



Some pages of this thesis have been removed for copyright restrictions.

If you have discovered material in AURA which is unlawful e.g. breaches copyright, (either yours or that of a third party) or any other law, including but not limited to those relating to patent, trademark, confidentiality, data protection, obscenity, defamation, libel, then please read our [Takedown Policy](#) and [contact the service](#) immediately

THE UNIVERSITY OF ASTON IN BIRMINGHAM

DISCOURSE VARIATION IN MEDICAL TEXTS:

Schema, Theme and Cohesion in Professional and Journalistic Accounts

Ph.D. 1989

Volume 1

KEVIN NGOZI NWOGU

Doctor of Philosophy

THE UNIVERSITY OF ASTON IN BIRMINGHAM

May 1989

This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with its author and that no quotation from this thesis and no information derived from it may be published without the author's prior, written consent.

THE UNIVERSITY OF ASTON IN BIRMINGHAM

DISCOURSE VARIATION IN MEDICAL TEXTS:
Schema, Theme and Cohesion in Professional and Journalistic Accounts

Kevin Ngozi Nwogu

Ph.D.1989

SUMMARY

This study is undertaken as an exercise in Comparative Discourse Analysis. Working within the framework of the branch of Discourse Analysis known as Genre-Analysis, this thesis examines differences in the organization of discourse in three parallel genres of written medical texts — the Abstract (ABS) accompanying a medical research paper, the Research Article Proper (RAP) and the equivalent Journalistic Reported Version (JRV) of the research article in a popular science magazine or in a newspaper.

The corpus for the study consists of 45 texts: 15 triadic sets of parallel texts representing an equal number of texts from each of the three genres on which the study is based. The approach to the analysis of features of discourse organization is eclectic and comparative. Thus, texts in the corpus are analysed for differences at three levels of organization — the levels of schematic structure, Thematic Progression (TP) and Cohesion.

A number of differences are observed across the three genres. First, the genres are found to differ distinctively in the way information is organized into hierarchical schematic units. Less distinctive pragmatic differences are observed in the realization and distribution of formal syntactic elements which constitute theme. Similarly, very subtle pragmatic differences are observed in the realization and function of Reference, Conjunction and Lexical Cohesion items in texts across the three genres. The results indicate that there is a great similarity in discourse organization between the professional genres (ABS and RAP) and that both genres differ distinctively from the popularized genre (JRV).

This thesis provides conclusive evidence to support the view that subject-matter is not the sole determinant of genre; that subject-matter is only one alongside a number of other contextual factors such as purpose, audience and medium of discourse by which genres must be characterized. To this end, this thesis is a contribution to the newly developing field of genre-analysis. But, more importantly, it is a contribution to research in the field of science popularization.

Key words:

Comparative Discourse Analysis, Genre Analysis, Medical Discourse, Science Popularization

ACKNOWLEDGEMENTS

My deepest gratitude goes to Mr Tom Bloor, my supervisor for inspiring me at every stage in my research programme with his wealth of experience, dedicated guidance and incisive criticisms. More importantly, I am very grateful to him for demonstrating unwavering faith in me and my work and showing enormous interest in my welfare and those of members of my family at all times.

My gratitude is also due to Professor Jan Firbas, Brno University, Czechoslovakia and Professor Betty Lou Dubois, New Mexico State University, U.S.A., for their contributions in helping me clarify issues on FSP conceptions of Theme and Thematic Progression. The same gratitude is due to Peter Ghosh, Science Editor, New Scientist for his positive response to my questions on aspects of science popularization in his magazine.

I am particularly grateful to the Association of Commonwealth Universities and the Commonwealth Scholarship Commission in the U.K. for sponsoring my research programme at Aston University and to Professor Sir Frederick Crawford, Vice-chancellor, Aston University for supporting my nomination.

My very sincere gratitude goes to Emilia, my wife, Chigozie, Ogechi, Ihuoma and Ndidiamaka, our children for their love, care, support and understanding.

Finally, I wish to thank Tracy Knight for the good job she has done in typing this thesis.

LIST OF CONTENTS

Title Page	1
Thesis Summary	2
Acknowledgements	3
List of Contents	4
List of Tables	9
List of Figures	11
Chapter One: Introduction	12
1.0 Background and Purpose of Study	12
1.1 Need for the Study	15
1.2 Statement of the Problem	20
1.3 Basic Hypotheses	23
1.4 Design of the Study	23
1.5 Outline of Chapter Content	25
Chapter Two: A General Review of Related Literature	31
2.0 Introduction	31
2.1 Approaches to Functional Variation in Discourse	31
2.2 Functional Variation and the Analysis of Scientific/Medical writing	33
2.2.1 The Description of Scientific Texts as Register	34
2.2.2 The Grammatical-Rhetorical Approach to EST Discourse	38
2.2.3 The Analysis of Scientific/Medical Texts as Genre	43
2.2.2.1 Genre as Literary Art	43
2.2.3.2 "Real-Life" Genres	45
2.3 Genre-Analysis and EST Discourse	48
2.3.1 Types of EST Discourse	48
2.3.2 The Structure of Scientific Research Papers	49

2.4 Investigations into Medical Research Report Writing	59
2.5 Investigations into Abstracts of Research Articles	65
2.6 Investigations into Science Popularization	68
2.6.1 Historical Perspectives	68
2.6.2 The Need for Science Popularization	69
2.6.3 Types of Science Popularization	70
2.6.4 Science Topics in the Print Media	71
2.6.5 Some Approaches to the Study of Science Popularization	73
Chapter Three: Research Methodology	83
3.0 Introduction	83
3.1 The Corpus	83
3.1.1 Sources	84
3.1.1.1 Coded List of Texts and Sources	85
3.2 Criteria for Data Identification	88
3.2.1 Representativity	88
3.2.2 Reputation	91
3.2.3 Accessibility	92
3.3 Criteria for the Selection of Corpus	94
3.4 Theoretical-Methodological Framework	95
3.4.1 Variation	95
3.4.2 The Comparative/Contrastive Framework	96
3.4.3 The Discourse Framework	98
3.5 Procedures for the Analysis of Data	99
Chapter Four: Variation in Schematic Structure	104
4.0 Introduction	105
4.1 The Concept of Schemata	104
4.2 The Surface Organization of Information in the Genres	108

4.2.1 The Research Article Proper (RAP)	109
4.2.2 The Abstract (ABS)	112
4.2.3 The Journalistic Reported Version (JRV)	116
4.3 The Schematic Organization of Information in the Genres	126
4.3.1 The Identification of Discourse "Moves"	126
4.3.1.1 Explicit Prefacing or Preparatory Expressions or Statements	128
4.3.1.2 Explicit Lexemes	129
4.3.1.3 Discourse Conjuncts	130
4.3.1.4 Grammatical Specifiers of Content Relations	132
4.3.1.5 Summary Statements	133
4.4 Description of MOVES	135
4.4.1 Outline of Canonical MOVES	135
4.5 Differences in the Schematic Organization of Information	161
4.5.1 Differences in the Number of MOVES	161
4.5.2 Differences in the Constituent Elements of MOVES	170
4.5.3 Differences in the Linguistic Features of MOVES	179
4.5.4 Differences in the Order of Presentation of MOVES	188
Chapter Five: Variation in Thematic Progression Patterns	196
5.0 Introduction	196
5.1 The Theme-Rheme Concept	198
5.2 Approaches to the Characterization of Theme	198
5.2.1 The Systemic-Functional Approach	199
5.2.2 The Functional Sentence Perspective (FSP) Approach	201
5.2.3 FSP and Danes' Theory of Thematic Progression	206
5.3 Some TP-Based Analysis of Discourse	210
5.4 The Verbal and Situational Context	214
5.5 The Identification of Utterance Themes	218

5.5.1 Types of Utterance Themes in Text	222
5.5.2 The Identification of Theme in Some "Exceptional" Utterances	225
5.5.2.1 Existential and Extraposition Utterance Themes	225
5.5.2.2 Interrogative Utterance Themes	228
5.6 Cross-Genre Comparison	229
5.6.1 The Realization of TP Patterns Across Genres	230
5.6.2 The Distribution of TP Patterns	237
5.6.3 The Characteristic Elements which Realize Theme Across Genres	255
Chapter Six: Variation in Textual Cohesion	266
6.0 Introduction	266
6.1 The Concept of Cohesion	268
6.2 Approaches to the Study of Cohesion	270
6.3 Halliday and Hasan (1976) Model of Cohesion	278
6.4 Methodological Issues	282
6.5 The Analysis of Cohesion in Sample Texts	286
6.5.1 Cohesion in the RAP	287
6.5.2 Cohesion in the Abstract	299
6.5.3 Cohesion in the Journalistic Reported Version (JRV)	304
6.6 Cross-Genre Comparison	309
6.6.1 Differences in Reference Cohesion	311
6.6.2 Differences in Conjunction Cohesion	323
6.6.3 Differences in Lexical Cohesion	331
Chapter Seven: Summary, Conclusions and Implications	342
7.0 Introduction	342
7.1 Summary of Results	343
7.1.1 Summary of Results on Schematic Structure	343
7.1.2 Summary of Results on Thematic Progression	347

7.1.3 Summary of Results on Cohesion	350
7.2 Conclusions	354
7.3 Theoretical Implications	357
7.3.1 Theoretical Implications for Discourse Organization and Varieties	357
Differentiation	357
7.3.2 Some Wider Implications	361
7.4 Limitations and Suggestions for Further Research	365
List of References	368
Appendices (Volume 2)	1
Title Page	1
List of Content	2
Appendix 1: The ABS And RAP Data	3
Appendix 2: The JRV Data	74
Appendix 3: Manual of Style for Contributors (<i>New Scientist</i>)	89
Table 14: Classification of Move 9	155
Table 15: Classification of Moves 10 and 11	159
Table 16: List of Moves and their Constituent Elements	160
Table 17: Constituent Elements in the ABS	171
Table 18: Constituent Elements in the JRV	172
Table 19: Constituent Elements Deleted from the ABS	173
Table 20: Constituent Elements Deleted from the JRV	173
Table 21: Average Sentences per Move	180
Table 22: Average Length of Sentences per Move	181
Table 23: Hesse (1984) Revised Categories of Logical Cohesion	285
Table 24: Sample Analysis of Cohesion in Introduction Section of the RAP	288
Table 25: Sample Analysis of Cohesion in Methods Section of the RAP	291
Table 26: Sample Analysis of Cohesion in the Results Section of the RAP	293
Table 27: Sample Analysis of Cohesion in the Discussion Section of the RAP	297

LIST OF TABLES

Table 1:	List of Texts and their Sources	85
Table 2:	The RAP: Distribution of Moves per Text	136
Table 3:	The ABS: Distribution of Moves per Text	136
Table 4:	The JRV: Distribution of Moves per Text	137
Table 5:	Canonical RAP Moves	138
Table 6:	Exemplification of Move 1	140
Table 7:	Exemplification of Move 2	143
Table 8:	Exemplification of Move 3	144
Table 9:	Exemplification of Move 4	147
Table 10:	Exemplification of Move 5	148
Table 11:	Exemplification of Move 6	150
Table 12:	Exemplification of Move 7	153
Table 13:	Exemplification of Move 8	154
Table 14:	Exemplification of Move 9	155
Table 15:	Exemplification of Moves 10 and 11	159
Table 16:	Outline of Moves and their Constituent Elements	160
Table 17:	Constituent Elements in the ABS	171
Table 18:	Constituent Elements in the JRV	172
Table 19:	Constituent Elements Deleted from the ABS	173
Table 20:	RAP Constituent Elements Deleted from the JRV	173
Table 21:	Average Sentences per Move	180
Table 22:	Average Length of Sentences per Move	181
Table 23:	Hasan (1984) Revised Categories of Lexical Cohesion	285
Table 24:	Sample Analysis of Cohesion in Introduction Section of the RAP	288
Table 25:	Sample Analysis of Cohesion in Methods Section of the RAP	291
Table 26:	Sample Analysis of Cohesion in the Results Section of the RAP	293
Table 27:	Sample Analysis of Cohesion in the Discussion Section of the RAP	297

Table 28:	Sample Analysis of Cohesion in the ABS	300
Table 29:	Sample Analysis of Cohesion in the JRV	305
Table 30:	Conjunction Cohesion Types in the RAP	325
Table 31:	Conjunction Cohesion Types in the JRV	328
Table 32:	Lexical Chains in a Sample RAP Text	336
Fig 2:	Complex TP with a Different Theme	209
Fig 3:	Complex TP with a Realization of a Split Rheme	210
Fig 4:	Realization of Simple Linear TP in the RAP	235
Fig 5:	Realization of Simple Linear TP Patterns in the RAP	240
Fig 6:	Realization of Simple Linear TP Patterns in the JRV	241
Fig 7:	Realization of Simple Linear TP Patterns in the ABS	241
Fig 8:	Realization of the Complex TP Pattern in the RAP	247
Fig 9:	Realization of the Complex TP Pattern in the JRV	248
Fig 10:	Realization of the Complex TP Pattern in the ABS	248
Fig 11:	Realization of the Complex TP Pattern with Embedding	254

LIST OF FIGURES

Fig 1:	Danes' Simple Linear TP Pattern	208
Fig 2:	Danes' Constant TP Pattern	208
Fig 3:	Danes' TP with a Derived Theme	209
Fig 4:	Danes' TP with Exposition of a Split Rheme	210
Fig 5:	Realization of Derived TP in the RAP	235
Fig 6:	Realization of Simple Linear TP Patterns in the RAP	240
Fig 7:	Realization of Simple Linear TP Patterns in the JRV	241
Fig 8:	Realization of Simple Linear TP Patterns in the ABS	241
Fig 9:	Realization of the Constant TP Pattern in the RAP	247
Fig 10:	Realization of the Constant TP Pattern in the JRV	248
Fig 11:	Realization of the Constant TP Pattern in the ABS	248
Fig 12:	Realization of the Derived TP Pattern with Embedding	254

CHAPTER ONE

INTRODUCTION

1.0 BACKGROUND AND PURPOSE OF STUDY

The extent to which the linguistic properties of professional texts written by specialists for fellow specialists in the fields of Science and Technology differ from their equivalent popular versions published in popular science magazines and newspapers for less-specialist or lay audiences has not been extensively researched. The few studies in this area, some of which have an overt media orientation have tended to focus primarily on the sociological factors which constrain the production of popular science materials (see Jones et al, 1978; Karpf, 1988; Taylor, 1975 etc). Others, which adopt a distinctive sociolinguistic and discourse-based approach have tended to emphasize differences in syntax and vocabulary (see Adams-Smith, 1986; Varantola, 1987; Dubois, 1985). To the best of my knowledge, no study has attempted a detailed examination of the difference which exists between professional science texts and their equivalent popular versions in terms of their patterns of discourse organization. The investigation reported in this thesis is an attempt to fill this gap. Two professional genres and one popular science genre are examined. The ABSTRACT (ABS) accompanying a research article, the RESEARCH ARTICLE PROPER (RAP) and the equivalent JOURNALISTIC REPORTED VERSION (JRV) of the research article in popular science magazines and newspapers are examined for differences and relationships in their patterns of discourse organization.

Written discourse of all kinds is known to possess distinctive features of discourse organization (see Cooper and Greenbaum, 1986, Blakemore, 1988). Professional research papers published in scientific journals are made up of sections and sub-sections which reflect stages in the development of the experiment they report upon. The structure imposed on the research paper by these sections is referred to in

this study as its rhetorical structure; that is, the functional division of the research paper into sections (Introduction, Materials and Methods, Results, Discussion) which reflect a rigid format for the presentation of details of a scientific experiment in a professional journal. This structure is almost always constant and similar in professional research papers across a wide range of scientific disciplines. It is analogous to the rhetorical structure posited for simple stories by Mandler (1978) and for expository texts by Meyer (1975).

In addition to having a rhetorical structure, professional research papers also have a discourse structure. That is, a structure which indicates the pattern of communication of the information contained in them, and which accounts for the internal coherence of their parts. This structure may differ from discipline to discipline and from genre to genre, even within the same discipline. For instance, all three genres in this study belong to the field of medicine. Yet, they are considered distinct genres because they represent distinct recognized communicative events with known public purposes, as well as exploiting forms and styles familiar to those who use them. Therefore, we are concerned in this study with comparing the organizational features of three related or parallel genres of written medical discourse. The genres are considered parallel because their texts report on the same research investigation, in the same discipline, but in different contexts and for different purposes.

Since this is a comparative study of the discourse structure of parallel genres in a specific discipline or profession, it is essential that an approach which makes the comparison both feasible and illuminating is adopted. Three related approaches to the analysis of discourse organization are adopted in this study. First, texts in the corpus are analysed for the hierarchical organization of information into schematic structures, referred to in this study as **Moves** (Chapter 4). A Move is conceived in this study as a text segment made up of a bundle of linguistic features (Lexical meanings, propositional meanings, illocutionary forces, etc) which give the segment a uniform

orientation and signal the context of discourse in it. Second, texts in the genres are analysed for the organization of utterance "Themes" into identifiable progression patterns (Chapter 5) in line with Danes' (1970, 1974) theory of **Thematic Progression** (TP). Finally, sentences in texts in the genres are examined for the semantic and grammatical choices which writers make in using sentences to construct text, in what Halliday and Hasan (1976) refer to as **Cohesion** (Chapter 6).

To sum up the discussion in this section, the purpose of this study is to compare three parallel genres of written medical discourse with a view to ascertaining how they differ in terms of the schematic organization of information, the thematic organization (Thematic Progression) of information and the patterns of textual cohesion. The aim is to show how discourses can be the same in terms of subject-matter but different in that they must incorporate information from their context in different degrees, some incorporating less, others incorporating more; that is, to explain how the "same" discourse subject-matter may vary when presented under different pragmatic conditions. Thus, the goal of this research project is to show the extent to which these three genres are the same and how they are different, to identify and show how the differences are necessary, given the pragmatic conditions of their production.

