or *therefore* is often possible, as it is here. At times the result relationship may be so weak that the connector can be omitted altogether:

The soup was too hot. We could not eat it.

Similar remarks apply to the reason connector *for*:

I gave in to his demands, for there was nothing else to do,
I gave in to his demands. There was nothing else to do.
I gave in to his demands because there was nothing else to do.

(where in the latter case ‘reason’ has been intensified to ‘cause’).

Other non-temporal connectors are those of concession (*although*, *even though*, *nevertheless*, *yet*, *in spite of*), condition (*if . . . then*), correlation (*as . . . so*), coordination (*and*), alternation (*or*, *either . . . or*, *alternatively*), antithesis (*but*, *on the contrary*), contrast (*but*, *by contrast*), and denial of expected consequence (*but*, *however*). The different senses of the connector *but* may be illustrated by the following examples:

He is not dead, but [on the contrary] he is alive
My horse is black, but [by contrast] yours is white

They set out for Paris, but [however] they did not arrive

It should be noted that a number of connectors may be used to underline the thematic progression of a text. For example, the introduction of themes in the development of a split rheme (fig. 3) may be signalled by *alternatively* or *on the other hand* if the relation between sections is one of alternation, by *however* if an expected consequence is denied, etc. Likewise, parallel progressions (fig. 2) may be signalled by *furthermore*, *in addition* or by the numbering of points *first*, *secondly*, *thirdly*. Other connectors with essentially textual functions are those of paraphrase (*in other words*, *in brief*), of illustration (*for example*), and of recapitulation (*I repeat, once again*).

One point should be stressed about semantic progression. There is no necessary correspondence between the sequence of sentences or clauses in a text and the logical or chronological sequence of their respective propositions, actions or events. Thus a speaker may invert the 'real' chronology (*Before X there was Y*) or he may invert the 'logical' conditional-before-consequent sequence (*Y if X*). Some connectors may thus be grouped in antonymic pairs; a familiar temporal example is the pair *before* and *after*, a teleological example is *so* and *for*:

I gave in to his demands, for there was nothing else to do

There was nothing else to do, so I gave in to his demands

Such inversions may involve larger segments of text; for example, the 'topic sentence' of a thematic progression might be placed at the end of a paragraph, where it could then serve as a kind of 'summary' (introduced perhaps by *in other words*).

4. This brings us to the macro-structural organisation of text. Here too we find features of thematic and semantic progression. It is quite natural to expect to find that the general principles of thematic progression should extend beyond the level of the paragraph to larger segments of text. Grady (1971) has shown how Christensen's analysis of paragraph structure can be applied to complete texts. Just as the paragraph has a 'topic sentence' which provides the foundation for its succeeding sentences, so too the text as a whole has an initial section of one or more paragraphs in which the 'topics' to be discussed are presented. Subsequent paragraphs (or groups of paragraphs) are related to this introductory section either coordinatively or subordinatively, in the same way as the sentences are in an individual paragraph. A similar picture results from an extension of Daneš' notion of thematic progression to relations among paragraphs. Just as the thematic elements of a sentence may be related to either the rheme or the theme of its immediately preceding sentence giving either a linear or a parallel progression respectively, so may the 'theme sentence' of a paragraph be related either to the 'new' parts of its preceding paragraph or to some part of its 'theme sentence'. The result is either a linear progression or a parallel progression of paragraphs, and the initial paragraph or paragraphs may be regarded as the 'theme paragraph(s)' of the text as a whole.

Clearly this can be no more than a broad sketch of the global thematic organisation of texts. It would need to be supplemented by other features, such as the notion of 'participant orientation' as described by Grimes (1975). In any single episode of a story there are rarely more than three participants directly involved (whether the participants are human or inanimate). The narrator will generally have something to say about each of these participants as 'themes', but often he will concentrate for a while on just one of them (as in a parallel progression) or on two of them alternating as 'themes' (as in a linear progression). The third participant may then be introduced either in the rhematic part of a sentence or as the 'theme' of a new sentence. The former brings in the third party in a natural unhurried way, but the latter can have a dramatic effect – particularly if the participant has not been mentioned for some time. Grimes has noticed that the thematic re-orientation of participants in this way often signals a break in the story and is often associated with the beginning of a new section or paragraph in a written narrative.