Attention is often paid to lexical and syntactic changes when texts are being rewritten for different audiences and for different purposes. For instance, in its manual of "Style for Contributors" (See Appendix), the *New Scientist* enjoins its authors to "avoid passives" because "They are ambiguous and they can make an article sound cautious and flat". It also informs them that "adjectival nouns stacked in front of a noun sound ugly", adding that "a reader disentangling a long list (of adjectival nouns) is doing unnecessary work". However, little attention is paid to discourse structure in making these recommended alterations. It should, therefore, be interesting to examine how differences in discourse structure in the three parallel genres is

considerations of audience and purpose of discourse despite the fact that writers may not necessarily pay conscious attention to this as they write.

1.1 Need for the Study

The problem of discourse organization in scientific discourse has been the subject of investigation in many areas of linguistic research, such as discourse analysis, composition theory, reading research, language teaching for specific purposes, automatic text processing and documentation. Most of the studies conducted in these areas have tended to emphasize the practical application of linguistic theories to the processing of scientific information. The aim, for most of them, has been to facilitate the teaching, to students of Science and Technology, of the art of reading and writing texts in their subject areas.

Unlike most of these studies, the immediate concern of this study is primarily theoretical. To a large extent, this study is a theoretical exercise in Text Linguistics and Applied Discourse Analysis. Therefore, its foremost concern is to contribute to these related fields of linguistic inquiry. The need to conduct theoretical research into scientific and technical discourse has been justified (Meyer, 1987; Anderson et al, 1983). According to Meyer (op cit:8),

practical applications of text-linguistic approaches in the fields of, for example, languages for specific purposes or machine aided document description, can still profit from deeper theoretical insight into the problems of technical discourse or discourse in general since the models applied so far have been rather crude and unsophisticated, given the vast complexity of the structure of any discourse, even the most simple and stereotyped one.

In the same vein, Anderson, Brockmann and Miller (op cit:10) contend that "scholarly researches can contribute to scientific and technical communication by building further theory, supporting and correcting teaching strategies and curriculum design and informing the practices and problem-solving strategies of the communication professional". In line with these views, this study is of relevance to the teaching of

English not only to students of Science, Technology, Technical Communication and Media Communication studies, but also to those interested in machine aided document description. However, for operational reasons, such as the need to make the findings pertinent to specific language users and learners, this study examines genres of written medical discourse selected to represent parallel professional and journalistic popularized medical writings.

Professional research papers, in this case, Research Articles and their accompanying Abstracts are selected for this study because they are representative of written scientific and technical discourse. As Meyer (op cit: 8) points out:

Technical discourse is not just one type of discourse among many others. It is the prototype of coherent discourse; that is our notions of coherence have been shaped on the model of technical, namely expository, discourse.

Furthermore, professional research papers are selected for analysis in this study because they are a primary source of knowledge in the academic scientific world. Therefore, they are an important source of knowledge to both specialists and learners in the field. A study which seeks to characterize the linguistic properties of professional scientific, and/or technical research papers will, no doubt, be of interest to scientists (professionals and learners) who are constantly engaged in either writing or reading them. It will also interest non-scientists, especially linguists who are in the business of teaching English to Science and Technology students.

More importantly, this study examines professional and popularized science (medical) texts because of the need to fill the gap already mentioned in 1.0 above. That is, the lack of a sufficient number of studies on science popularization in general and on the comparison of professional and popularized science texts in particular. This need is underlined by an editorial in an issue of the influential UNESCO journal *IMPACT of Science on Society* devoted to science popularization, in which the editor states as follows:

This issue deals with an activity that is increasingly carried out but rarely written about; science popularization. Although it has a long history (see Jack Meadow's article, p.341), few textbooks exist to instruct those who wish to popularize science, and the subject is infrequently explored in print, outside of the publication of organizations directly devoted to it. This issue of *IMPACT* may well be the first of any journals predominantly devoted to the theme from an international point of view.

(IMPACT of Science on Society, No.114, 337)

Reference has been made in Section 1.0 above to a few discourse-based studies which have sought to compare professional and popularized science texts. Of those, Myers (1986, 1989) remain the most outstanding and the more frequently cited. As we shall see in Chapter 2 on the review of literature, Myers (1986) studied professional and popularized articles by the same authors and on the same topics. He made a number of interesting observations regarding the way in which the same writer, writing on the same topic for different audiences is constrained to produce two difference narratives — the narrative of science and the narrative of nature. Concluding his study, Myers (op cit: 38) suggests that:

Further studies would be needed to show whether the distinction between the narrative of science and the narrative of nature applies to popularizations written by science journalists in general interest magazines and newspapers.

It is primarily in response to this need that the present study examines **Journalistic Reported Versions (JRV)** of research articles published in popular science magazines and newspapers and compares them with their equivalent research articles and their accompanying abstracts.

Just as important as the needs discussed in the preceding paragraphs is the contribution which this study can make to the newly emerging area of Genre-Analysis. As has been pointed out (Section 1.0), each of the "text-types" examined in this study

is considered a genre because each of them is a standardized communicative event with a known public purpose. Genre-analysis is the only approach capable of revealing to us much more clearly something of the organization of discourse in genres. But, since its inception, genre-analysis has tended to concentrate almost exclusively on the characterization of the linguistic properties of single genres. This tendency is attested to by the numerous genre-based studies which focus on individual genres (Swales, 1981a; Dudley-Evans, 1986; Pindi, 1988; Juigfu, 1987; Crooks, 1984; Anderson and Dudley-Evans, 1988). It is, however, the contention in this study that in addition to claiming to say something about individual genres, genre-analysis can also say something about the relationships which exist between one individual genre and another with which it shares some characteristic features. This contention is based on the belief that it is by studying related or parallel genres that we can truly begin to understand what the term "genre" is meant to cover.

One way in which this study attempts to make this understanding possible is by adopting a comparative approach to the problem of discourse organization in genres. The other is by adopting an eclectic or integrated approach to the analysis of discourse organization in genres. The need for an eclectic approach to the analysis of text structure has been pointed out (Meyer and Rice, 1985). According to Meyer and Rice (op cit:342):

It is possible to describe a passage from the point of view of the content it presents, from purely structural standpoint (as in story grammars), in terms of cohesive relations which tie sentences together, or from the point of view of the emphasis patterns being used by the author to underscore his/her own interpretation of the relative importance of the ideas expressed. At present, no single system attempts to combine all these possible organizations.

Therefore, there is the need to extend the horizons of genre-analysis beyond its present pre-occupation. It is, perhaps, in recognition of this need that Dudley-

Evans (1987:5) contends that, "an approach to genre-analysis that combined Hoey's work on clause relations with an analysis of MOVES as proposed by Swales provides a very useful insight into the workings of an individual text". This study is an attempt to provide this very useful insight into the workings of, not just an individual genre, but of parallel genres, through an examination of the relationships and differences in the structure of discourse at three levels of organization: schematic, thematic and cohesive.

The levels of organization examined in this study correspond with those identified by Hartman (1980, 1987) in his theory of contrastive textology, viz:

- Text pragmatics (or communicative textology)
- Text syntax (or combinational textology)
- Text semantics (or referential textology)

Therefore, this study perceives the need for intensive research into what may be termed "Comparative Genre-Analysis" along the same lines as those proposed by Gleason (1968) for comparative discourse analysis. This proposition is based on the same observations as those which prompted Gleason (op cit:58) to state that:

We now have a framework that provides a better starting point than any we have had before for systematic contrastive work. It allows us to focus on what may well prove to be the most interesting of all contrastive problems, the differences in the way that organization is signalled to the hearer or the reader.

As we have pointed out in this discussion, the differences in discourse organization examined in this study have significant relevance to the writing, reading, and documentation of scientific and technical information. They also contribute to a better understanding of what the term "genre" is meant to cover.

1.2 STATEMENT OF THE PROBLEM

This study is related to recent attempts in discourse analysis to characterize discourse by reference to its users in what has been termed DISCOURSE COMMUNITIES (Swales, 1986; Freed and Broadhead, 1986; Cooper, 1986 etc). Furthermore, it relates to sociological investigations into the transfer of science information from a professional to a popular medium and from a specialist to a less-specialist audience (Karpf, 1988; Jones et al, 1978; Taylor, 1975). It also identifies very strongly with the discourse-based investigations into the construction of popular science (Myers, 1986, 1987, 1988; Adams-Smith, 1986, 1987; Dubois, 1985 etc). As with most of these studies, a lot of insight is drawn from studies on the sociology of science (Bazerman, 1983, 1988; Knorr-Certina, 1981; Gilbert and Mulkay, 1981; Knorr-Certina and Mulkay, 1983 etc). However, unlike most of these studies, this study is concerned with the linguistic problems associated with the construction of professional and popular science texts based on the concept of "genre" and genre-analysis.

The linguistic problems which can be investigated in a discourse-based analysis of the construction of professional and popular science texts may vary depending on the purpose for which the study is conducted. In line with the purpose stated in Section 1.0 above, this study seeks to:

- Examine the extent to which the three text-types can be characterized as distinct genres despite the fact that they report on the same topic or subject-matter in the same field or discipline. That is, to account for the fact that though the genres report on the same subject matter, they vary in terms of the amount and pattern of organization of the information they contain.

- Establish that the difference in the pattern of discourse organization across the three genres is due to the pragmatic conditions of their production. That is, the influence of purpose, audience and medium of discourse and the assumptions of the writer about the reader.

In order that these problems can be put in perspective, answers are sought to the following questions:

- Do the three text-types (Research Article Proper, Abstract, Journalistic Reported Version) belong to the same genre or are they three distinct genres? In posing this question, cognisance is taken of the fact that a genre is a structured and standardized communicated event with constraints on allowable contributions in terms of their positioning, form and intent (Swales, 1985). The question also takes cognisance of the fact that a genre may have a number of sub-genres. It also endorses the view that "genre-analysis" may be used as a classificatory system, revealing the essential differences both between the genre studied and other genres and also between the various sub-genres (Dudley-Evans, 1987: 2).
- Assuming that they are three distinct genres, are there commonalities and differences in the patterns of discourse organization of utterances and sentences (Thematic Progression and Cohesion) in the genres? This question is meant to focus on an aspect of the problem which is concerned with determining whether it is also possible to differentiate between related genres in terms of Thematic Progression and Cohesion.
- Are the similarities and differences random or systematic?

Assuming that the similarities and differences are systematic, can they be adequately accounted for in terms of the assumptions which writers have about their audience based on the level of knowledge which they share with that audience?

The questions posed above are based on the following general assumptions:

- That the form and content of parallel genres (Research Article Proper, Abstract, Journalistic Reported Version) are most effectively cross-compared by a systematic examination of the discourse organizing features which are known to be common to all texts. This assumption is based on the understanding that texts constitute genres which reflect identifiable units of information known as MOVES; that the Moves in turn govern the unfolding of the text into a hierarchy of schematic structures which may embody a progression of themes and of cohesive texture.
- That a full interpretation of the differences between parallel genres can only be achieved when an attempt is made to compare the form or structure which the genres reveal, especially within the context and situation in which each genre is realised. This assumption is based on the understanding that "discourse is the totality of undifferentiated language within a context that subsumes the totality of pragmatic/semantic and communicative variables, a structure that is a hierarchy of texts and a structure that subsumes patterns of cohesion" (Hatim, 1981: 36). In line with Hatim (op cit), the contention is made in this study that a true picture of the differences between parallel genres can only emerge when the structure of the mass of seemingly undifferentiated language is revealed; that the process of revealing this structure is enhanced by reference to the context and situation of discourse in the genres. It is the context and the situation alongside the linguistic clues

that enable the identification of information units referred to in this study as MOVES.

1.3 BASIC HYPOTHESES

In line with the question raised in 1.2 and the assumptions stated in 1.2, this study seeks to test, in qualitative and descriptive terms, the following basic hypotheses:

- That texts of a given genre (eg. RAP) will display striking similarities in the pattern of organization of schematic structures, thematic progression and cohesion that will set them off as different from texts from a parallel genre (eg. JRV, ABS).
- That the differences between the genres is directly related to the context and situation of discourse in each of the genres.

1.4 DESIGN OF THE STUDY

Since this is a comparative study, it is necessary to explain the format in which the discussion in the core chapters is presented. Essentially, the approach adopted in reporting the results embodied in the core chapters is qualitative and discursal. Therefore, discussions of results are accompanied by exemplifications or extracts from portions of texts to which the results relate. In addition, results are authenticated by reference to relevant literature in the fields of Text Linguistics and Applied Discourse Analysis.

The thesis is divided into chapters which in turn are divided into sections. Each chapter can be said to be complete in itself, and therefore, can be read as an independent unit which contributes to the overall purpose of discourse in the thesis.

In general terms, the core chapters or the result chapters, as they are sometimes known, consist of:

- an introduction section
- a section on the review of the concept being discussed in the chapter
- a section on the review of other approaches to the concept
- a section on the specific methodological approach adopted in the chapter
- a section on the analysis of data
- a section on the comparison of results.

Since this is a comparative study, there is the inevitable repetition of results and examples which is noticeable between the sections on the analysis of data and that in the comparison of results. Again, since the core chapters examine the same problem - discourse structure - albeit from different perspectives and at different levels of organization, it is possible for results to overlap across the three chapters. Effort has been made to limit the chances of this happening. To this end, the discussion in each chapter is limited to the most important causes of difference which can be identified as being peculiar to the particular model of discourse organization which a chapter examines.

Finally, the linguistic concepts which this study adopts to explain differences in discourse structure across the three genres are functional categories. Therefore, they are not as discrete as syntactic or grammatical ones. To a large extent, the concepts were originally developed from highly restricted and in some cases, made-up examples. Therefore, they can be fuzzy and slippery when applied to different texts

and to an extensive body of data. This notwithstanding, they have been successfully applied to varieties of discourse and to large bodies of naturally occurring language data. Nevertheless, they still remain delicate concepts which require careful handling and application. In recognition of this fact, special emphasis is placed in this thesis on an explanation of the principles and methods used in the identification of concepts which are applied in each core chapter. A lot more emphasis has been placed on this issue in the chapters on Schematic Structure and Thematic Progression both of which are capable of generating arguments over the appropriacy of MOVE boundaries and the extent to which thematic information can be deemed recoverable from preceding context. Despite this fact, it is impossible to completely eliminate areas of disagreement in discussions based on concepts such as those handled in this thesis, though they can be minimised.

1.5 OUTLINE OF CHAPTER CONTENT

Chapter Two is an attempt to review aspects of the literature which have general relevance to the topic of this thesis. In general terms, the discussion in the chapter is limited to a review of three main areas of language use: English for Science and Technology (EST), English for Medical Purposes (EMP) and Science Popularization. In specific terms, Chapter Two begins with a review of approaches to varieties differentiation in English for Science and Technology (EST). Three approaches - Register Analysis (Halliday et al, 1964); Grammatical-Rhetorical Analysis (Selinker et al, 1972); Genre-Analysis (Swales, 1981a) - are reviewed in association with studies which have attempted to apply them to the analysis of scientific texts. The extent of the relevance of each of these approaches to the corpus of this study is also discussed. Following a review of these approaches is a discussion of researches on written medical texts which have relevance for the kind of problems discussed in this thesis. Selected for review are some genre-based analyses

of aspects or sections of medical research papers as well as studies which attempt the characterization of the linguistic features of medical texts. The aim is to indicate the general trend of research in the analysis of medical research papers, discuss the limitations of such studies and indicate how this study fits into, as well as contributes to studies in this field of linguistic research. Finally, the chapter ends with a review of studies on science popularization starting with a historical review of the origin of science popularization through brief comments on the state of science popularization in developing countries to an account of a number of sociological studies on science popularization. The section ends with an extensive discussion of some sociolinguistic and discourse-based studies on science popularization. Considerable attention is paid to Myers (1986) account of the "social construction of popular science". Further reviews of literature which are considered important for putting the discussion in the core chapters in perspective are also provided in Chapters Four, Five and Six.

Chapter Three is an account of the general issues underlying the methods adopted in the investigation. In view of the seemingly independent nature of the concepts examined in the core chapters, it was not considered methodologically sound to concentrate discussions on all issues of research methodology in a single chapter. To do so would mean having to constantly refer back to the methodology chapter and losing touch in the process with the thread of discussion, particularly on issues requiring an understanding of principles and approaches on which some assumptions and contentions are made. Therefore, Chapter Three spells out only aspects of the method of research which relate in a general way to the purpose of this thesis. To this end, the chapter is an attempt to highlight the conceptual framework upon which the data was selected, delimited and analysed. Thus, it concentrates on the following issues:

- sources of texts in the corpus

- criteria for the selection of texts
- basic methodological principles
- methodological approaches to the analysis of the organizational features of texts.

More specific issues of research methodology relevant to the needs of the core chapters are discussed in those chapters.

Chapter Four reports on the analysis of variation in schematic structures across the three genres. It begins with a brief introduction which provides background information on what the chapter is about. This is followed by a brief account of the concept of schema and an indication of what the concept is taken to mean in the context of discourse in the thesis. The chapter also contains extensive discussions on the principles upon which schemata are identified in the texts which make up the corpus. Having discussed the principles for the identification of schematic units, the description of canonical schematic units or MOVES follows. The description of schematic structures requires the extraction of large portions of text from the corpus into the thesis in order that effective discussion of the linguistic properties of the schematic units can be carried out. This requirement means that the chapter on schematic structures may be very lengthy indeed. But, since I am working within given guidelines with regards to the length of thesis and in order to cut down on incidences of repetition which may occur from reporting on the analysis of schematic structures in each genre, I have had to limit the detailed account of the analysis of moves in the genres to those realized in the RAP. These are regarded as canonical moves from which those in the ABS and the JRV are derived. However, a brief description and outline of Moves in the ABS and JRV is also presented. The chapter ends with a discussion of the differences which exist between the genres in their patterns of schematic structures. The differences are discussed under three main

headings: differences in the realization of Moves, differences in the pattern of organization of Moves, and differences in the linguistic features by which Moves are realized.

Chapter Five, is an account of differences which exist between the genres in terms of their patterns of Thematic Progression (TP). The realization and functions of Danes' (1970, 1974) Thematic Progression Patterns (simple linear, constant, derived or hypertheme and split-rheme TP) are examined in texts across the three genres. The chapter begins with a brief description of the concept of theme as a means of pointing out why the concept is intuitively appealing as an approach to the characterization of discourse organization in text. This is followed by a review of two traditions to the characterization of theme in discourse - Halliday's Firthian tradition and the Functional Sentence Perspective (FSP) tradition of the Prague School of Linguistics. A review of FSP conceptions of theme is provided, with some emphasis on Firbas' theory of Communicative Dynamism (CD). However, a lot more emphasis is placed on Danes' theory of Thematic Progression. Some relevant FSP-based studies are also reviewed. Following this is an account of the criteria used in the identification of theme in the corpus. The criteria adopted are: identical wording, synonymous expression, semantic inference and paraphrase. Themes identified in utterances in the corpus are classified into: simple, multiple and complex themes. The approaches adopted in the characterization of theme in what has been termed "exceptional utterances" in this study - existential, extraposition and interrogative utterances - are also discussed. Finally, the chapter ends with a discussion of cross-genre differences of TP patterns in texts in the corpus. The genres are compared according to the following headings:

- The realization of TP patterns across genres
- The distribution of TP patterns
- The characteristic elements which realize theme across genres.

Chapter Six examines differences which exist between the three genres in terms of textual cohesion. The model of cohesion adopted in the investigation is that of Halliday and Hasan (1976). As in Chapters Four and Five, Chapter Six begins with an account of the concept underlying the study - the concept of cohesion. Cohesion is characterized in relation to associated concepts such as text, texture and genre. A brief account of other approaches to the characterization of cohesion in discourse is provided as a means of pointing out how they relate to or differ from Halliday and Hasan's model. The approaches reviewed are: Grimes (1975), Longacre and Levinsohn (1977), de Beaugrande and Dressler (1981) and Gutwinski (1976). This is followed by a detailed account of Halliday and Hasan's model of Cohesion in English. For the purpose of this study, the model is modified in accordance with insights derived from Halliday (1985a) and Hasan (1984). After this, the approach adopted to the analysis of cohesion in sample texts from the corpus is presented. Text extracts from each of the genres and tables indicating the analysis carried out for each text are also presented. These are meant to serve as examples of the approach adopted in the analysis of all texts in the corpus. Finally, the chapter ends with a comparison of differences which exist between the genres in the cohesion types they characteristically realize. The discussion is organized according to the following headings:

- Differences in reference cohesion
- Differences in conjunction cohesion
- Differences in lexical cohesion.

Chapter Seven is in four parts. The first part is a summary of the main differences observed in the analysis of Schematic Structures, Thematic Progression and Cohesion in the genres. The second part is an account of five main conclusions

arrived at in the study based on the method of analysis adopted in the study and the results obtained. In the third part, the theoretical and practical implications of the results obtained in the study for discourse organization, varieties differentiation and genre-analysis in EST discourse are discussed. Finally, the fourth and final part points out the limitations of the study and identifies areas for future research.

It was pointed out that the study is concerned with describing related or parallel genres of written medical texts in terms of their patterns of discourse organization. The study structure, thematic progression and cohesion. As a comparative study of parallel genres, the study is an attempt to establish the extent of differences which exist between these related genres.

In this chapter, an attempt will be made to further explicate the concepts mentioned in chapter 1 in order to determine how this study fits into current trends of research in the field of discourse analysis with particular reference to the analysis of scientific texts. To this end, an overview in this chapter will focus on the concept of functional variation as applied to English for Science and Technology (EST) discourse; the concept of genre analysis and approaches to the analysis of medical research reports, scientific research abstracts and journalised reported versions or popularized scientific research papers.

APPROACHES TO FUNCTIONAL VARIATION IN DISCOURSE

In sociolinguistics and functional descriptions of language, the term 'variety' tends to be associated with the characterization of functional varieties of language in particular contexts such as 'region, social group, field of discourse, institutional activity' (Duck et al. 1983: 34). One of the most notable notions developed in sociolinguistics for the description of functional varieties in language is the term 'register'. Although one of the earliest known uses of the term is Hall (1936) who used it to characterize changes in the language behaviour of a single individual,

... the term is also used to refer to the language behaviour of an individual in any

CHAPTER TWO

A GENERAL REVIEW OF RELATED LITERATURE

2.0 INTRODUCTION

In Chapter 1, a number of principles and concepts underlying this study were introduced. It was pointed out that the study is concerned with comparing related or parallel genres of written medical texts in terms of their patterns of discourse organization - Schematic Structures, Thematic Progression and Cohesion. As a comparative study of parallel genres, the study is an attempt to establish the extent of variation which exists between these related genres.

In this chapter, an attempt will be made to further explicate the concepts introduced in Chapter 1 in order to determine how this study fits into current trends of research in the areas of discourse and genre analysis with particular reference to the analysis of scientific texts. To this end, the discussion in this chapter will focus on the concept of variation as it has been applied to English for Science and Technology (EST) discourse; the structure and organization of information in EST discourse; the concept of genre and genre-analysis and approaches to the analysis of medical research reports, scientific research abstracts and journalistic reported versions or popularized scientific research papers.

2.1 APPROACHES TO FUNCTIONAL VARIATION IN DISCOURSE

In sociolinguistics and functional descriptions of language, the term 'variation' tends to be associated with the characterization of functional varieties of language in use based on variables such as "region, social group, field of discourse, medium, attitude" (Quirk et al, 1985: 16). One of the most notable notions developed to account for the existence of functional varieties in language is the term **register**. Amongst one of the earliest known uses of the term is Reid (1956) who used it to characterize changes in the language behaviour of a single individual. Reid (op cit) points out that "the linguistic behaviour of an individual is by no

means uniform", contending that placed in what appear to be linguistically identical conditions, the individual will use a number of distinct "registers". Following from Reid (op cit), the concept has been widely used in linguistic descriptions of language varieties classified in terms of mode, genre and style (Hill, 1958); role, style and medium (Catford, 1965); field of discourse, mode of discourse and style of discourse (Halliday, McIntosh and Strevens, 1964); field of discourse, mode of discourse and tenor of discourse (Halliday, 1979, 1985a, 1985b; Gregory and Carroll, 1978). However, the most well-known characterization of the concept of register is that by Halliday et al (1964). In their seminal work on the concept, Halliday et al contend that:

Language varies as its function varies: it differs in different situations. The name given to a variety of a language distinguished according to use is 'register'.

(Halliday et al, 1964: 87-89)

Elsewhere, Halliday (1985b: 38) defines register as a configuration of meanings that are typically associated with a particular situational configuration of field, mode and tenor which also expresses "the lexico-grammatical and phonological features that typically accompany or REALISE these meanings". Explaining, Halliday (op cit: 43) states that:

Register is "what you are speaking (at the time)", ie. determined by what you are doing (nature of activity in which language is functioning).

adding that:

Register reflects social order in the sense of social process (types of social activity).

Halliday (1985b) refers to variation based on the register characteristic of language identified in the definitions above as **functional variation** - "the systematic variation of language according to what we are doing - what we are using language to achieve".

Halliday's notion of "register" is also related to what Stubbs (1986: 18) refers to as "diatypic variation". Stubbs defines a diatype as "a variety defined according

to use."

Explaining the concept of a diatype, Stubbs (op cit: 19) states that:

Although, generally speaking, individual speakers have native command of only one dialect, they have command of many different diatypes - that is, they change their language according to who they are talking to, what they are talking about, where they are, and so on.

Thus, to a great extent, Stubb's concept of a diatype is related to Halliday et al's concept of register and his conception of "diatype variation" similar to Halliday (1985b) conception of "functional variation".

Having briefly outlined the basic principles underlying the concept of register and functional variation, we shall attempt in the next section to examine the extent to which these and other approaches have influenced research into the analysis of scientific/medical texts.

2.2 FUNCTIONAL VARIATION AND THE ANALYSIS OF SCIENTIFIC/MEDICAL WRITING

As the brief outline of attempts at varieties differentiation presented in 2.1 above shows, language varies very distinctly based on usage and the communicative purpose to which it is put. Commenting on this fact, Ellis and Ure (1979: 251) state as follows:

Linguistic (and literary) scholarship has always recognised, without exhaustively systematizing, variation in the usage of linguistic items. This awareness is reflected in such terms as 'literary language', 'elevated language', 'colloquial language' etc.

The idea that there exists such things as "literary language", "elevated language", "colloquial language" etc, which is reflected in the way linguistic items are used, probably influenced linguists, especially those working in the area of English for Specific Purposes to characterize the linguistic properties of "specialized" languages used in different contexts and for different purposes. In no area of specialized language use has this influence been more markedly demonstrated than in science and

technology, thus giving rise to the most influential branch of ESP known as English for Science and Technology (EST). The interest which this influence generated has also meant that a critical look has had to be taken at the extent to which the concepts underlying earlier attempts at varieties differentiation have been adequate for characterizing variation in specialized use of language, such as in scientific/medical discourse.

Three broad approaches to the description of variation in scientific/medical discourse may be identified. The first approach represents the earliest attempts to characterize the language of what is generally referred to as "scientific texts" (Barber, 1962; Huddleston, 1971; Gopnik, 1972). The theoretical construct upon which this approach is based is the concept of "language variety" which has been identified in this study as "register". The second approach is represented by the works of Selinker, Trimble and Lackstrom et al (1974, 1978, 1979, 1981) which they refer to as the **grammatical-rhetorical** approach. It draws heavily from the classical theories of rhetorics as well as the theories of composition teaching as practised in the United States. The third and more recent approach is represented by the works of Swales (1981a, 1981b, 1985), Dudley-Evans (1986, 1987), Adams-Smith (1987), Jacoby (1987) etc. The theoretical framework upon which this approach is based is the concept of "genre" (Swales, 1981a; Brown and Yule, 1983).

2.2.1 The Description of Scientific Texts as Register

Before the development of the concept of register, there appeared not to have existed any suitable notion for accounting for the kind of texts studied in ESP/EST situations. So, the notion of register, especially as put forward by Halliday et al (1964), Halliday (1979) provided a theoretical basis for initial attempts at the characterization of scientific language as a distinct form of language in use.

Halliday (1979: 35) identified extreme cases of register as "restricted languages and languages for specific purposes" and "an excellent example of register variation" as Jean Ure's "lexical density and register differentiation", thereby creating

the impression that register analysis is essentially an exercise in isolating the lexicogrammatical properties of language varieties. Widdowson (1979: 22) makes a similar observation when he points out that:

The argument upon which the notion of register is based rests on a double fallacy. On the one hand the fact of language variation is thought to entail the existence of separate language varieties within a language and on the other hand functional variation is thought to entail the existence of these varieties as formally distinct subcodes. It seems to me that there is a confusion here between language and a language and between form and function.

In place of Halliday et al's characterization of language varieties on the basis of formal linguistic properties, Widdowson (1979: 23) suggests a shift of theoretical orientation from linguistic form to communicative functions.

However, despite these observations, a large number of the earliest descriptions of the language of science seem to have been concerned with the characterization of the formal properties of "scientific language". One of such studies is Barber (1962). Barber examined three scientific texts in the fields of Electronic Engineering, Biochemistry and Astronomical Instrumentation and arrived at the conclusions (among others) that out of the 264 subordinate clauses in his data, 102 were adjective clauses out of which 52 were defining and 50 non-defining; that the commonest relative pronouns used to open the relative clauses is '*which*': that of the total number of verb forms in the texts, 61% were finite verbs and 39% non-finite verbs. Finally, he found that in order of ranking, the tenses in his material might be ordered according to: (1) present simple active (64%), (2) present simple passive (25%), (3) future simple active (3.7%), present perfect passive (1.7%) etc.

Barber's pioneering work has been commended for being a clear demonstration that the descriptive technique of modern linguistics, as most influentially represented in the linguistic science by Halliday, McIntosh and Stevens (1964) could be successfully applied to the language of science and technology" (Swales, 1985: 1). Its statistical details also provide evidence which, according to Swales (op cit) suggests that "Scientific English" is different from

"General English" or "Literary English". However, this last contention is highly debatable, especially since it is not based on data from non-scientific English. Besides it is not clear what Swales means by "General English".

Despite these commendations, Barber's work lacks any serious reasoned explanation of why these structural features occurred at the rate they did. As Swales (op cit: 9) points out: "A simple statement of the frequencies of Active and Passive can be misleading, especially if the percentage of passive can be taken as an indication of the degree of impersonality ..." The real question, Swales contends, is "what the textbook writer does when he has a choice of either active or passive, and why he chooses one or the other". This is related to a similar observation by Widdowson (1979: 55). According to Widdowson (op cit):

The fact that scientific English text exhibits a relatively high proportion of certain syntactic features and a relatively low proportion of others may be useful for identifying scientific English texts should we ever wish to do so. But it cannot reveal the communicative character of what is written. It cannot of its nature deal with discourse.

In addition to these criticisms, Barber's (op cit) work suffers from the common problem associated with most register-oriented studies - ie. the tendency to oversimplify and overgeneralize on issues of varieties differentiation on the basis of limited samples and a highly contentious methodological procedure. In other words, it sounds too ambitious, in view of the present level of knowledge in discourse studies, to make generalizations concerning "scientific prose" on the basis of three text-types in only three fields of science. Swales (op cit: 4) underlines this contention when he states that not many people today would consider Electronic Engineering, Biochemistry and Astronomical Instrumentation to be interdisciplinary fields.

Another significant study of the language of science which is heavily influenced by Halliday's et al (op cit) conception of register is White (1974). According to White (op cit: 402):

If a successful and enlightening attempt is to be made to discover distinctive characteristics of registers, more attention should be given to studying the characteristics of language determined by

the major contextual dimensions which influence linguistic form, viz, Field, Mode and Tenor, which Crystal and Davy have supplemented by such additional dimensions of constraint as individuality, singularity and province.

This is characteristic of the trend that dominated linguistic research, especially in the area of ESP/EST for many years. As in most of these studies, White (op cit) set out to examine the frequency of occurrence of grammatical features, specifically verb phrase and non phrase forms in "Scientific English". He found that laboratory reports tend to make more use of passive forms than either science textbooks or general English textbooks.

White's (op cit) work shares a number of shortcomings with Barber's (op cit) and others (Wingard, 1981; Heslot, 1982; Hanania and Akhtar, 1985, etc) which tend to approach textual variation from the point of view of the statistical counting of selected grammatical features and vocabulary items. Prominent amongst these shortcomings is the non-representativeness of the corpus already referred to in the discussion of Barber (op cit). The materials for analysis were drawn from (1) science and technology textbooks which represent "two kinds of standard language which foreign students wishing to study at a British university would have to cope with," (2) under-graduate laboratory reports, representing "the performance of native speaker students", (3) texts from non-scientific subjects which represent "a general profile of textbook material." The rationale for the selection of the materials appears justifiable, but the selection itself appears indiscriminate and unmotivated because it tends to bring together for comparison disparate fields such as science and literature, which ostensibly display different rhetorical modes. Since the aim of the study is not to show that the text types differ in terms of their "communicative purpose" and since the method of analysis used cannot be said to show sensitivity to variation in terms of the communicative purpose of the text-types, the exercise becomes a sterile investigation into linguistic variation, its main merit lying in its ability to provide a taxonomic description of the register of science.

Secondly, the study's aim of examining grammatical features has been described as a meaningless, futile exercise unless the various aspects being examined

are related to some overall strategy relevant to some vital aspect of the communicative process (Hatim, 1981: 9). As Hatim (op cit) points out:

... assuming that the discourse value 'hypothesizing' has been found to correlate with the construction would, verb and that this construction has been frequently observed to be the textual reflex of 'hypothesizing' it would be revealing indeed to embark on an investigation of the construction of a large sample of related texts, using a number of analytical tools, including frequency counts. But to produce frequency counts in a complete contextual vacuum is utterly misleading.

This study shares the views expressed by Hatim (op cit) - a view which serves to explain the emphasis which this study places on the qualitative descriptive approach to the analysis of discourse organization, "using a number of analytical tools".

2.2.2 The Grammatical-Rhetorical Approach to EST Discourse

The inadequacies of the traditional register analysis model to the description of scientific discourse must have prompted an early search for a more viable alternative. The alternative approach which has received widespread attention from ESP/EST practitioners is the "grammatical-rhetorical" model proposed in Lackstrom, Selinker and Trimble (1972), Selinker, Trimble and Vroman (1974), Lackstrom (1978) and Trimble (1985). According to Trimble (1985: 14) the model is built around three main rhetorical concepts:

- The nature of the EST paragraph.
- The rhetorical techniques most frequently used in written EST discourse
- The rhetorical functions most frequently found in written EST discourse.

The aim of the "grammatical-rhetorical" approach is to show that correct grammatical choices in a written medium cannot be taught apart from consideration of rhetoric and subject-matter. Lackstrom et al (1972) set out to show the following:

- That the choice of tenses in the medium is dependent not on "time lines", as most textbook presentations suppose, but on rhetorical and subject-matter consideration.
- That specific grammatical choices involving definite and indefinite articles depend on rhetorical and subject-matter principles.
- That choices involving adverbs, aspects, agent phrases and normalization often demand contextual directives.