Semantic progression at the macro-structural level needs little demonstration. It is clear that paragraphs and longer text segments can be linked by temporal, spatial, teleological and other

connectors in just the same way as the individual sentences and clauses of a single paragraph. This is most obvious in texts relating a series of events in chronological sequence, where each event is described in a paragraph or group of paragraphs. Linking one event to another are frequently found such temporal connectors as *then, afterwards, at the same time*, etc. Equally familiar is the semantic progression of a descriptive passage in which paragraphs are linked by spatial connectors: *to the rear, on the other side, in the next room*, etc. In texts of an expository nature the connectors are mainly, as we should expect, those of cause, result, condition, circumstance, contrast, etc. Thus we find semantic progressions in which, for example, paragraphs are linked to their predecessors by connectors such as *for this reason, in these circumstances, alternatively, by contrast*, and in which a sequence of paragraphs is linked by the connectors *first, secondly, thirdly*, or by *on one hand, on the other hand*, and so forth.

However, semantic progression as described above represents only one aspect of the semantic organisation of text. Another aspect concerns what Kinneavy has called the 'modes of discourse'. Kinneavy (1971:36) identifies four basic attitudes or approaches of speakers or writers to the matter about which they wish to say something: telling a story, describing features, analysing and categorising, and evaluating and criticising. These four approaches are manifested by four discourse modes: narrative, descriptive, classificatory and evaluative. Texts are rarely if ever composed entirely in one mode, although often one mode will be dominant. In narrative texts (stories, histories) the 'narrative mode' is dominant; in topographical works the 'descriptive mode'; in many scientific studies it is often the 'classificatory mode'; and in works of criticism or theory it is the 'evaluative mode'. But however dominant one mode may be, the other modes are normally present to a greater or lesser extent: descriptive passages are common in narrative texts, classification and analysis are almost inevitably present in predominantly 'evaluative' texts, historical background and description are rarely absent from essentially 'classificatory' texts, etc. Each discourse mode may be characterised, as we might expect, by the type or types of connectors which occur most often. In the narrative mode it is of course the temporal connectors that are most prominent; in the descriptive mode the spatial connectors; in the classificatory mode the connectors of correlation, coordination, alternation, antithesis and contrast; in the evaluative mode those expressing concession, condition, circumstance, cause, result, etc.

5. Both semantic progression and discourse mode are largely determined by a further more 'global' aspect of the semantic organisation of text: the arrangement and distribution of the actual 'content' of the text. A good point of departure for our discussion is the analysis by Vladimir Propp of the Russian folk-tale, a work familiar to all concerned with the structural analysis of narrative texts (Propp 1968).

Propp showed that all Russian folk-tales which he examined could be analysed into a basic series of episodes. Episodes were loosely defined as sections of the tale which concerned an event involving principal participants of the tale and which could be summarised by a relatively simple statement or phrase. These he called 'functions'. Examples are: the treachery of the villain, the rescue of the victim by the hero, the struggle between hero and villain, the victory of the hero, and so forth. In all Propp identified 31 different 'functions' which always occurred in an invariant sequence. In any particular tale not all 'functions' would be present, but those which did occur would always appear in an order conforming to the observed sequence.

Since Propp's 'functions' can be expressed by sentences (simple clauses in most cases), it is evidently a simple operation to transform them into a reasonable summary of the plot of a tale. All that needs to be done is to go through the list of those 'functions' which occur and insert the appropriate temporal connectors. The result is a 'schema' of the semantic progression of a tale in terms of connectors, e.g.

- (4) *In the beginning* (description of setting)
 One day (an interdiction addressed to the 'victim' is violated)

and (the villain captures the victim)
Later (the hero is approached by the victim's family)
Then (the hero sets out on a journey to find the villain)
 etc.