The grammatical-rhetorical approach is based on a framework referred to as the 'rhetorical process chart'. The rhetorical process chart is divided into four levels, A-D, indicating the different levels at which EST rhetoric may be realized. Level 'A' is concerned with indicating the purpose of the total discourse as represented by information contained in the introductory section of most written scientific texts. Level 'B' is made up of those aspects which relate to the method of organization of information in the discourse, usually in the form of headings and sub-headings. Level 'C' consists of specific rhetorical functions which are common within EST discourse. Level 'D' identifies rhetorical techniques most commonly used by writers to present information relating to functions identified in level 'C'.

As far as the analysis of scientific discourse is concerned, the "grammatical-rhetorical" model presents a radically different approach from the "Register Analysis" approach in which "Scientific English" is characterized only in terms of its distinctive grammatical and vocabulary features. Swales (op cit: 59) points out that the work by Lackstrom et al (op cit) opened the door to the rhetorical explanation of scientific discourse, while also half-closing the door of statistical or surface grammatical features as the basis for characterizing EST discourse by showing that "the frequency work of the type undertaken by Barber and others could have descriptive validity but little explanatory force".

Despite the obvious advantages which the grammatical-rhetorical approach has over the "register analysis" approach, the "Rhetorical-Process Chart" on which the model is based has some shortcomings. One is the restricted nature of the data used in the study. The "Rhetorical-Process Chart" was developed from a study based on the analysis of scientific textbooks in engineering and on "supplementary reading" materials. It is not clear what texts are subsumed under the category "supplementary materials". They are supposedly textbook related materials. Scientific textbooks are only one of the many sources of the language of science and technology which, as Stevens (1973) has shown, are of different types. However, the data on which the G-R model is based can be said to be representative of the language of science/technical textbooks. The extent to which the categories which it embodies can be said to apply to all forms of science/technical text is subject to further verification.

Another limitation of the Grammatical-Rhetorical model is its tendency to emphasize the learning of one EST skill over all others. As Selinker, Trimble and Trimble (1978: 37) point out, the model is essentially designed to tackle problems of reading amongst non-native speakers learning to cope with the language of Science and Technology. Selinker et al (op cit) paraphrase the applied hypothesis underlying the series of researches which culminated in the G-R model thus:

Errors in understanding implicit information by the non-native speaker attempting to read EST are often related to a lack of understanding of the relationship between individual clauses consisting of supporting information, and between those clauses and core generalizations.

Thus, it would appear that the model proposed by Selinker et al is not designed for the teaching of writing skills. Its combined use of rhetorical functions (definition etc) and rhetorical techniques (time order etc) as approaches to the study of EST texts while being useful for teaching reading comprehension may not account adequately for the way information in EST texts is organized and the perspectives and assumptions of the writer vis-a-vis his audience and purpose.

Despite these limitations, a number of studies have indicated the strength of the grammatical-rhetorical model in explaining interaction in written EST discourse. For example West (1980) has shown that in the biological sciences, the density of "that-nominals" differed significantly between rhetorical sections of the text. In his study West (op cit) examined 15 biological sciences articles from "The Proceedings of the National Academy of Science of the United States of America (76) 1979". He bases the rhetorical divisions in the articles on the conventional sections of a research paper - Introduction, Materials/Methods, Results, Discussion. West also made use of quantificational methods, T-unit per rhetorical section, analysis of variance etc. He contends as follows:

Besides supporting the Lackstrom, Selinker and Trimble claim that "grammatical-rhetorical dependencies (1970: 130) exist in scientific and technical writing, this study also provides evidence for a hierarchy of that-nominal densities in the four rhetorical divisions of research papers. The Introduction section has more that-nominal constructions than the Results section; the Introduction section has approximately the same number of that-nominals constructions as the Discussion section; the Methods section has fewer that-nominal constructions than any of the other three sections.

(West, 1980: 486)

The work by West (op cit) suffers mainly from an internal contradiction arising from its methodological approach. While claiming inter alia to be conducting a grammatical-rhetorical study of biological science articles, the study ends up looking like a classical register analysis. It is questionable whether the study is strictly in line with Lackstrom et al cited in West (op cit: 486) above. First, its association of rhetorical categories with the conventional (IMRD) sections of the Research Article is neither consistent with Lackstrom et al (op cit), nor with modern views about the rhetorical/functional organisation of information in Research Articles (Swales, 1981a). Secondly, its results which are based on the statistical quantification of that-nominals across sections of the Research Article does not reveal much about the contextual and situational influences which constrain the occurrence of this construction across the various sections.

Another related, though more recent study based on the grammatical-rhetorical model of EST discourse is Hanania and Akhtar (1985). The study was aimed at examining "variation in verb use across rhetorical divisions of discourse within a single genre, thesis writing" (Hanania and Akhtar, op cit: 50). It was based on 20 Master of Science theses taken from the area of Biology, Chemistry and Physics. According to Hanania and Akhtar (op cit):

A two way classification was used for the analysis: rhetorical and grammatical. The discourse material in each thesis was first divided into five sections, each representing a broad rhetorical function: Introduction, Review, Methods, Results and Discussion.

As in West (op cit), T-Unit (Hunt, 1970) and clause length were used to measure syntactic complexity and counts of these parameters were made on a reduced sample of about 20% of the whole sample. Based on the quantification of verb forms, Hanania and Akhtar arrived at the conclusion that in overall terms, active verbs exceed passive verbs and that active verbs appear mainly in the simple present tense across the data. In specific terms, they found that the Introduction section makes more use of the present tense; the Review section more of the past; the Methods section more of the past passive, the Results and Discussion sections more of present tense and modals. They contend that these choices are influenced by the communicative function of the rhetorical divisions of the texts. Comparing the results obtained across the three disciplines they found a similarity in the pattern between Biology and Chemistry, especially in the higher use of active verb forms over passive verb forms in the four sections of text in each discipline. Though the physics data showed a similar dominance of active over passive in the four sections, there was a smaller shift (10%) from active to passive in the methods section as opposed to a 40% shift from active to passive in both Biology and Chemistry.

Hanania and Akhtar's work is definitely an improvement on earlier studies of this type (West, 1980; Wingrad, 1981) in that it attempts to combine statistical counting of grammatical features across rhetorical divisions with some form of explanation based on communicative function. For instance, they made the following

observation (Hanania and Akhtar, op cit: 54).

... the choice of the verbs is influenced by the communicative function: the simple present expresses generalizations and conclusions based on the result of research, and modals help to qualify interpretations and conclusions.

Despite this improvement, there is an apparent danger posed by explanations which tend to concede everything to the communicative function of rhetorical divisions as broad as those upon which the study by Hanania and Akhtar is based (ie. Introduction, Review, Methods, Results, Discussion). Such explanations either tend to be too simplistic or highly generalized. One such simplistic and generalized explanation is the statement that: "The use of qualified generalization is a common feature of advanced writing about scientific research" (Hanania and Akhtar, op cit: 54). Questions such as how common these linguistic features are, circumstances under which they occur and why they are used are neither investigated nor explained. In addition, as in West (op cit), the division of the rhetorical structure of the texts into Introduction, Review, Methods, Results and Discussion appears too broad and could have the effect of marginalizing and obscuring the rhetorical function of relatively high but not statistically significant grammatical forms by those considered 'more significant' features. Thus, the results obtained within such broad rhetorical divisions could be misleading, especially in teaching people to read and write texts upon which the study is based. It is, in view of criticisms such as this, that the study reported in this thesis is based on schematic units of text - units which tend to capture the communicative functions of segments of text and the relations which exist between them in the overall organization of information in text.

2.2.3 The Analysis of Scientific/Medical Texts as Genre

2.2.3.1 Genre as Literary Art

Traditionally the word *Genre* is associated with literary works of art. Its development as a literary theory may be traced to the works of Aristotle and Horace who identified two major forms of literary genres as Tragedy and Epic. Of course

Aristotle in particular also recognized fundamental distinctions between categories of genre such as that between drama, epic and lyric (Welleck and Warren, 1973: 227). Following from these, attempts have been made to further characterize literary genres. In modern times, genres have been classified into "Fiction (novel, short story, epic), Drama (whether in prose or in verse), and Poetry (centering on what corresponds to the ancient lyric poetry)" (Wellek and Warren, op cit). The argument, in literary circles, concerning what constitutes a genre appears inconclusive, but in the opinion of Welleck and Warren (op cit: 231):

Genre should be conceived, we think, as a grouping of literary works based, theoretically, upon both outer form (specific metre or structure) and also upon inner form (attitude, tone, purpose - more crudely, subject and audience). The ostensible basis may be one or the other (eg. 'pastoral' and 'satire' for the inner form; dipodic verse and pindaric ode for outer); but the critical problem will then be to find the other dimension, to complete the diagram.

The distinction is also made between classical genre theory and modern genre theory. As Welleck and Warren (op cit) point out, 'classical theory is regulative and prescriptive: it "not only believes that genre differs from genre, in nature and glory, but also that they must be kept apart, not allowed to mix" giving rise to the famous doctrine of the "purity of genre". It was based on aesthetic principles, an appeal to specialization as well as on social differentiation (eg. tragedy and epic deal with the affairs of kings and nobles, comedy with those of the middle class).

Modern genre theory, on the other hand, is descriptive. According to Wellek and Warren (op cit: 235):

It doesn't limit the number of possible kinds and doesn't prescribe rules to authors. It supposes that traditional kinds may be 'mixed' and produce a new kind (like tragi-comedy). It sees that genres can be built up on the basis of inclusiveness or 'richness' as well as that of 'purity' (genre by accretion as well as by reduction). Instead of emphasizing the distinction between kind and kind, it is interested after the romantic emphasis on the uniqueness of each 'original genius' and each work of art - in finding the common denominator of a kind, its shared literary devices and literary purpose.

This distinction between classical and modern genre theory is important for understanding the application of the word **genre** in modern linguistic descriptions in general and in the area of ESP research in particular.

2.2.3.2 "Real-Life" Genres

The concepts underlying modern genre theory (literary) identified by Wellek and Warren (op cit) appear to have influenced the use of the term in linguistics, especially in applied linguistics researches in ESP and EST where it has proved invaluable as a useful concept for characterizing texts. The first known reference to 'genres' in ESP work has been traced to Swales (1981a) (Dudley-Evans, 1987: 1). In his seminal work on Article Introductions, Swales (1981a: 4) states as follows:

I believe it is important to recognize that this investigation is genre-specific. By genre I mean a more or less standardized communicative event with a goal or set of goals mutually understood by the participants in that event and occurring within a functional rather than a social or personal setting.

Echoing Swales (op cit), Dudley-Evans (op cit: 1) defines genre as:

A typified society recognised form that is used in typified society circumstances. It has characteristic features of style and form that are recognised, either overtly or covertly, by those who use the genre.

It would appear from these definitions that there is a slight distinction between the conception of genre in literature and its use in applied linguistics. It is, perhaps, in recognition of this fact that Swales (1985, 1986) attempts to distinguish between genre as art and what he terms "real-life" genres, thus:

- (a) A genre is a recognized communicative event with a shared public purpose and with aims mutually understood by the participants within that event.
- (b) A genre is, within variable degrees of freedom, a structured and standardized communicative event with constraints on allowable contributions in terms of

their positioning, form and content.

- (c) Overt knowledge of the conventions of a genre is likely to be much greater in those who routinely or professionally operate with that genre rather than in those who become involved in it only occasionally.
- (d) Discourse and speech communities give genre names to types of communicative event that they recognise as recurring. Named genres are manifested through spoken or written texts (or both) and their associated text-based tasks.
- (e) Modified genre-names (Survey article, issue memo, panel discussion) indicate features that the community finds salient and thus provide a way into a typology of genres.

It would appear that there is a relationship between Halliday's concept of register and the concept of genre as outlined by Swales in the sense that they are both concerned with providing an account of the functional analysis of language in a communicative setting. However, proponents of both concepts have attempted to show that they are not identical notions. Halliday (1985b: 6) argues that the notion of register as incorporated in ESP "has tended to be interpreted in the rather narrow sense of "subject-matter". While conceding that "the subject-matter is, certainly, one aspect of the 'field' of discourse - it is one component in the nature of the social activity that is taking place", Halliday points out that subject matter 'is only one aspect even of the field; and the field, in turn is only one alongside two or three other factors that together form the background to the discourse". He further contends that one has to take account of the tenor, the social relationships among the participants; and the mode, the specific role that is being allocated to the text".

In his own argument, Swales (1981a: 10) contends that genre analysis differs from "traditional register or sub-register analysis in the importance they attach to communicative purposes within a communicative setting". Echoing Swales' views, Ferguson (1983: 155) criticised register variation models of discourse which provide a taxonomic grid of several major dimensions or parameters as field, mode, participants, tenor etc (Leech, 1966; Gregory, 1967; Halliday, 1968; Ellis and Ure, 1969) as unsatisfactory. He contends that "the parameter approach often fails when applied to particular cases because the situational or functional characteristics that correlate with structural differences in the language may cut across, lie outside or be nested within the proposed parameters". Ferguson advocates the adoption of an approach that tries to "locate" a presumed register by identifying situational or functional features that seem to characterize a recognizable kind of language". This contention is related to a similar one made by Swales (op cit) that "it is only within **genres** that viable correlations between **cognitive, rhetorical and linguistic** features can be established, for it is only within genres that language is sufficiently conventionalized and the range of communication purpose sufficiently narrow for us to hope to establish pedagogically-employable generalizations that will capture certain relationships between form and function".

From the foregoing, it would appear that there is a distinctive difference between the terms "register" and "genre". It is a difference which is based on the scope of linguistic data encompassed by each term as well as the relative emphasis which each attaches to the influence of different contextual features in characterizing discourse types. Based on the highly specific nature of genres, it can be said that the genre-approach has obvious practical advantages over the traditional register analysis approach. As Swales (1987: 11) points out, a genre approach not only makes sense as a way of structuring the narrow world of the classroom, but also reflects a reality beyond that narrow world. If practical relevance is an advantage, it could also be a disadvantage. Since the purpose of genre analysis, especially in ESP work is often pedagogical, there is the danger that it would almost certainly be prescriptive.

However, as Dudley-Evans (1987: 3) points out "its prescription is not in the old fashioned prescribing certain grammatical forms to be the correct forms, but rather more flexible prescription based on analysis that makes suggestions about the layout, ordering and language appropriate to a particular writing or speaking task". Furthermore, since the concept of genre is highly dependent on the communication purpose of texts, it may not account adequately for certain categories of discourse which do not have any identifiable communication purpose such as day-to-day conversations and ordinary narrative accounts. Swales (1985, 1987) appears to have sensed this problem, hence he identifies these two as instances of "non-genre-specific" discourse. Explaining why they are non-genre-specific Swales (1987: 12) states as follows:

It seems to me that these two verbal activities (day-to-day conversation and ordinary narrative accounts) are prototypical or pre-generic in a number of ways: in the probable evolution of human societies, in an individual's acquisition of language, in the sheer amount of time given to chat and story-telling, and in the fact that most human beings are psychologically dependent on such interactions and exchanges. As a result, the interest for the genre-analyst lies in the kinds of 'rules', patterns and procedures of general conversation and narrative and genre-specific communicative events.

2.3 GENRE-ANALYSIS AND EST DISCOURSE

2.3.1 Types of EST Discourse

Two broad types of EST discourse - science as discipline discourse and science as subject discourse - have been identified (Widdowson, 1979: 52). While the former is expository and characterized by the discourse of research articles, the latter is didactic and represented by scientific textbook discourse. Glaser (1982: 69-83) has also characterized them as the academic, scientific and technological style on the one hand and the didactic style on the other. According to Glaser (op cit), the academic, scientific and technological style represent the discourse in monographs, articles in special journals, conference papers and in academic lectures. Glaser distinguishes these highly technical forms which are normally written for "insiders" from the

"didactic style" which "observes the principles of intelligibility of text with an interspersing of technical terminologies with common core words". In a related development, Heslot (1982: 85-105) characterizes what he refers to as the "discipline-type academic scientific and technological discourse" into "primary" and "review" articles. He distinguishes between them on the basis of their function, contending that while "primary articles bring new experimental results, review articles review results from various scientists and evaluate current state of the art".

Apart from these, more or less, formal types of scientific discourse, there exists yet another form which is unlikely to be found in professional journals or in scientific textbooks. As Widdowson (1979: 52) points out, the term "scientific discourse" may also apply to the treatment of scientific topics such as are found in popular journalism - an aspect of scientific discourse that has been characterized as telling of "the narrative of nature" (Myers, 1986).

The foregoing illustrates that there is not just one entity known as "scientific discourse", but instead "types of scientific discourse", each capable of realising a number of distinct genres. Genre-analysis provides the theoretical framework for characterizing these types of scientific discourse and their constituent genres, mainly for pedagogical purposes in EST materials development and writing.

2.3.2 The Structure of Scientific Research Papers

Swales (1981: 8) identifies three main areas in which scientific research papers may belong as:

- (i) the hard sciences
- (ii) the biology/medical field
- (iii) the social sciences.

Of these three, only the first two fall within the framework of our conception of EST discourse, with the second being directly related to the data on which this

study is based. Swales' categorization, is broad, requiring further specification. Besides it can apply to any type of EST discourse and not just science research papers.