The nature of the texts being analysed enabled Propp to define his 'functions' precisely, but this very precision precludes any easy extension to other kinds of narrative text. More abstract and therefore more flexible in this respect is the kind of analysis applied to a wide variety of stories, which seeks to identify a common pattern in the sequence of episodes. One familiar pattern often suggested or implied by many analysts of novels and short stories may be summarised, following Longacre (1974), as: Aperture, Setting, Inciting moment, Developing conflict, Climax, Denouement, Final suspense, Closure. The Aperture is one of the conventional openings of stories, e.g. *Once upon a time ...*, *There was once a beautiful princess...*, and it may often be absent, particularly in written literature. The Setting consists of those parts introducing the main characters and the principal location. The Inciting moment sets in motion the story proper; some event, object or person is introduced that disturbs the situation described in the Setting. In the Developing conflict the disturbance becomes more critical, problems and complications become more involved. The crisis intensifies until everything comes to a head in the Climax. Something then happens in the Denouement that makes possible a resolution of the conflict, a way out can be seen. But in the final suspense the outcome is kept in doubt by fresh complications, until finally everything is brought to a happy (or unhappy) end in the Closure.

6. In the analysis of the episodes of an actual story, the use of such a pattern requires the analyst to work at a higher degree of generalisation than he need to do when analysing a folk-tale into Propp's sequence of functions. The identification of a Propp function involves only a relatively simple 'matching' of the content of a text segment and the content of a 'function'. But the assignment of an episode to its appropriate place in an abstract story-pattern must follow a close examination of the syntagmatic relationships of all the episodes; the narrative content of the episode is less important than its role in the story as a whole. Propp himself made some tentative steps towards the identification of syntagmatically interdependent 'functions'; he recognised, for example, that a villainous act demands some later retribution, that the heroic rescue of someone in danger requires an appropriate reward, etc. The 'coupling' of episodes ('functions') can be pursued at higher levels of generality until an abstract pattern emerges which embraces all variants of the folk-tale. This path was followed by Greimas (1966) who reduced the main Propp functions to the sequence:

- | | | |
|-----|--------------|---|
| (5) | Orientation | (initial state of equilibrium) |
| | Complication | (rupture of this situation) |
| | Evaluation | (arrival/trial of hero) |
| | Resolution | (beneficent action of hero) |
| | Coda | (re-establishment of initial state, and recompense of hero) |

There are striking similarities between Greimas' sequence and the common story pattern given above, suggesting that there may be general principles of organisation underlying the macro-structure of all narrative texts. The search for such principles lies at the heart of the French structuralist tradition. In this spirit comes Brémond's notion of the 'narrative cycle'. It is summarised by Hendricks (1972: 102) thus: "According to Brémond every narrative integrates a succession of events of human interest oriented towards a goal. These events can be classified into two categories of elementary sequences: amelioration and degradation". At the beginning of a narrative there exists either a state of deficiency or a satisfactory state. From a state of deficiency or disequilibrium there is a movement towards a state of equilibrium or a satisfactory state, i.e. there is an amelioration. From a state of equilibrium there is a movement to a state of

disequilibrium, i.e. there is a deterioration or degradation. A number of such cycles may occur successively in a narrative, and one cycle may be embedded within another.

Brémond himself demonstrated the theory on an analysis of French fairy-tales (Brémond 1970), and Hendricks showed its applicability to an analysis of a short story by Ambrose Beirce (Hendricks 1972). Greimas' display of the basic Propp functions above reveals two cycles: from equilibrium (Orientation) through degradation (Complication) to disequilibrium (Evaluation) and then through amelioration (Resolution) back to equilibrium (Coda). The same two-cycle pattern can be applied to Longacre's plot sequence without much difficulty

(6)	Setting	Equilibrium
	Inciting moment	} Degradation
	Developing, conflict	
	Climax	Disequilibrium
	Denouement	Amelioration
	Closure	Equilibrium

The 'final suspense' episode may be regarded as an interruption of the last cycle, temporarily holding up the process of amelioration by reintroducing a state of disequilibrium.