O'Connor et al (1976) have attempted a more specific categorization of scientific research papers into:

- (i) Biomedical papers
- (ii) Descriptive science papers
- (iii) Theoretical papers

They differentiate between these types of research papers based on their structure or form, thus:

1 Biomedical papers

- Introduction
- Materials and Methods
- Results
- Discussion

2 Descriptive science papers

- Introduction
- Materials and Methods
- Geographical Context
- Analysis of Data
- Results
- Discussion
- Conclusion

3 Theoretical Papers

- Introduction

- Theoretical analysis
- Applications
- Conclusions

Although O'Connor et al (op cit) successfully identify types of science research papers, it would appear that the basis for their distinction is more formal than conceptual and discorsal.

A more conceptual categorization of scientific research papers is provided by Gopnik (1972). Gopnik classifies scientific research papers into:

- (i) The controlled experiment type
- (ii) The hypothesis testing type
- (iii) The technique description type.

These are distinguished according to the function of information they contain, thus;

- (1) The controlled experiment type.
 - Establishment of experiment (the subject to be examined and treatment to be performed)
 - Statement of techniques and methods
 - Results of measurement and comparisons.
- (2) The hypothesis testing type
 - 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 declining between

hypotheses.

The structure of the "technique description type" was not discussed. Despite this shortcoming, Gopnik's categories provide a more discourse-based framework for the analysis of the structure of information in science research papers. This notwithstanding, Hutchins (1977) thinks that a more discourse-oriented classification of the structures in Gopnik's categories may be represented thus:

	- current hypothesis/paradigm
'the problem'	- demonstration of inadequacies
	- statement of problem
	- statement of new hypothesis or of alternative hypothesis
' the solution'	- testing hypothesis or hypotheses
	- proof of hypothesis or of one alternative hypothesis

Implication of "solution"

Hutchins (op cit) categorization is related to that of Winter (1976), which was taken up and developed by Hoey (1983) as the "problem-solution" model.

Winter used the problem-solution approach to account for a reader's interaction with written scientific discourse. It was conceived of as a text-attack strategy designed to capture the subconscious questions which readers ask as they decode a text. Thus, the model is built around the following questions:

What is the situation?

What is the problem to be solved?

What is the solution?

How is the solution to be achieved?

These questions are summed up into "speech act" categories which relate to the structure and function of information contained in the text as follows:

(1)	Situation	(2)	Situation	(3)	Situation
	Problem		Problem		Evaluation
	Solution		Evaluation		
	Evaluation				

The principles of the problem-solution approach have been adopted and used by some linguists (Jordan, 1982, 1984, 1985, 1986; Ghadessy, 1983; Hoey, 1979, 1980, 1983) for the analysis of both scientific and non-scientific texts. In the process, a number of modifications to the original model have been proposed. For instance Hoey (1983: 82) proposes the concept of "multilayering" to account for the fact that 'evaluation' in the 'problem-solution' model need not necessarily be positive; that it could also be negative or both. Hoey (op cit) points out that the assumption underlying the 'problem-solution' model is that the evaluation of the solution will always be positive, but contends as follows: "The truth of course is that it is quite common for evaluations to be negative". According to Hoey, these two facts can be reconciled by the phenomenon of multilayering. Multilayering is based on the simple fact that a negative evaluation is a signal of problem. As Hoey (op cit: 83) points out:

It remains true that negative Evaluation is a signal of Problem even when the Evaluation is not local but the Evaluation of a Response. Thus, when a Response is negatively Evaluated, another Problem is normally signalled, except where the Negative Evaluation does not allow for any further Response (eg. "This killed him"). In other words, a principle of recursiveness is introduced.

Hoey (1986) has also attempted to relate the 'Problem-solution' model to the concept of clause-relations (Winter, 1971, 1982). Winter (1971) defines clause relation as follows:

A clause relation is the cognitive process whereby we interpret the meaning of a sentence or group of sentences in the light of its adjoining sentence or group of sentences.

Hoey adapts this definition and modifies it to account for the fact that each clause produced in discourse "is offered in the light of its predecessors" (in the case of revised writing, its successors also). Therefore, Hoey defines clause-relation, thus:

A clause relation is also the cognitive process whereby the choices we make from grammar, lexis and intonation in the creation of a sentence or group of sentences are made in the light of its adjoining sentence or group of sentences".

(Hoey, 1983: 19)

Thus, Hoey (1986) points out that patterns of discourse organization "are not on a different level from clause relations; nor are they on a higher rank", contending that "problem-solution patterns, and indeed all other patterns, are composed of clause relations".

Some weaknesses of the problem-solution model as well as of its modifications by Hoey (op cit) have been pointed out. Jarvis (1983: 108) points out that the model is too all-embracing in its characterization of information structure. Ghadessey (1983: 47) states that the model does not clarify its unit of analysis, ie. word, clause-sentence or chunk. Ghadessey is also of the view that it may be difficult under the problem-solution model to determine where one category ends and where another begins. Commenting on Hoey's (1983) use of "paraphrase as a means of clarifying clause-relations", Reddick (1986: 32) contends that Hoey is not proposing a means of clarifying clause relations, but a means of making as explicit as possible a particular discourse fragment, pointing out that the relations identified by Hoey in his exemplifications do not exhaust all possibilities of the use of the term.

It must, however, be pointed out that Reddick's criticism is not directed purposely at Hoey's model, but at pointing out, using Hoey (1983) that "whether we are examining clause relations or zones of turbulence, the 'linguistic structure of text' depends upon the text linguist's interpretation in that it is that interpretation which assigns structure to that text". Therefore, Reddick (op cit: 41) contends that the complexity of the task of interpretation must be faced by discourse analysts if in Hoey's term we want "to discover what in the discourse allows the reader's acts of interpretation to take place."

However, despite its shortcomings, the problem-solution model remains a very useful conceptual framework for accounting for the structure of information in text. Recent studies (Adams-Smith, 1987; Proctor, 1988; Jordan, 1985) indicate that the model applies to a wide variety of texts. For instance, Proctor (op cit: 34) contends that Hoey's suggested analytical procedures have been found applicable and productive in the identification of the overall discourse patterns in his data.

Notwithstanding the compelling evidence provided by these studies, the 'problem-solution' model or its modified version which embraces the concept of clause-relations was not found suitable for the data in this study because "problem-solution" is an abstract superstructural concept which cannot provide an adequate and detailed description of the differences in communicative functions which exist between related text types. It is for this reason that the 'genre-analysis' approach was preferred in this study. The relevance of the genre-analysis approach to the purpose of this study and the data on which the investigation is based may be succinctly summed up by the observation by Dudley-Evans (1987: 5) to the effect that:

Genre-analysis may be used as a classificatory system, revealing the essential differences both between the genre studied and the other genres and also between the various sub-genres.

As has been pointed out elsewhere in this thesis, the first known study to be characterized as an exercise in genre-analysis is Swales (1981a). Swales applied the approach to the analysis of 48 Article Introductions, selected as follows:

- (a) 16 from the 'hard sciences', of which 10 are from Physics, five from Electronics, and one from Chemical Engineering.
- (b) 16 articles from 'the Biology/Medical field, made up of one from Chemical Pathology, five from Dermatology, five from Radiology and one from Plant Biology.
- (c) 16 articles from the Social Sciences, from the areas of Educational Research (4), Educational Psychology (1), Management (2), Business Finance and Accountancy (2), Language (7).

Adopting strategies similar to those employed in schema-theoretic models, Swales (op cit) posits a four 'move' schematic structure for Article Introduction as follows:

Swales Four Move shcema for Article Introductions:

MOVE ONE: Establishing the Field

- (a) Showing centrality
 - (i) by interest
 - (ii) by importance
 - (iii) by topic-prominence
 - (iv) by standard procedure
- (b) Stating current knowledge
- (c) Describing key characteristics.

MOVE TWO: Summarizing Previous Research

- (a) Strong author-orientations
- (b) Weak author-orientations
- (c) Subject orientations.

MOVE THREE: Preparing for Present Research

- (a) Indicating a gap
- (b) Question-raising
- (c) Extending a finding

MOVE FOUR: Introducing Present Research

- (a) Giving the purpose
- (b) Describing present research
 - (i) by this/the present signals
 - (ii) by move 3 take-up
 - (iii) by switching to first person pronoun.

Swales four-move structure begins with "Establishing the Field" (chart above). The purpose of this opening move, according to Swales (op cit: 23) appears to be to attract readership and to indicate that the researchers are "working in a lively or significant research tradition, even if the appearance of their title might indicate the contrary". Having established the field, the author summarizes previous research in Move Two. This may involve a discussion of authors and their specific contributions to the field, what Swales refers to as "strong author orientation" or a citation of the contributions, but with the authors' names in brackets, referred to by Swales as "weak author orientation". Yet another procedure involves merely stating findings as qualified results without identifying who made the findings. In Move Three, the author "prepares for present research" by showing the need for the study. He may also do this by either or all of three ways: by indicating that a gap exists in the present state of knowledge about aspects of the field, by raising some questions about previous studies, by indicating the need to extend or apply the results of previous research findings. Finally, in "Move Four", the author attempts to present his research, its purpose and what it is all about.

Swales' study is not just an attempt to chunk texts into identifiable knowledge structures. It is also concerned with characterizing the linguistic features of each 'Move' and the means by which information in the 'Moves' is signalled.

Swales' Genre-Analysis approach has received wide-spread commendation from applied linguistics and ESP practitioners. Widdowson (1983: 102) sees it as providing a "characterization of the communicative conventions associated with particular areas of language use" which moves beyond merely itemising forms and functions and "into larger schematic units upon which procedural work can effectively operate". Bley-Vroman and Selinker (1984) believe the method could be useful as a strategy for formalising research into ESP rhetorical grammatical analysis. Jarvis (1983) considers the method as providing a framework for understanding "why, when and how particular functions are used in article introduction". He also thinks that it provides further "evidence of the productive nature of a concentration on discourse organisation for understanding communicative function".

The short-comings of Swales approach have also been highlighted. For instance, Bley-Vroman and Selinker (op cit), and Crooks (1984) have pointed to the difficulty of demarcating 'moves', especially 'Moves 1 and 2' following Swales four-move categorization. But as Dudley-Evans (1986) reports, Swales (personal communication) now accepts that these two moves can be conflated to a single move, referred to as 'Handling Previous Research'. Lopez (1982) found only three rather than four Moves in his study of Spanish research articles and Crooks (1984) found no Move 3 in seven of 24 Article Introductions studied using Swales' model. Similarly Hopkins (1985) examined 5 conference papers on Irrigation and Drainage and found that only one of the papers contained all four moves. These findings would seem to suggest that Swales' model does not apply to all texts, thus calling into question the criteria upon which Swales selected his data. These apparent weaknesses do not necessarily invalidate Swales' model. A look at Swales detailed report would show that the four move structure was not identified in all the texts in his data. Therefore, the four-move structure represents the commonest pattern (Swales, op cit: 15), "of

which part two would occur *a priori*, and the others it seemed would be found sufficiently often for us to state that there was a predominating pattern" (op cit: 19).

Another area of criticism of Swales model has been its methodology. Bley-Vroman and Selinker (op cit: 3) argue that Swales method of data selection should have been less random and that Article Introductions should not have been removed from their main articles. Valid as these observations are, it needs to be pointed out that Swales (op cit: 8) at least recognizes the former and attempts to justify it thus:

The random selection had two pre-conditions attached. First, only introductions that contained at least one reference to previous research were included as, after all, the DPR was the starting point for the exercise.

He also acknowledges that (op cit: 9)

A sample of 48 introductions (of usually between 100 and 500 words) is small but not microscopic. I suspect that a corpus of this size lies somewhere between accidental exemplification and a justifiable basis from which to propose adequately-supported generalizations.

In summary, approaches to the analysis of the structure of information in scientific papers may vary depending on the nature of the data being examined and the purpose of the study.

However, neither the approach by O'Connor et al (op cit), Gopnik (op cit), the Rhetorical-Grammatical approach of Trimble et al (op cit) nor the problem-solution approach of Winter (op cit) as modified by Hoey (op cit) would be as adequate as Swales' (op cit) Genre-Analysis approach for characterizing the data in this study. The reasons for this contention may be found in the shortcomings of each approach vis-a-vis the purpose of this study.

2.4 INVESTIGATIONS INTO MEDICAL RESEARCH REPORT WRITING

In relative terms, there appear to be more studies in the area of spoken Medical discourse (see Candlin et al, 1980; Freeman, 1987; Hein and Wodak, 1987; Messaide, 1987 etc) than there are of written medical communication, and, even then, there are

much fewer studies which are based on the analysis of written medical research reports. The interest in the analysis of face-to-face interaction between medical practitioners and patients in medical settings is understandable. As Freeman and Heller (1987: 1) point out, "medical interaction is, by its nature, non-trivial, since communicative lapses and failures may result in potentially serious outcomes". Though not necessarily implied, the relatively fewer studies of written medical communication would seem to suggest that nothing can be said about the writings of medical practitioners.

But as the relatively few studies into written medical discourse have shown, this is far from being the case. For instance, Adams-Smith (1981) citing McPhill (1911) points out that the phenomenon of bad writing in medical journals is a long standing one. She points out in her study that bad writing in medical journals may be attributed to any of the following features: jargon, careless diction, qualifiers, poor flow of ideas, verbiage, pomposity, cliches and vague words, barbarisms, impersonal, indecisive writing resulting from the use of passive rather than active verb forms, euphemisms and genteelisms, punctuation. She suggests that English Language tutors teaching foreign medical students should ensure that materials they select for teaching reading are checked for readability using the criteria listed above. Adams-Smith also recommends that writing should be taught using models which emphasize logic, precision, clarity, directness and brevity together with points on readability, organization and the integration of non-verbal data.

These observations have some relationship with that by Hemminki (1982) whose investigation into results of reviews of quality clinical trial reports shows that most clinical trials were judged to be uncontrolled or poorly controlled. It also relates to the observation by Hibberd and Meadows (1980) that the information contained in clinical trial reports of major medical journals are generally not well reported in the article summary including dosage regimen and duration of therapy.

In line with these observations, investigations into written medical discourse have tended to focus on linguistic features of text. Pettinari (1983) examines the function of grammatical alternation in surgical reports based on functional-rhetorical differences between sentences with indefinite subjects and those with *there* as subject.

She demonstrates that the discourse function of such grammatical alternations "is to provide an alternation between thematic and non-thematic information, pointing out that "the alternation is ultimately controlled by the extralinguistic situation of the surgery itself". The extralinguistic situations are identified as the surgical procedure and how it is sequenced. It also includes features in the situation considered essential to the procedure. On this basis, she concludes that indefinite subjects are thematic and indefinite non-subjects in *there* sentences are non-thematic. Pettinari uses the term "thematic" in the Prague school, tradition as a means of accounting for the way in which textual organization is linked to the extralinguistic situation. Thus, in line with linguists of the Prague School, Pettinari contends that textual organization can be analysed by an examination of *shared* and *non-shared* information, and in addition, by an examination of information "which is crucial to the development of a *text* or *subject*". This latter proviso determines her characterization of thematic information as "that knowledge which the surgeon assumes is relevant to the goal of the communicative event of describing the surgical procedure for medical record purposes."

This categorization of thematic information appears not to be consistent with known definitions of the term. Besides, there could be problems with the identification of "that knowledge which the surgeon assumes is relevant to the goal of the communicative event", since that which is important to the communicative event may not necessarily be the theme of the utterance in Functional Sentence Perspective.

Dubois (1981) examines the discursual functions of noun phrases in Biomedical Journal Articles and observes that prenominal modification of head nouns is a stylistic feature of Scientific English; that there is a pile-up of content modifiers left of the noun which can be avoided by reduction or deletion of terms, use of pronouns and coding of complex meaning into simpler synonymous expressions. Dubois concludes that construction of NPs can serve the discursual function of delimiting a section or subsection of a journal article.

Bruce (1983) examines the rhetorical constraints and information structure in Medical Research reports based on a selection of articles from the *British Medical Journal*. The study was aimed at showing how "the syntactic form of a message is determined by certain discourse connections and communicative constraints such as type of general rhetorical function and level of conceptual sophistication". Bruce argues that each section of the research report features different rhetorical functional patterns which are realised linguistically by different structural patterns. Although Bruce does not label his approach as the rhetorical-grammatical approach, contentions such as the one above suggest that his study relates to that tradition. Another unsurprising observation by Bruce (op cit) is that there is no neat given-new cycle, either at the sentence level or at the NP level; that "what linearity there is lies at a non-linguistic level, discernable in a schematization of logical procedures in the methods section, and, in the Results section, in a deliberate series of measurements taken of combinations of variables to arrive at the proof or disproof of target hypotheses". It has been shown (Danes, 1974) that the progression of information in discourse can take any of five patterns, of which one is the simple linear pattern. Therefore, there can hardly be a neat "Given-New" Circle in any piece of discourse.

Salager (1986) adopts a rhetorico-grammatical approach to the analysis of infinitive clauses in Medical English literature. The study which was based on a corpus of 32,000 words yielded the following results at the level of grammatical analysis.

60% of the infinitive clauses were verbal complements, 22% adjectival complements and 18% nominal complements.

At the rhetorical level, the study indicates that:

- The verbal antecedents express notions of purpose, cause-effect, methodology and result of investigation.

- The adjectival antecedents indicate possibility, probability or necessity.
- The nominal antecedents are abstract items bearing a modal connotation and being mostly derived from the aforementioned verbal and adjectival roots.

Salagar's study is interesting for the notional, functional and semantic interpretations it brings to bear on a grammatical item in a specific area of language use. Its conception of rhetorical functions, which is not restricted to the formal structural features of the medical research paper (IMRD structure) but to notions such as 'purpose', 'probability' and 'necessity' is a significant departure from the original rhetorical-grammatical model of Lackstrom et al (op cit).

In a study based on the analysis of articles from the *British Medical Journal* and the *New England Medical Journal*, Adams-Smith (1984) examines aspects of author's comment in Medical discourse. The purpose of the study was to determine "how, why and where the subjective element is introduced in the main categories of articles in medical journals". The study indicated that the medical research report writer introduced the subjective element mainly by means of verbal modality - the modal auxiliaries such as may, should; non-verbal modality, especially adverbs (eg. probably, possibly, certainly); attitudinal markers (eg. surprisingly, appear). Adams-Smith observes that editorials are more author marked, "while clinical case notes and research papers contain sharp differences between the objective recounting of the methods and results sections and the more subjective discussion or comment section". The shortcomings of this study have been pointed out by Swales (1987). Swales points out that the method of data selection for the study is highly debatable since all the texts used in the study come from a single journal and are limited to six examples of each of her three categories. However, the study has its merit which as Swales (op cit: 124) points out, has to do with "the often-misunderstood question of quantification. Commenting, Swales (op cit: 124) states that:

Adams-Smith uses frequency data to allow the reader to appreciate the variable extent of the phenomena she investigated, most tellingly in the final table entitled "Incidence and Ratio per

most tellingly in the final table entitled "Incidence and Ratio per Line of Author's Comment by Category and Section of Article" (1984: 35). She does not use quantitative methodology as discovery procedure but as evidence which shows, *inter alia* that the presence of personal emphasis, advice, and evaluation in medical writing is under-recognized in training materials.

McKinley (1983) attempts a genre-based analysis of the Discussion section of 30 articles taken from six medical journals. The study, which is closely tailored along Swales (1981a) model of schematic organization of information in scientific texts, identified five "moves", thus:

- Move 1: Comment on Methodology
- Move 2: Results and comment
 (reference to other work)
- Move 3: Results and comment
 (reference to other work)
- Move 4: Results and comment
 (Hypothesis)
- Move 5: General conclusion.

A major shortcoming of McKinley's study is that the 'Moves' appear to be highly generalized as no indication was given of the sub-categories which make up each of them. Following this, it is difficult to say whether Moves 2, 3 and 4 are really separate moves or whether they are sub-categories of a single move. Furthermore, it is not clear whether it is being suggested that the choice of hypothetical expressions and lexical items in Moves 3 and 4 imply that a hypothesis is being stated, confirmed or rejected. These weaknesses notwithstanding, the study is interesting in its detailed analysis of linguistic signalling in the texts examined.

The studies discussed in this section point to the importance which the analysis of medical writing has for applied linguistics and for medical practitioners in the field. But, how do these studies compare with the present research investigation into related genres of written medical discourse? First, none of them attempts an analysis of related genres in terms of patterns of the organization of information at the level of schematic structures. Second, none of them examines the problems of the presentation of information in medical discourse at the micro-structural level. Third, none of them attempts to compare these patterns across genres of written medical

research papers. Of course, all of them were addressing part of the problem, but none of them attempts to present the picture in a manner presented in this study.

2.5 INVESTIGATIONS INTO ABSTRACTS OF RESEARCH ARTICLES

Very few discourse-based studies have been conducted into the linguistic properties of research abstracts (see Graetz, 1982; Gopnik, 1972; Rounds, 1981, 1982; Urquhart, 1986). All but one of these studies were based on the study of Abstracts taken out of the context of their original articles. This may well be a methodological flaw, but it may also relate to the primary purpose of the researches themselves. Evidence from the studies would seem to support the latter, for of these studies, only Graetz and Urquhart set out to investigate the pattern of organization of information in Abstracts. The others used Abstracts as data for the investigation of other linguistic problems, and merely touched on the structure of information in the texts which constitute their corpus.

Graetz (op cit) and Urquhart have posited different structures for research Abstracts. Graetz (1982) identifies three possible ways of organizing information in Abstracts as:

A	B	C
1 Objectives	1 Objectives	1 Objectives
2 Purpose	2 Methods	2 Methods
3 Experimental design	3 Results	3 Results
4 Methods		4 Validity of results
5 Results		5 Conclusions
6 Conclusions		6 Applications

Graetz's study is based on 87 samples of abstracts found in 41 arbitrarily selected scholarly journal articles, thus: 8 from Health Sciences journals, 13 from Social Sciences, 5 from Education and 15 from Humanities. The arbitrariness in the method

of data selection raises a few methodological questions. For instance, it is not quite clear why five Abstracts were selected from Education and why as many as fifteen were selected from Humanities. It is also unclear why Abstracts from other branches of science were excluded from the study. On the other hand, the fact that the Abstracts were studied in relation to the research articles which follow them, arguably, argues well for the validity of the results because it ensures that the analyst can confirm his intuitive judgements by matching information in Abstracts with their equivalents in the Research article. However since the data for the study is drawn from a wide range of disciplines, one would have thought that an attempt would have been made to relate the possible structures identified in the analysis to particular disciplines or a particular group of disciplines. Since this was not done, it is not clear whether Graetz intends that these three possible structural types could apply to all the disciplines which were examined in the study. Our intuition about variations in the organization of information in texts would certainly suggest that there might be differences.

Urquhart (1986) examines 50 texts selected at random from Genetics Abstracts and identifies 10 possible functions of Abstracts as:

Problem; Reference to past literature; Gap-Indication, Reference to Current Knowledge; Reference to the Project in Progress; Textual Reference; Aim; Procedures, Findings; Conclusions.

He finally reduced the categories into 4 functions:

- Aim
- Procedures
- Findings
- Conclusions.

Urquhart's results are an improvement on those of Graetz in that his categories are stated in functional terms and because his data is "genre-specific". Since as the study by Graetz has shown, there can be no one structure for all forms of Abstracts, Urquhart's study would seem to have greater practical applications than Graetz's because it characterizes the discourse functions of Abstracts in terms of their

communicative functions. However, one major shortcoming of Urquhart's study is that it is based on Abstracts from an Abstracting journal. These abstracts are most often written by professional abstractors who may or may not be specialists in the field or may not even be familiar with the topic of discourse. Besides, such abstracts tend to be written for slightly different purposes from those which accompany research articles. While the latter set of abstracts function to indicate and predict the structure and content of the text which follow them (Graetz, 1982), the latter tend to be written primarily with indexing and information retrieval as primary purposes. Therefore, it is not clear to what extent results based on abstracts from Abstracting Journals can be said to be representative of the structure of Abstracts in general.

One problem which can be posed by an analysis of Abstracts which have been taken out of their contexts is that "analysis of further abstracts might lead to more functions being described" (Urquhart, op cit). Perhaps one approach that could help to resolve the problem of what Urquhart refers to as "more minor (less frequent) functions" is the examination of abstracts which accompany articles with the aim of determining the relationship existing between them. It is reasonable to argue that a research design which attempts to account for the structure of abstracts by direct reference to their accompanying articles is more likely to give us an accurate description of the linguistic functions for the organization of information in Abstracts, by highlighting the relationship or fit existing between both genres. Such an approach would also enable us to validate the results of these previous studies which are based on Abstracts from Abstracting Journals, as well as "validate the advice given by pedagogical rhetoricians with regards to organization " (Urquhart, 1984).

2.6 INVESTIGATIONS INTO SCIENCE POPULARIZATION

2.6.1 Historical Perspectives

The need for science popularization arose towards the end of the seventeenth century when the emergence of the quantitative, mathematical approach to knowledge of the physical world left behind the majority of readers (Meadows, 1987; 341). Thus, it became necessary to present science in a popular form to non-specialists for most of whom the content of science was becoming increasingly impenetrable. By the nineteenth century the need for popularization had become much stronger as the publication in journals gradually replaced publication in book form for the different branches of science. This shift from publications in book form to publication in research journals is seen as an indication of "the growing professionalization and specialization of science", which in turn, is related to the "increasing difficulty of the subject for non-scientists" (Meadows, op cit). In the twentieth century, the scope for science popularization further widened with the recent advances in communication technology in the form of Radio and Television science broadcasts. While the trend towards popularization has continued to grow in the industrialized nations of the west, it would appear that it is yet to take firm roots in the developing countries. In an attempt to account for the imbalance in the state of scientific knowledge and science popularization between the developed and the developing countries, Sharafuddin (1987) states that:

The invention of printing, and the consequent ease in the dissemination of knowledge; the habits of logical thinking, accurate observation and measurement; the development of new instruments for observation and measurement such as the clock, the telescope and the microscope; and above all, a questioning mind that places experimental evidence above authority and dogma - all of these factors led to the Industrial Revolution.

In the absence of any parallel development in the Third World countries most of which were meanwhile colonized by the industrially advanced countries, they gradually remained immersed in their past traditions.

(Sharafuddin, 1987: 348)

And on science popularization today in the Third World, Sharafuddin (op cit: 349) further adds:

One big advantage the Third World countries have today is that it is no longer necessary to argue the case for science and technology ... However, this unconscious acceptance of science does not necessarily make it easy to popularize science and its methods among the mass of people. Their lack of education and acquaintance with the basic principles of science, coupled with their traditional reverence for the old ways, in fact makes the task extremely complex.

While these observations paint a very grim picture of the state of science popularization in the developing countries, they also point to positive prospects for the future, thus stressing the need for the study of science popularization by linguists and media practitioners in these countries.

2.6.2 The Need for Science Popularization

Having very briefly reviewed the history and development of science popularization in developed and developing countries, a pertinent question that needs to be asked is: Why bother to interpret science to the layman? Part of the reason lies in the fact that there is such an overwhelming increase in the amount of scientific and technological knowledge available today that even scientists are unable to keep abreast of it. This state of affairs is even more daunting for the layman whose knowledge of science is most likely to be rudimentary. But, despite this fact, the layman has to contend with the domination which science and technology have over their lives in the modern world. For this reason, it is necessary for the layman to be constantly made aware of the results of the numerous scientific research projects and how they could affect their lives.

It is ironical that despite the extent of domination which science and technology have over the life of the layman, there appears to be a general lack of interest among laymen about the science underlying things which shape the world they live in. This lack of interest would appear to have created a gap between scientists and the public which science is supposed to serve. Such a gap is definitely not in the

interest of science nor of the public. It is a gap which could badly affect the development of research in science because ignorance could lead to negative public reactions to the financing of research projects through public funds. Issues such as these and more have been succinctly summed up in Burkett (1973: 39), where six reasons given for why science should be interpreted to the layman are identified as: the importance of science as part of the general cultural knowledge of man, political considerations for funds, financial support from laymen, the need for science to show what it is doing in a direct and explicit way, the need for compatibility between the aims of journalists and scientists, and the need for scientists to bridge any gap in understanding between the arts and sciences. Similar reasons have also been given by Burkett whose views are related to those of Jean Rostand, cited in Burkett (op cit: 41). According to Rostand, popularization is necessary because it fills the gaps left in school education which lags behind progress, arouses the desire to take up research, acquaints the public at large with the power and efficiency of creative science, creates a link between specialists in different disciplines and keeps politicians informed.

These views by Burkitt and Rostand are eloquent testimonies of the need for science and its popularization; reasons one would have thought would have guaranteed considerable interest in the production of popularized texts in the mass media as well as the study of popularized texts by linguistics, media practitioners and sociologists. But as the issue of the UNESCO journal "IMPACT" which was devoted exclusively to science popularization indicates, "few textbooks exist to instruct those who wish to popularize science; and the subject is infrequently explored in print, outside the publications of organizations directly devoted to it" (1987: 337)

2.6.3 Types of Science Popularization

About a hundred years ago, it was customary for scientists to popularize their own works. Meadows (op cit: 344) cites examples of Michael Faraday's *The Chemical History of a Candle* and Thomas Huxley's *On a Piece of Chalk* as attempts by scientists to popularize their works by means of easily understood public lectures.

In modern times, with the expansion of popular magazines and newspapers, there has emerged a group of professionals whose duty it has become to popularize science to the public. Jones et al (1978) refer to this group of professionals as "the entrepreneurs of science". They include "journalists, producers, etc who act as intermediaries between the scientists and the general public". And as Pallab Ghosh, science editor of *New Scientist* says (personal communication,), contributions to the science section of the magazine fall somewhere in the middle between scientists who have left research to become journalists or who may still be actively involved in research and journalists who have followed a particular field for many years.

Science popularization may be classified according to the mass media channels in which the popularization takes place, eg. the print media (newspapers, magazines) or the electronic media (radio and television) (Jones et al, 1978). It may also be classified according to the status of the popularizers, ie. science researchers who are called upon to popularize their works on the one hand and the so called "entrepreneurs of science" or science writers on the other. Popularization of the former category has been studied by Myers (1986). This study was based on articles in the *Scientific American*. A study based on the latter category of popularization is Adams-Smith's (1986/87) analysis of "derived" materials from the *New Scientist* and *Times* newspaper.

2.6.4 Science Topics in the Print Media

Attempts have been made to examine the distribution of science information in newspapers and magazines. In a study of four British newspapers - *Times*, *Daily Mirror*, *Daily Express* and *Leicester Mercury*, Jones et al (1978) examined items devoted to scientific or science related issues in the newspapers and arrived at the following distribution:

Medicine (including Human Biology)	75%
Behavioural Science	6%

Engineering/Technology	5%
Biology (Natural History)	4%
Space	3%
Earth Sciences	3%
Physics	1%
Chemistry	1%
General (science policy etc)	1%
Unclassified (science fiction etc)	1%

Farago (1976: 32) states as follows:

A recent survey analysed the subject-matter of science stories appearing in three major American newspapers during 1973. In a one-week period the Los Angeles Times published 22, the New York Times 32, and the St Louis Post-Dispatch 16 science articles, the great majority of them being less than 20 paragraphs. Between a third and a half of the articles dealt with medicine and public health; between 10 per cent and a quarter with agriculture and biology; while all the others, for example the earth sciences represented well under 10 per cent each.

A striking similarity between these two studies is the predominance of medical items over items from other branches of science. Much earlier surveys (Davis and Sklair, 1972; Kriegbaum, 1967) have also made similar observations. This may, perhaps, explain why the few studies on science popularization tend to be based on Bio-medical texts (see Myers, 1986, 1988; Dubois, 1985; Adams-Smith, 1986, 1987, 1988).

It has been suggested (Jones et al, op cit: 4) that the popularity of medicine in newspapers may be due to the fact that it has "human interest", and "it is concerned with issues - such as health, birth, death and sex". This is seen to be compatible with the basic aim of newspapers - that of catering for the human interests of their readers. The human interest factor would also seem to explain the assumptions which the "entrepreneurs of science" have about their readers, ie. that the reader is more likely to find popularized science material interesting if the results of the research investigation affects him directly even if he does not fully understand the abstract principles

underlying it.

2.6.5 Some Approaches to the Study of Science Popularization

Science popularization has been studied from a number of perspectives - the sociology of health (Karpf, 1988); the media (Jones et al, 1978; Taylor, 1975); Semiotics (Khan, 1983; Funkhauser et al, 1970); discourse analysis (Adams-Smith, 1986/1987/1988), the sociology of science/discourse analysis (Myers, 1986). In this section, I shall examine a few of these studies in order to determine the extent to which they relate to this research project.

Karpf (1988) examines the reporting of health and medical issues on radio and television broadcasting. Karpf's central aim is to determine whether the persistent frames used in the media coverage of health and illness "permit a healthy debate, and why and how alternative frames or approaches are regularly excluded". She classifies radio and television programmes on health and medical issues into four: the medical approach, the consumer approach, the look-after-yourself approach and the environmental approach. Karpf characterizes the medical approach as that in which the doctor's authority over what constitutes diseases and how it should be treated is considered legitimate. According to Karpf, "the medical approach validates experts, especially the expertise of the doctor, and the visual grammar of medical programmes tends to reinforce the doctors centrality and authority". Unlike the medical approach, the consumer approach is characterized as that in which the patient plays the leading role in the programme. Thus, "many consumer programmes give information, the sort which the doctor rarely gives but which could help the patient make choices". As Karpf (op cit 15) points out:

It's consumer medical programmes and writings which have been pejoratively labelled "doctor-bashing". Doctors are often the baddies of these programmes and, as with the medical approach, they're assumed to be uniquely powerful, their inadequacies held accountable for the inadequacies in medical care.

The third approach, "the look-after-yourself" approach is a new approach which emphasizes preventive medicine. The cause of illness is not seen as a purely

biological thing in which medicine is the only solution. Rather, as Karpf (op cit: 18) points out:

The look-after-yourself approach, on the other hand, reinstates the mind and will of the patient: illness is the result of harmful individual habits or a 'life style' undertaken voluntarily - eating the wrong foods, drinking too much, smoking, lack of exercise and stress. Where the consumer programme saw patients as innocent victims, the look-after-yourself programme sees them as active, though perhaps unwitting, self-harmers.

The "environmental approach" is very much like the look-after-yourself approach in that "both reject notions of pathology and cure endemic to the medical programme, and turn instead to the preventable causes of illness. Both believe that there are limits to medicine". But unlike the "look-after-yourself", the environmentalists believe "today's chronic and degenerative diseases to be as environmentally or socially-caused as the disease of the past. They challenge the belief that a wealthier society is a healthier one, and implicate the social structure itself". Karpf is very critical of the look-after-yourself approach in particular, arguing that they may even help make people sick. Other issues examined by Karpf include media coverage of childbirth, alternative medicine and media "moral panics" over surrogate mothers, herpes and AIDS. Karpf argues that:

in spite of greater diversity in media's reporting of health and medical issues over the past decade, medical definitions and perceptions still prevail, and squeeze out more contentious, oppositional viewpoints which take an environmental approach and look at the politics of health. (op cit: 2)

and concludes that in order to bring about change in the media's coverage of health and medicine,

Journalists would need access to alternative sources of information, and would have to consider them legitimate. They'd need training in questioning dominant views, requiring a whole new ideology of journalism.