7. To what extent do these analyses of narrative texts help us to understand the semantic organisation of expository texts? At first sight they might seem to offer little assistance, since as we have pointed out earlier expository texts are characterised by non-temporal relationships whereas the temporal succession of episodes is a basic feature of all narratives. It is true that 'succession' in some form is intrinsic to all these analyses, but nevertheless it does seem possible to apply the general principles to texts where the semantic progression is non-temporal.

Gopnik (1972) has identified three basic types of scientific paper: the 'controlled experiment' type, the 'hypothesis testing' type and the 'technique description' type.

In the 'controlled experiment' type of paper an experiment is described in which two or more groups are compared, e.g. one is tested and the other is not (the control group), or one is given one treatment and the other is given a different treatment. Gopnik found that the structure of such papers follows the pattern:

- (7) Establishment of experiment (the subjects to be examined and the treatment to be performed)
 Statement of techniques and methods
 Results of measurements and comparisons
 Conclusions, in the form of law-like statements

Applying Brémond's cycles, we notice that we have here a progression to a state of (reasonable) satisfaction, in so far as the author feels able to draw 'law-like' conclusions. When setting up the experiment there was, by contrast, a certain amount of confusion, some doubt about what the results might show: a state of 'disequilibrium'. The intervening sections represent steps on the way to the 'amelioration' of this situation.

In the 'hypothesis testing' type the basic structure is described thus:

- (8) Report of somebody's hypothesis, or of a number of alternative hypotheses
 Statement of methods and techniques used to test or compare the hypothesis or hypotheses
 Results of measurements and comparisons
 Conclusions confirming or disconfirming hypothesis, or deciding between hypotheses

This pattern amalgamates three kinds of progression. In one case the 'solution' proposed by another person (i.e. as establishing a state of equilibrium) is shown by the author to be unsatisfactory; the 'problem' remains to be solved (i.e. a state of disequilibrium). In another, a proposed hypothesis (by the author or by somebody else) which is still necessarily tentative (i.e. at

least partly deficient, in disequilibrium) is shown to be correct and a ‘solution’ to some problem is given further confirmation (i.e. some degree of equilibrium is achieved). In the third case, the unsatisfactory co-existence of alternative (conflicting) hypotheses is ameliorated by demonstrating which of them is to be preferred.

In the ‘technique description’ type no hypothesis is put forward. A new technique is tested on subjects not previously tried, or it is tested in new circumstances, and its effectiveness is measured. In terms of Brémond’s cycles there are two possibilities: either a previously satisfactory technique is found wanting (degradation of equilibrium), or a previously insufficiently tested technique is proved to be effective (amelioration of disequilibrium).

At a less abstract level of analysis we can interpret these text outlines in terms of the semantic progression from one section to another, as we did for the plots of Propp’s folk-tales. Here, of course, the connectors will generally be non-temporal, e.g. connectors of alternation, contrast and antithesis. For example, we could derive the following schema of semantic progression for Gopnik’s ‘hypothesis testing’ type of scientific text:

- (9) *Either* (hypothesis X) *or* (hypothesis Y)
 If (we apply tests A,B, C)
 so that (we have results D, E, F)
 then (X is true, Y is false)

We should note also that each of these types of scientific paper displays a variety of ‘discourse modes’. Whereas in narrative texts the ‘narrative mode’ is dominant, in these texts we find a mixture of the ‘descriptive’, ‘evaluative’ and ‘classificatory’ modes with occasional ‘narrative’ sections. In passages devoted to the account of tests of hypotheses and techniques the dominant mode is the ‘descriptive’; the assessment and comparison of results involves mainly the ‘evaluative’ mode; and the theoretical passages are largely in the ‘classificatory’ mode. But in practice many sections will have an intricate interweaving of all three modes.

8. Gopnik’s three types of scientific paper can all be reduced to a single Brémond cycle, either of degradation or of amelioration. A more complex paper might exhibit two cycles; the following structure seems to underlie many scientific papers:

- (10)
- | | | |
|----------------|---|--|
| the ‘problem’ | { | ‘current’ hypothesis/paradigm
demonstration of inadequacies
statement of ‘problem’ |
| the ‘solution’ | { | statement of ‘new’ hypothesis or of alternative hypotheses
testing of hypothesis or hypotheses
‘proof’ of hypothesis or of one of alternative hypotheses
implications of ‘solution’ |

The basic direction is from disequilibrium (the ‘problem’) to equilibrium (the ‘solution’). The elaborations arise from the author’s decision to explain the motivation behind his own proposal. Therefore he starts by outlining the current approach to the scientific phenomena with which he is concerned. He then demonstrates (or reviews the findings of others which demonstrate) the inadequacies of this approach. He can then state the nature of the ‘problem’ to be tackled. In this background (or ‘historical’) section of the paper the movement is clearly one of degradation from a supposedly satisfactory state to an unsatisfactory ‘problem’ state. In the body of the paper it is the author’s objective to ameliorate this state of affairs. As in Gopnik’s ‘hypothesis testing’ type above, he first proposes an hypothesis or some alternative hypotheses. He then provides evidence to test his hypothesis or to choose between the alternatives; and concludes either that his hypothesis is correct or that one of the hypotheses is to be preferred. In the final coda he may offer some comments on the implication of this ‘solution’ in some other problem areas, showing perhaps where it may solve other ‘problems’ (amelioration) or where it may cause further disturbance of the ‘current’ paradigm (degradation).

As with Gopnik's type (8) we may analyse the text structure in terms of connectors in a semantic progression, with the following result:

- (11) (it is generally held; X is true)
but
because (it has been found that A, B, C)
(the problem is Y)
Therefore (hypothesis Z is proposed)
If (we apply tests D, E, F)
so that (we have results G, H, J)
then (Z is true)
and
if (this 'proof' is valid)
then (V may be true)
and (W may be false)

The 'logical' structure of the schema underlines the care taken by scientists to present their arguments coherently and persuasively. (Whether the same pattern can be identified in expository texts outside the natural sciences must remain an open question.)

9. However, this structure has other facets. Firstly, it demonstrates that the scientific paper can be regarded as an integral part of the pursuit of 'normal' science, as defined by Kuhn (1970). An inevitable feature of 'normal' science is that problems arise in the interpretation of certain data within an accepted paradigm. The aim of 'normal' science is to find an explanation which does not disturb the fundamental premises of the science. Within these constraints hypotheses may be formulated and tested, and the successful ones are incorporated into the growing body of scientific 'knowledge'. The scientific paper is the record of the process of solving problems by putting forward and testing hypotheses, and as a record it becomes part of the 'objective' public knowledge of a particular science (Popper 1972, Ziman 1968).

Secondly, the scientific paper may be seen as continuing a long rhetorical tradition. Kinneavy (1971) has demonstrated the remarkably similar organisational principles underlying the structure of expository texts (i.e. texts which discuss, describe and evaluate natural phenomena and human behaviour) from the time of Plato and Aristotle, through Aquinas and Bacon, to the writings of modern philosophers and scientists. Basically there are two types of 'scientific' texts, those which search for generalisations and explanations from the observation of particular instances and those which seek to demonstrate the 'truth' of assertions or hypotheses. The latter display the characteristic logical progressions of deduction and induction, i.e. testing the consequences of a theorem and confirming the validity of a generalisation. They are typified by Gopnik's 'hypothesis testing' scientific paper.

The other type of 'scientific' text is exploratory in nature and illustrates the processes of 'abduction', the search for an hypothesis (or model) to explain or account for apparently conflicting observations and evidence. In this 'exploratory' type the author discusses the current 'dogma' on a particular issue, he observes that some evidence is not consistent with the predicted consequences of the dogma, and he searches for a better explanatory framework. His search may often take the form of a successive testing of various hypotheses until one is found to be reasonably satisfactory; this is put forward as a new 'model'. In Plato the search is frequently conducted as a dialogue between Socrates and one or more interlocutors; a succession of hypotheses is proposed and each in turn is rejected or modified (e.g. in *Laches* a series of definitions of 'courage'). Thus, the 'exploratory' type may incorporate a number of 'hypothesis-testing' sections, each characterised by deductive and inductive arguments. But the overall objective of the exploratory text is the search for a new 'model' to replace (or improve) one that has been found deficient; or in terms of Hegel's

dialectic, to find a new ‘synthesis’ which combines the best of the old dogma (‘thesis’) and the evidence which has revealed its weaknesses (‘antithesis’).