(op cit: 239)

Karpf's study is interesting, especially in its categorization of approaches to

the reporting of health and illness in radio and television programmes which might prove useful in the study of science popularization in newspapers and magazines. Similarly, ideological and social insights which Karpf brings into her study might also prove useful for interpreting linguistic evidence from an analysis of similar or related materials. However, Karpf's work tends towards speculation and makes use of a lot of emotive language in condemning opposing views and projecting alternative ideological approaches. It would, therefore, appear that Karpf uses the popularization of health and illnesses in the media as an "inferential framework" or as "expression" for other issues in society such as ideologies, power, racism and sexism. Thus not surprisingly, Karpf's study has greater sociological than linguistic import. This, however, is not to suggest that the study has no relevance to a linguistic analysis of science popularization. Indeed, as Van Dijk (1985: 3) points out, no full-fledged theory of media discourse is adequate without explicit inferences of the type provided in Karpf's work.

Jones, Connell and Meadows (1978) approach science popularization from a strictly media perspective. They point out that their study is the first "integrated study of the presentation of science by the main mass media channels (viz: newspapers, radio and television) in the UK". As has already been pointed out, their analysis of science in newspapers is based on four British newspapers (2.6.4 above). Jones et al conducted a content analysis of the four newspapers over a period of three months and observed that medical items tend to dominate scientific news in newspapers. In specific terms, they found that medical and behavioural science items accounted for the percentage of total science items in each newspaper.

Based on their observations, Jones et al arrived at the conclusion that:

- (i) The more popular the newspaper, the less science it makes available to the public.
- (ii) The more popular the newspaper, the higher the proportion of biomedical items it surveys.

- (iii) The more popular the newspaper, the more limited, therefore, its coverage of science.

Jones et al (op cit: 2)

Apart from these conclusions, they also found that only 8% of the science items in the newspapers they examined were "pure exposition of scientific information", the rest of the items being associated with issues of social, as well as scientific relevance.

Jones et al sum up their observations, thus:

- Science is not widely represented in the newspapers unless it touches on questions of general societal interest.
- A large proportion of the items in the press are concerned with the moral and ethical aspects of science.
- Some degree of suspicion and criticism often underlines the presentation of popular science in newspapers.

As their results indicate, the study is more of a content analysis than a linguistic investigation of science popularization in the media. Its merit lies in its ability to provide partial insight into properties of media discourse, such as the prevalent themes in a particular type of media discourse, the kind of actors in the discourse, advertisements, TV programmes or style, or style in the headlines. As Van Dijk (1985: 3) points out, "content analysis is interested in principles of description when it tries to account for the basis of unitization, and in this way also structural properties of media discourse can be attended to, at least from a quantitative point of view". However, the shortcomings of the study, from a discourse point of view, is that it does not account for the underlying meanings or processes realised in the texts examined.

A study that attempts to combine content analysis with some measure of description of underlying processes in the discourse is Khan's (1983) semiotic comparison of textbooks of medicine and popular home medical books. As with most semiotic analyses of text, the study compares the two text-types in terms of social, historical and critical contexts in which they occur. It also cross-compares what Khan refers to as "the text (that portion of a book extending from the beginning of the first chapter until the end of the final chapter)"; the table of contents and the "paratext" defined as everything in the text. By means of quantitative methods, Khan examines the distribution of lexical items and non-discursive elements such as reference-to-page ratio, illustration-to-page ratio and author-to-page ratio. He also used mapping methods to analyse selected headings and schemata.

In terms of style, which is the most relevant to this study, Khan (op cit: 233) contends that textbook style is characteristically formal, authoritative and high in accountability compared to popular style which may be described as informal, authoritarian and low in accountability.

Khan (op cit: 237) also argues that the relationship between author and the reader in textbooks is one of "master" to "apprentice" while that in popular books is between a "friendly consultant" and a "concerned client".

The major weakness with Khan's study lies with its methodology which is highly quantitative, involving the counting of words per sentence. As a result, it does not say much about the actual interaction in the texts being compared. Another weakness of Khan's work lies in the numerous unexplained terminologies which it introduces into the study. For instance, it is not clear what Khan means by "impervious" and "non-impervious" sentences. Khan's characterization of "schemata" as "bold-type headings in chapters in both books" does not appear to be consistent with the use of the term in cognitive psychology. Apart from the little section devoted to the analysis of style, the vast proportion of the study is not concerned with the message or discourse in the texts compared. Therefore, the analysis of the message or

discourse is only incidental and subordinate to the analysis of the formal properties of the texts analysed. This notwithstanding, the study provides a partial insight to an understanding of a particular type of popularization - the popular home medical book.

Unlike Khan (op cit), Adams-Smith (1987) is a discourse-based analysis of 'source' and 'derived' discourse aimed at examining the differences between scientific research papers and their popularized versions. Adams-Smith examined features of these text-types such as field - lexical and syntactical choices, lexical density and the framing of the subject matter; features of tenor - including tentativeness, formality and author orientation; features of mode - overall organization of the text, clause relations and lexical signalling.

Adams-Smith observed that source discourses make use of multiple pre-head modification, many nominal groups and passive or stative verb forms while their derived equivalents make use of simpler pre-head modification and post-head modification and post-head qualifiers, many verbal groups and more active verb forms. It was also observed that both forms of discourse have high lexical density.

The analysis of the overall organization of text is based on Winter/Hoey's Problem-Solution model. As Adams-Smith (op cit: 19) points out:

I began by basing my analysis on the hypothesis that research papers would fit into the problem-solution pattern described by Hoey.

... In the absence of evidence to the contrary, I assumed that the derived discourses would reflect the clause relational patterns of the source papers.

On the basis of these assumptions, Adams-Smith observed that the research paper is organized according to this pattern:

Situation

Unexplained

Procedure

Findings (explanation)

Findings (evaluation)

And the derived discourses, according to the same pattern but with a modification - superimposed on the predominantly "unexplained-explained" pattern of the source discourses is the problem-solution format. Thus, according to Adams-Smith (op cit: 29), the primary distinction between both discourse types in terms of the overall organization of texts is that:

The source articles, taking as their starting point a disputed topic or a gap in knowledge, without exception reflect the unexplained-explanation pattern of clause-relations, although some contain hypothetical - real sequences. In general, members of the set of derived discourses reflect the problem-solution pattern.

It would appear that Adams-Smith's work has some terminological and methodological weaknesses. The parameters upon which the study is based would seem to suggest the view that it is a register analysis of source and derived discourses, but as Adams-Smith (op cit: 13) claims, the study is an attempt to cast more light on patterns of variation typically present in field-related genres. As the results from the study show, differences in discourse organization between the two text-types in the study are not based on genre theory but on the "problem-solution" model of text structure. Dudley-Evans (1987: 5) has pointed out that the problem-solution model differs from the genre analysis approach because it attempts to describe relations found in all texts and not those found in specific genres. Thus, the contention that "the source articles ... without any exception reflect the unexplained-explanation pattern of clause relation" and that "members of the set of derived discourses reflect the problem-solution pattern" is to make a statement that can apply to research papers in any form, field or situation and derived discourses of all types. This does not tell us much about the differences existing between the two text types in terms of the characteristics which make the text-types unique as distinctive communicative events or genres.

Despite the methodological and terminological shortcomings pointed out above, Adams-Smith clearly demonstrates the enormous potential which a linguistic study of "source" and "derived" texts of the type she studied has for an understanding

of the changes which take place in the construction of texts for a popular audience. A further indication of this potential is indicated in Adams-Smith (1988).

A very significant contribution to the understanding of the process of science popularization is Myers (1986) which, in many respects, relates to his earlier works (1985a, 1985b). Myers (1986) differs from the studies so far reviewed in this section in its approach to the study of science popularization. As Swales (1987: 126) points out, Myers avoids categorization and the use of any formal analytic apparatus and adopts an approach that "is somewhat akin to those associated with rhetoric or literary criticism". At the same time, he draws heavily from insights derived from studies on the sociology of science as well as discourse analysis.

Myers' basic argument is that science articles for professionals and science articles for the general public present two incompatible views of what a scientist does, and that both views play complementary roles in creating the cultural authority of science, (Myers, op cit: 1). Based on this assumption, he makes the contention that articles for the general public tell a different story from articles for professionals. According to Myers (op cit: 2):

The professional articles create what I shall call a narrative of science, they arrange time into a parallel series of simultaneous events all supporting their claim, and emphasize in their syntax and vocabulary the conceptual structure of the discipline. The popularizing articles, on the other hand, present a sequential narrative of nature in which the plant, or animal not the scientific activity, is the subject, and in their syntax and vocabulary they emphasize the externality of nature to scientific practice.

Myers' study is based on an analysis of articles on Evolutionary Biology in *Science* and *Evolution*, as well as articles by the same authors on the same topics for more popular journals, *Scientific American* and *New Scientist*. The purpose of the study is to determine textual differences in narrative structure, in syntax and in vocabulary. Adopting an ethnomethodological approach by engaging in conversations with authors of the professional and popular articles, Myers observed that the authors of the popular articles "describe the differences between their articles for professionals

and their articles for popular audiences in terms of level of information". But, Myers shows that in addition to levels of information, the differences can also be seen in textual terms, in titles, abstracts, opening sentences, organizational devices and illustrations.

Myers findings may be summed up thus:

- Titles of the professional articles states an index heading and a claim to be proved (op cit: 8) while titles of popular science articles tend to link areas in the research report which are most intriguing to a general reader (op cit: 9).
- Abstracts of professional articles confirm emphasis on the work of the scientists and its importance to other scientists. Abstracts of popular articles, on the other hand, emphasize nature and its fascinations (op cit: 9).
- The opening sentence in the introduction section of the professional articles tend to outline a problem for biologists (p 11) and to place the article immediately in the context of the literature. The opening sentences in the introduction to the popular articles tend to present a problem for mankind, not for biologists (p 11) and in so doing the reader is drawn into the article (p 12).
- In terms of organization, the narrative in the professional articles depends on rearranging a number of events into a simultaneous order of arguments, that is, not chronologically, but in a hierarchial order following the argument (p13-14). The popular articles, on the other hand, organize the narrative following the activities of the organisms or objects in the research which are often subordinate to all accounts of experiments.

Not much can be said about the shortcomings of Myers work. Perhaps the only criticism which can be levelled against it as a study which Myers would like to

refer to as a sociological discourse analysis (Swales, 1987: 127) is its "avoidance of categorization and/or formal analytic apparatus (Swales, op cit: 126). Despite this fact, the work has been acclaimed as a huge success. According to Swales (op cit: 127):

Not only does Myers patently succeed in his hope, but the paper adds a whole new dimension to the long-standing debate regarding the appropriacy of popularizations as material for English for Academic Purposes Courses.

The text which forms the corpus for this study are drawn from three genres of written medical discourse. The main body of the Research Article, otherwise referred to as the "Research Article Proper" (RAP) is identified as a distinct genre aside from its accompanying Abstract which is also assigned the status of a distinct genre. The Abstract, otherwise referred to as the 'ABS', is assigned the status of a genre in this study because as will be shown throughout the study (see page 4) it

- (i) ...
- (ii) ...
- (iii) ...
- (iv) ...
- (v) ...
- (vi) ...
- (vii) ...

Further statistical details will be provided in Chapters Four, Five and Six.

3.1 THE CORPUS

The text which forms the corpus for this study are drawn from three genres of written medical discourse. The main body of the Research Article, otherwise referred to as the "Research Article Proper" (RAP) is identified as a distinct genre aside from its accompanying Abstract which is also assigned the status of a distinct genre. The Abstract, otherwise referred to as the 'ABS', is assigned the status of a genre in this study because as will be shown throughout the study (see page 4) it

CHAPTER THREE
RESEARCH METHODOLOGY

3.0 INTRODUCTION

This is a corpus-based analysis of discourse in genres of written medical texts. As in almost all corpus-based studies, a lot depends on the extent to which the texts which make up the corpus enable an investigation of the kind of problems which the study purports to examine. In other words, much of the reliability and validity of the procedures adopted in the process of analysis and the results obtained therein depend very strongly on selecting texts from a 'representative' sample, based on criteria consistent with the objectives of the research investigation and on methods which enhance the realization of these objectives. Therefore, consistent with this fact, I attempt in this chapter to outline details of:

- (i) Sources of texts which make up the corpus
- (ii) Factors influencing the identification and selection of texts
- (iii) Basic methodological principles underlying the study
- (iv) Methodological approaches to the analysis of the organizational features of texts
- (v) Criteria for the identification of organizational features of text.

Further methodological issues will be treated in Chapters, Four, Five and Six.

3.1 THE CORPUS

The texts which form the corpus for this study are drawn from three genres of written medical discourse. The main body of the Research Article, otherwise referred to as the "Research Article Proper" (RAP) is identified as a distinct genre aside from its accompanying Abstract which is also assigned the status of a distinct genre. The Abstract, otherwise referred to as the 'ABS', is assigned the status of a genre in this study because as will be shown elsewhere in this study (chapter 4) it

is a recognised communicative event, with explicit constraints on the positioning form and content of materials in it relative to the RAP. The third genre in the study is referred to as the "Journalistic Reported Version" or JRV. It is a rewritten version, for a less professional audience, of the medical research investigation presented in the RAP. The JRV is almost invariably written by science journalists and most often appears in the "science" columns of popular science magazines and newspapers. Thus, the three genres in this study together constitute a group of triadic texts which by virtue of the relatedness in their cognitive content form what Hartman (1980) refers to as "parallel" texts.

Therefore, the corpus for this study consists of fifteen (15) triadic, but parallel texts which together add up to a total of forty-five (45) texts. They represent the sum of an equal number of texts from each of the three genres of written medical discourse referred to above. The forty-five texts contain an estimated 75,000 words. They were selected from medical research journals, magazines and newspaper reports published between April 1985 and January 1987.

3.1.1 Sources

The texts in this study may be broadly categorized into: (a) professional articles and (b) popularized articles.

The professional articles are taken from five sources:

- 1 The Lancet
- 2 The New England Journal of Medicine
- 3 The Journal of the American Medical Association
- 4 The British Medical Journal
- 5 The Journal of Clinical Investigation

The popularized articles, are taken from three main sources:

- 1 New Scientist
- 2 The 'Times' Newspaper
- 3 'Newsweek' Magazine.

3.1.1.1 Coded List of Texts and Sources

The 15 triadic texts which form the corpus for this study are numbered in groups from 1 to 15 with each group being made up of three texts - that is, one text from each genre in each group of triadic texts.

Table 1: List of Texts and Their Sources

Group	Genre	Title	Source
1	RAP 1	Controlled Trial of Treatment of Recurrent Spontaneous Abortion by Immunisation with Paternal Cells.	The Lancet, 27 April 1985
	ABS 1	Same as above.	Same as above.
	JRV 1	Immunisation Can Prevent Miscarriage.	New Scientist, 9 May 1985
2	RAP 2	Oral Contraceptives and Breast Cancer in Young Women.	The Lancet, 2 November 1985
	ABS 2	Same as above.	Same as above.
	JRV 2	Pills Use and the Risk of Breast Cancer.	New Scientist, 7 November 1985
3	RAP 3	Low-dose Aspirin Prevents Pregnancy-Induced Hypertension and Pre-eclampsia in Angiotensin-sensitive primigravidae.	The Lancet, 4 January 1986
	ABS 3	Same as above.	Same as above.
	JRV 3	Aspirin prevents hypertension in pregnancy.	New Scientist, 6 March 1986

Group	Genre	Title	Source
4	RAP 4	Epstein-Barr Virus Receptors on Human Pharyngeal Epithelia.	The Lancet 1 February 1986
	ABS 4	Same as above.	Same as above.
	JRV 4	Epstein-Barr Virus Finds a New Route to Cancer	New Scientist, 6 March 1986.
5	RAP 5	Infant Mortality, Childhood Nutrition, and Ischaemic Heart Disease in England and Wales.	The Lancet 10 May 1986
	ABS 5	Same as above.	Same as above.
	JRV 5	Heart Disease Could be in Decline.	New Scientist, 8 May 1986
6	RAP 6	Association Between Transfusion of Whole and Recurrence of Cancer.	British Medical Journal, 30 August 1986
	ABS 6	Same as above.	Same as above.
	JRV 6	Bad Blood for Cancer Patients	New Scientist, 11 September 1986
7	RAP 7	An Outbreak of Meningocele Disease in Gloucestershire.	The Lancet, 6 September 1986
	ABS 7	Same as above.	Same as above.
	JRV 7	Meningitis check "being hindered".	The Times, 13 September 1986
8	RAP 8	The Relation of Apolipoproteins A-1 and B in Children to Parental Myocardial Infarction.	The New England Journal of Medicine, 18 September 1986
	ABS 8	Same as above.	Same as above.
	JRV 8	Heart Disease Link is Confirmed	The Times, 1 October 1986
9	RAP 9	Optic-Nerve Degeneration in Alzheimer's Disease	The NEJM 21 August, 1986
	ABS 9	Same as above.	Same as above.
	JRV 9	Eye-test clue to diagnosing Premature Senility	The Times, 23 August 1986

Group	Genre	Title	Source
10	RAP 10	Human Papilomavirus in Clinically and Histologically Normal Tissue of Patients with Genital Cancer.	The NEJM, 23 October 1986
	ABS 10 JRV 10	Same as above. Virus Theory in Doubt for Cancer	Same as above. The Times 23 October 1986
11	RAP 11	Is Relationship Between serum Cholesterol and risk of Premature Death from Coronary Heart Disease Continuous and Graded?	Journal of American Med Association; 28 Nov, 1986
	ABS 11 JRV 11	Same as above. Cholesterol: still risky at low levels.	Same as above. Newsweek, 8 December, 1986
12	RAP 12	Linoleic and Eicosapentanenous Acids in Adipose Tissue and Platelets and Risk of Coronary Heart Disease	The Lancet 24 January 1987
	ABS 12 JRV 12	Same as above. Why Essential Fatty Acids are Essential.	Same as above. 29 January 1987 New Scientist
13	RAP 13	Detection of Tissue Kilikrein in the Bronchoalveolar Lavage Fluid of Asthmatic Subjects	Journal of Clinical Invest January 1987
	ABS 13 JRV 13	Same as above. Peptides Implicated in Asthma Attacks	Same as above. New Scientist, 26 February 1987
14	RAP 14	A Progress Report on the Treatment of 157 Patients with Advanced Cancer Using Lymphokine-Activated Killer Cells and Interleukin-2 or High-Dose Interleukin-2 alone	The New England Journal of Medicine April 9, 1987
	ABS 14 JRV 14	Same as above. The Body Fights Back Against Cancer	Same as above. New Scientist, 16 April 1987

Group	Genre	Title	Source
15	RAP 15	Effects of Enalapril on Mortality in Severe Congestive Heart Failure	The NEJM 4 June 1987
	ABS 15	Same as above.	Same as above.
	JRV 15	Enzyme blocker offers relief for heart patients	New Scientist 18 June, 1987

3.2 CRITERIA FOR DATA IDENTIFICATION

Three major issues guided the identification of sources from which texts which make up the corpus for this study are based. They are: Representativity, Reputation and Accessibility.