Kinneavy shows that texts of this exploratory type are founded on a basic pattern: Dogma - Dissonance - Crisis - Search - New Model. The pattern clearly underlies the type of scientific paper above (fig. 10); the current hypothesis represents the ‘dogma’, its inadequacies constitute the ‘dissonance’, the problem is the ‘crisis’, the proposal of a new hypothesis or of alternative hypotheses and its (their) testing represent the ‘search’ which is concluded by a solution put forward as a ‘new model’. To underline the correspondences, it can readily be confirmed that this basic pattern for the ‘exploratory text’ can be interpreted in terms of Brémont’s cycles, and that the result matches almost exactly the similar interpretation of the ‘scientific paper’ in fig. 10. The state of equilibrium represented by current ‘dogma’ is disturbed by conflicting evidence producing a ‘crisis’ state of disequilibrium, then efforts are made to ameliorate the position by the establishment of a ‘new model’ restoring some measure of equilibrium. We may summarise these correspondences by the following table:

(12)	Rhetorical tradition (Kinneavy)	Scientific paper (fig. 10)	Brémont’s cycles
	Dogma	Current approach	Equilibrium
	Dissonance	Demonstration of inadequacies	Degradation
	Crisis	Statement of problem	Disequilibrium
	Search	Statement of hypotheses	Amelioration
	{ Hypothesis	{ Testing of hypotheses	
	{ Testing		
	New model	Solution: ‘proof’ of Hypothesis	Equilibrium

10. The agreement is striking and suggestive. We would seem justified in believing that there are general principles of structural organisation underlying the form of many scientific and ‘exploratory’ texts. But we may well in fact go further. The organisational principles of thematic and semantic progression were shown earlier to be present at all levels of text structure. Does the same apply to these macro-structural patterns we have been discussing? Do they also have analogues in shorter text segments? There is, for example, a suggestive parallel between Brémont’s cyclic progression and thematic progression. We may regard the initial state of a cycle (whether disequilibrium or equilibrium) as the ‘foundation’ or ‘point of departure’ of a movement towards a ‘new’ state of equilibrium or disequilibrium. Interpreted in this way, a Brémont cycle reflects the movement from an initial ‘theme’ (i.e. the ‘theme paragraph(s)’ of a text) to the ‘new’ state of a ‘rheme’. Conversely, we might see the ‘theme-rheme’ articulation of sentences and paragraphs as reflecting micro-structural analogues of Brémont cycles.

Such speculations suggest the presence of some common factor binding the text-organising features of thematic and semantic progression and the exigencies of ‘content’ determining the (macro-structural) presentation of what the author wants to communicate. The common factor is perhaps involved in some way with what might be called ‘patterns of expectation’. If an author describes some ‘problem’, the reader normally expects him to put forward some ‘solution’; if he describes a certain ‘state of affairs’ the reader expects him to relate events which change the situation; if an hypothesis is proposed, the reader expects some testing of its validity; if an experiment is described he expects to be told the results; if a ‘topic’ is mentioned he expects to hear something ‘new’ about it.

Such patterns of expectations form a series of ‘oppositions’ (problem-solution, theme-rheme, etc.). Perhaps in the final analysis this is the essence of text structure: a network of interlocking, embedding, overlaying and underlaying oppositions at various syntagmatic levels. Some oppositions are concerned with ‘content’: problem - solution, question - answer, hypothesis -

test, experiment - result, initial state - change of state; some oppositions with semantic progression: condition - consequent, cause - effect, antecedent - subsequent, denial - assertion, etc.; and some with thematic progression: theme - rheme, given - new, topic sentence - commentary (or elaboration), theme paragraph - body of text. In each case the introduction of the first member raises the expectation that the second will follow. If it does not, then the 'text' is incomplete: just as a sentence is incomplete ('ungrammatical') if no rheme follows the theme, and just as an examination or trial is incomplete if the questions asked are not answered.

In brief, the structural form of a text is determined by the particular communicational functions it has to serve. The oppositions based on patterns of expectations provide the organisational principles underlying the formation of the kind of text we have been considering here (and perhaps of many other kinds of text). In their turn these oppositions are obviously related to more fundamental social and cultural determinants of behavioural and mental dispositions³. Just as language itself is fashioned by social needs (in the broadest sense), so 'texts' express and satisfy essential needs of social and cultural communication.

11. These tentative observations have much affinity to Halliday's conception of how the forms of language are determined by the functions they are to serve (Halliday 1973). He argues that there are three basic functions of language use which are embodied in the grammar as three components, the 'ideational', the 'interpersonal' and the 'textual', representing relatively independent formalised sets of options. The ideational component embraces those parts of grammar concerned with the expression of the external world and consciousness and the 'logical' relationships deducible from them; the ideational has thus two subcomponents, the experiential and the logical. The interpersonal component provides the means for expressing the speaker's role in the situation, his personal commitment and his interactions with others. And the textual component is concerned with the informational structure of the utterance, the relationship of each part to the whole and to the discourse setting.

At the level of the clause and sentence, the ideational functions are performed by case relationships (agent, patient, locative, etc.), verb categorisations (action, process, event), the voices (active, passive, middle), and so forth; the interpersonal functions are carried by 'modalities' (indicative, imperative, declarative, etc.) and realised by the familiar sequences of syntactic categories (subject, verb, object, etc); and the textual functions are borne by the articulation of theme and rheme and the distribution of 'given' and 'new'.

In this study we have, in effect, been discussing the operation of these functions at the higher levels of paragraph and complete text. Thematic progression belongs clearly to the textual component; semantic progression seems to belong most appropriately to the 'logical' subcomponent of the ideational; and what we have referred to as oppositions of 'content' (e.g. problem - solution, experiment - test) would seem to belong to the experiential subcomponent. The interpersonal functions seem to be present most clearly in the various 'discourse modes', narrative, descriptive, classificatory and evaluative, which reflect the speaker's (or writer's) stance or approach to the material or phenomenon he is talking about.

Each component offers the speaker or writer a number of options: the textual component allows him to choose, for example, a parallel thematic progression rather than a linear progression;

³ In particular, the oppositions identified here (and the 'patterns of expectations' to which they are related) would seem to underlie those mental activities which we call science, scholarship, research, study, etc. It may be legitimate to speculate that they represent constants of European-American 'scientific culture' analogous in some respect to the recurrent (hidden and subconscious) oppositions which Lévi-Strauss has 'discovered' in the myths of a wide range of primitive cultures. Whether syntagmatic oppositions of text structure such as problem-solution and hypothesis-test have a significance within our 'scientific' culture which is comparable to that of such paradigmatic oppositions as raw-cooked, fresh-putrid, profane-sacred, etc. in the myths of 'non-scientific' cultures is a question which would, of course, take us well beyond the speculative scope of the present essay.

the logical component offers a variety of semantic progressions (temporal, teleological, etc.); the interpersonal component provides a choice of discourse modes; and the experiential component allows him to choose from a range of possible macro-structures. The form of a particular text is largely determined by the writer's choice of one set of options rather than another.

12. This paper has thus been an investigation (in many respects preliminary and tentative) of some of the options available to writers of 'scientific' texts. It has shown how the structural features of certain common types of expository text are determined by the functions they serve. The selection of a particular global text strategy, within the experiential subcomponent, is made within the accepted framework of the procedures of 'normal science' and also in conformity with the 'traditional' forms of exposition and demonstration. As a result we may isolate certain basic patterns, the 'hypothesis-testing' type and the 'exploratory' type. The demands of 'scientific' argument affect also the range of options within the logical subcomponent, limiting the choice essentially to those semantic progressions appropriate to deductive, inductive and 'abductive' processes and to the description of causes, effects, results, conditions, consequences, etc. The particular function, of the 'scientific' text explains the preponderance of some discourse modes rather than others, the prevalence of evaluative and classificatory modes and the relatively subsidiary role of the narrative mode.

Whether any of the common features of text organisation which have been identified here represent 'universals' of a hypothetical 'text grammar' must, of course, remain an open question until more detailed and exhaustive analyses have been made of non-narrative texts of all kinds. In such examinations scientific texts must surely have a prominent part to play, not only because they constitute a significant proportion of all written texts but also because they are the expression of an activity whose procedures, techniques and objectives have been the subject of much detailed analysis. The influence of communicational function on textual structure is perhaps more direct and transparent in 'scientific' texts than in other types of text. For this reason, the linguistic analysis of such texts may well produce fruitful insights into fundamental and 'universal' aspects of linguistic systems and functions.

References

- Brémond, C. (1970): 'Morphology of the French folktale' *Semiotica* 2 (1970), 247-276
- Christensen, F. (1967): 'A generative rhetoric of the paragraph.' In his: *Notes towards a new rhetoric* (New York, Harper and Row, 1967), 52-81
- Daneš, F. (1974): 'Functional sentence perspective and the organisation of text' In: Daneš, F. (ed.) *Papers in functional sentence perspective* (The Hague, Mouton, 1974), 106-128
- Firbas, J. (1966): 'On defining the theme in functional sentence analysis' *Travaux Linguistique de Prague* 1 (1966), 267-280
- Firbas, J. (1974): 'Some aspects of the Czechoslovak approach to problems of functional sentence perspective' In: Daneš, F. (ed.) *Papers in functional sentence perspective* (The Hague, Mouton, 1974), 11-37
- Gopnik, M. (1972): *Linguistic structures in scientific texts*. The Hague, Mouton, 1972. (Janua Linguarum, Series minor, 129)
- Grady, M. (1971): 'A conceptual rhetoric of the composition' *College Composition and Communication* 22 (1971), 348-354
- Greimas, A.J. (1966): *Sémantique structurale: recherche de méthode*. Paris, Larousse, 1966
- Grimes, J.E. (1975): *The thread of discourse*. The Hague, Mouton, 1975 (Janua Linguarum, Series minor, 207)
- Halliday, M.A.K. (1973): *Explorations in the functions of language*. London, Arnold, 1973
- Halliday, M.A.K. and Hasan, R. (1976): *Cohesion in English*. London, Longman, 1976.

- Hendricks, W.O. (1972): 'The structural study of narration: sample analyses' *Poetics* 3(1972), 100-123. Reprinted in his: *Essays on semiolinguistics and verbal art* (The Hague, Mouton, 1973), 152-174
- Kinneavy, J.L. (1971): *A theory of discourse: the aims of discourse*. Englewood Cliffs, N.J., Prentice-Hall, 1971
- Kuhn, T.S. (1970): *The structure of scientific revolutions*. 2nd ed. Chicago, Ill., Chicago Univ.P., 1970
- Longacre, R. E. (1974): 'Narrative versus other discourse genre' In: Brend, R.M. (ed.) *Advances in tagmemics* (Amsterdam, North-Holland, 1974), 357-376
- Popper, K.R. (1972): *Objective knowledge: an evolutionary approach*. Oxford, Clarendon P., 1972
- Propp, V. (1968): *Morphology of the folktale*. Transl. by L. Scott, 2nd ed. rev. by L.A. Wagner. Austin, Univ. Texas P., 1968
- Van Dijk, T.A. (1972): *Some aspects of text grammars*. The Hague, Mouton, 1972 (Janua Linguarum, Series maior, 63)
- Ziman, J.M. (1968): *Public knowledge: an essay concerning the social dimension of science*. Cambridge, Cambridge Univ. P., 1968