3.2.1 Representativity

Since the study is based on a Corpus from a specific profession, texts which constitute the corpus must be seen to represent authentic discourse of that professional community. It must also represent authentic language for investigating the kind of problems identified in the research proposal. Finally the sources must be sufficiently varied to ensure that observations made are generalizable and not symptomatic of the 'housestyle' of a particular journal, magazine or newspaper.

Indications that the sources and texts of the professional articles (RAP and ABS) were systematically chosen to ensure a representative sample of the language of members of the medical profession is underlined by the following facts:

- (a) The texts are written by members of the medical profession . Although the authors are not usually identified by their titles, information from reprint requests at the bottom of the title page in most of the texts indicate that the author(s) are almost invariably specialists in various fields of the medical profession. For instance, in RAP 14, the information for reprint request identifies the title of the author to whom the application may be directed, thus:

"From the National Cancer Institute and Clinical Centre. National Institutes of Health, Bethesda, MD20892, where reprint requests should be addressed to Dr Rosenberg".

(b) Most of the authors of the texts work in university teaching hospitals, medical schools and research centres where they most probably engage in research, teaching and medical practice. Therefore, it can be said that all the authors to the texts used in this study are specialists in their fields. The authors of RAP 1 for instance, work at the Department of Experimental Pathology and Department of Obstetrics and Gynaecology, St Mary's Hospital Medical School, London.

(c) In accordance with what has become an established tradition in Bio-Medical research report writing, all the texts which make up the data for the professional articles are written by more than one author. The authors may be as few as two (in RAP 5) and as many as fifteen (in RAP 14). This suggests sufficient collaborative work, both in the process of experimentation and in the presentation of the "knowledge claims" (Myers, 1985a).

(d) As the names on the title page suggests, most of the authors of the texts are native speakers of English. This implies that they can be reasonably trusted to possess a very high intuition about how the language works which they can apply to communication in their profession.

(e) Finally, and more importantly, all the five sources identified in this study are professional, referred medical journals. Their number ensures that they are sufficiently varied to enable the selection of texts which would enhance the prospects of results obtained being made generalizable to related texts from other journals in the field. The process of referral and editing ensures that both in terms of knowledge claims and language use, the texts satisfy acceptable standards in the field.

The sources of the popular science texts (JRV) were also identified based on the need to ensure balance and representativity of popular science writings in the selection of texts. It has been pointed out elsewhere in this study (Chapter Two) that there exist different types of popular science writings. Part of the reason for the differences amongst the various types lies in the varying levels of popularization they exhibit, and consequently on the extent to which they assume an uninformed lay person as part of their audience. In other words, while all types of science popularization tend to employ a popular style of writing, some can be said to be more popular than others. This distinction is more directly illustrated by the distinction which exists between types of popular science writings found in newspapers and those in general science magazines. The need to ensure representativity of the different forms of popular science writings which relate to articles in the RAP and ABS necessitated the identification of one newspaper - *'The Times'* - One general interest magazine - *'Newsweek'* - and one general science magazine - *'The New Scientist'* - as sources for the selection of texts for the analysis of popular science writings. Of course, there are other other forms of science popularization not covered by this selection, eg. popularizations in the *Scientific American*. Part of the reason for choosing "*The Times*" as a source of my popular science texts is because it is the only paper in Britain that carries a daily science item outside its own internal lines of communication (Farago, 1976: 34) - that is, "The Times, Science Reports" column. The same reasons, with some modification, apply to the choice of "*Newsweek*" with its regular weekly "Science" section.

The "*New Scientist*" is an obvious choice for the selection of texts of the kind which match the popularized articles found in newspapers and general interest magazines - that is, articles written by science journalists or scientists who are not necessarily engaged in research. Articles which fall within this category were found in the "Science" section of the general science magazine - the "*New Scientist*". As has been pointed out elsewhere in this thesis (Chapter Two), the editor of the science section of the "*New Scientist*" Pallab Ghosh (personal communication) points out that

contributors to the science section lie somewhere between journalists and academics some of which have never been involved in research but are journalists who have followed a particular field for many years. The "*New Scientist*" also fits the earlier categorization of a form of science popularization which is written for a less popular audience than those found in newspapers or general interest magazines. Again, as Pallab Ghosh points out (personal communication):

Although, as editor of the section, my aim is to make the articles as accessible as possible to anyone interested in science, some of the articles will only be understood by those involved in the subject. Some, no doubt, will be lay people interested in keeping abreast of scientific developments. The majority of readers, however, will have scientific training to at least A level, will be science students, former science students or working researchers.

In summary, the particular sources used in this study were chosen because they are nationally circulated in Britain, because they represent science writing either labelled as such in popular science sources or presented on a regular basis in professional science journals and because they represent a broad spectrum of readership.

3.2.2 Reputation

Reputation has to do with the esteem in which members of an assumed readership hold a particular publication or group of publications. In the case of professional journals, reputation can be measured by, amongst other things, their levels of circulation and the rate of subscriptions by members of its target audience. In the case of newspapers, general interest magazines and general science magazines, it can be measured by the percentage of the readership of the "science" sections they publish which are students, academicians or people in scientific/technical occupations who form the bulk of their readership. Although no attempt has been made to measure these variables, the professional and popular science texts examined in this study are taken from highly reputable journals, magazines and newspapers (see Table 1).

3.2.3 Accessibility

This may be taken to refer to the ease with which texts which constitute the corpus of a particular study can be obtained by the researcher. It may also be construed to mean the extent of knowledge which the researcher has about the content of discourse in the texts being analysed. Both forms of accessibility are essential for an adequate analysis of discourse, albeit for different reasons.

The second form of accessibility is the subject of much controversy in ESP-related research. It is undeniably problematic, but not altogether impossible for an analyst to conduct an investigation into texts in a professional field in which he cannot claim specialist knowledge. One approach which has been suggested for overcoming this problem is the use of specialist informants (Selinker, 1979; Huckin and Olsen, 1984; Bley-Vroman and Selinller, 1984). While the use of SI's continue to remain an accepted practice in ESP-related research, its reliability has come under criticism in recent times. As Swales (1987: 125) points out:

Although Bley-Vroman and Selinker (1984) are clearly right to point out that, without specialist informants, analysts may well be in danger of "not knowing what they don't know", over-reliance on specialist informants may well produce the equally serious if opposite danger of analysts "believing all they hear". Further problems are that specialist informant work can be very time-consuming (Huckin and Olsen, 1984), and that it raises uncertainty when comparing texts from different disciplines (where instability is inevitably created by relying on different informants for each discipline).

Thus as Swales (op cit) concludes, the role of SI's can be seen only as partial and supportive as no controlled study has so far evaluated the use of informants "in terms of the trade-off between the rewards of 'behind the scenes' insight and professional confirmation and penalties of extra time and narrowed scope".

Since this study is based on an analysis of intuitively observable functional linguistic features of text ; since it is not concerned with making judgements regarding

the correctness or otherwise of the content of discourse in the texts, and since texts in the corpus are not restricted to one area of medicine, it was felt, in line with Swales that the use of subject-matter specialist informants on a formal scale might be time-consuming, might create uncertainties and instability in comparing texts from the different genres. However, informal discussions were held with medical practitioners at the Birmingham University Medical School on the fundamental principles underlying some of the medical topics contained in the texts.

A second approach is for the researcher to depend on his intuitions and knowledge about how language works and limit the analysis to intuitively observable linguistic features of text, not by counting the number of occurrence of items such as passives or verb forms, but by a qualitative account of the function which the linguistic features perform in text. For example, some linguistic features tend to indicate the beginning of a new schematic unit of text:

eg. In summary, a low concentration of linoleic acid ... (RAP 12)

In summary, this study has shown that ... (RAP 15)

The present report describes ...

some linguistic features tend to function to set the scene, thereby indicating the perspective from which the utterances are written.

eg. In 1983 two studies suggested that ... (RAP 2)

Formerly, this reflected differences .. (RAP 5)

This approach has its limitations, one of which is the fact that it tends to focus more on the function which language performs in the context, and therefore, it might not make highly detailed comments on the content of discourse in the texts. This limitation does not constitute a serious drawback, though, since the purpose of discourse analysis, and in particular genre-analysis, is not to analyse the correctness or otherwise of the cognitive content of text. That is a job for subject specialists and

editors of journals. This second approach has been the one on which much of the analysis in this study is based.

3.3 CRITERIA FOR THE SELECTION OF CORPUS

Having identified the sources for a possible selection of texts, steps were taken to assemble an appropriate corpus for the study. The process of corpus selection involved rational decisions over how to achieve comparability of texts across the three genres. Since the study is a qualitative rather than a quantitative one, what was aimed at was functional comparability. To this end, the major control on corpus selection was topic similarity across the three genres, where topic is taken to mean "subject matter" and topic similarity to mean the reporting of a research investigation with an accompanying Abstract in a professional medical journal and a parallel reporting of the same research report in a popular medium. In other words, the Research Article Proper (RAP), the Abstract (ABS) and the Journalistic Reported Version (JRV) had to report on the same research event, albeit in different situational contexts and for different purposes. Other measures which guided the selection of data include:

- (i) That the JRV must report on only one research event and not on a number of research investigations on a related medical problem.
- (ii) Again, the JRV must identify the source of the original research publication.
- (iii) Every RAP must have an accompanying 'Abstract' of not less than 100 words long.
- (iv) The ABS and the RAP must be written by the same authors, while the JRV would be written by a different author.
- (v) The texts need not necessarily be of the same length, but they all had to

report on the same research investigation.

3.4 THEORETICAL-METHODOLOGICAL FRAMEWORK

The discussion in this section will be concerned with the larger theoretical and methodological framework under which the analysis was carried out. To this end, theoretical issues, such as those related to an explanation of the notions of 'variation' and of comparative/contrastive discourse analysis adopted in this study would be provided. Also to be discussed is the basic discourse principle underlying this study and the linguistic theories employed in realizing it. This section will also examine procedures adopted in resolving a fundamental problem of research methodology - that of achieving comparability between text samples across the three genres in the analysis of thematic progression and cohesion.

3.4.1 Variation

A detailed review of 'variation in scientific discourse' has been provided elsewhere in this study (Chapter 2). That review was an attempt to critically examine possible approaches to the analysis of variation in scientific discourse. In this section, however, a discussion of a particular feature of language which has great relevance to an account of variation in discourse will be presented. That feature is **systematicity**. Systematicity relates to regularities which can be identified in language, and without which communication could not take place (Tarone, 1988:1). The regularities which exist in language constitute the differences between languages, thus accounting for the extent of systematic variation across them.

A central question which this study set out to answer is whether the differences and similarities between the three genres in the study is random or systematic. In other words, in what ways are there regularities in the way information is structured in one genre as opposed to another? The study seeks to answer this question within the 'form-function approach' with a view to accounting for the differences which exist when form and function are systematically related in the

organisation of information across related genres in different social contexts. This, in the main, is the view of variation adopted in this study. It is a view which concerns itself with descriptions of the regularities which occur in language in use and whatever speaker/writer knowledge may be inferred from these descriptions (Tarone, 1988:7). In order to ensure the realization of this notion of variation, the following steps are adopted in this study:

- (i) The research is limited to the analysis of language in use in one professional setting - medicine - and to related or parallel genres within this setting.
- (ii) Genres are compared on the basis of relatedness of the information they contain and the way it is organized based on inferences of writers' knowledge and assumptions, as well as readers' intuitive understanding of such knowledge.
- (iii) Genres are differentiated on the basis of medium of discourse or the social context of their production. That is, professional academic journals vs magazines and newspapers.
- (iv) The pattern of organization of information across the genres is differentiated by means of functional models of discourse structure - schematic structures, theme and cohesion.

3.4.2 The Comparative/Contrastive Framework

This study is constructed around the comparative/contrastive paradigm of discourse analysis.

The genres in this study can be said to be made up of related texts, with the ABS and the JRV representing developments from the RAP. The terminology used to

indicate this relationship in this study is the notion of "parallel texts". As has already been pointed out, the notion of "parallel texts" in a comparative/contrastive model of discourse analysis was put forward by Hartman (1980: 37). Hartman (op cit) categorizes parallel texts into three major groupings, thus:

- (i) Class A parallel texts which are "the result of a full-scale professional translation arrived at by conscious approximation processes in which the original message of the source-language text becomes a situationally appropriate target-language text".
- (ii) Class B parallel texts - "typically the result of a deliberate adaptation of a message in the respective conventions of two languages for the purpose of conveying an identical message to receivers of sometimes very different cultural background".
- (iii) Class C parallel texts - "typically unrelated except by the analyst's recognition that the original circumstances that led to the creation of the two texts have produced accidental similarities", for example comparing instances of a specific text-type across pairs of languages.

Although Hartman (op.cit) is talking about "pairs of languages" instead of "genres", it is to 'Class B parallel texts' that the notion of "parallel genres" used in this study is extended. Finally, the comparative framework upon which this study is based is one that recognizes the genres as being genetically related and one that attempts to account for that relationship in terms of a deliberate adaptation of the source message for the purpose of conveying an identical message to receivers of different cultural backgrounds.

3.4.3 The Discourse Framework

Central to this study is the notion of "discourse organization pattern". Two forms of discourse organization patterns are examined in the study - text structure and texture. Text structure refers to the global structure of the message form, while texture refers to the meaning relations which hold between parts of a text (Halliday and Hasan, 1985: 71). Although both patterns differ conceptually, they, nevertheless, are essentially related. A combination of both forms of discourse organization patterns is capable of providing meaningful insights into the nature of variation in text, especially since they are causally determined by contextual factors of text production.

Text structure is examined by means of a cognitive model of information processing - the schema-theoretic model. Schema-theory has been identified as one of the theoretical concepts that seem most useful to technical writers (Huckin, 1983). Swales (1981a) is one instance of a model of analysis based on the schema theory. Following from Swales (op cit), the texts in the corpus were analysed into hierarchical schematic units of information. Each unit in the schema is referred to as a **move**. Every 'move' is taken to embody a number of **constituent elements** which combine to form the information which make up a "move". Moves and their constituent elements are determined partly by inferencing from context, but mostly by reference to linguistic clues in the discourse.

Texture is approached from the point of view of two theories - the theory of **thematic progression** and the theory of **cohesion**. The notion of thematic progression adopted in this study is the one suggested by Danes (1970, 1974) within the framework of Prague School conceptions of "Functional Sentence Perspective" (FSP). In accordance with Danes (op cit) the texts in the corpus were analysed for relationships and differences in their patterns of thematic progression. Danes' three basic patterns - simple linear - constant theme and derived theme, as well as two others - TP with a split rheme and combinations of these - were examined in the corpus. The emphasis was on the identification of the utterance theme, since according to Danes, it is the theme and not the rheme that plays a more constructive role in pushing the

communication forward. Utterance themes were identified by means of inferences, identical lexical items, synonyms and paraphrases. In addition to these, considerations such as those involving subject/object distinctions, non-agent subjects and dummy subjects are also made in assigning thematic status to utterance sections. But above all, emphasis is placed on the notion of recoverability and that of givenness in the identification of theme. It is not unusual to employ all these approaches in the identification of an utterance theme. However, they are redundant in this case in the sense that an application of any one of them may be sufficient in revealing the theme of an utterance depending on the complexity of the information which is contesting for thematic status in the utterance. However, an overriding principle in Danes' model is recoverability of information from preceding context in the utterance following a particular point in text.

The analysis of cohesion is based on Halliday and Hasan's (1976) notion of 'Cohesion in English'. The analysis is limited to the main five categories listed in Halliday and Hasan - Reference, Ellipsis, Conjunction, Substitution and Lexical - as well as the number of sub-categories they embody. These cohesive properties of text are examined in the context of discourse in each of the three genres as a means of determining possible differences which contextual factors might have in the realization of cohesion across related genres.

3.5 PROCEDURES FOR THE ANALYSIS OF DATA

Different procedures were adopted for the analysis of the different types of organizational patterns analysed in this study. First, the analysis of schematic structures involve all the cognitive content in each text of the corpus. In other words, the whole text constitutes the domain of analysis in the investigation of the global structure of discourse in texts across the three genres.

The procedure adopted in identifying schematic structures in each text in the corpus consists of:

- (i) Focussing on the propositions in the texts to be analysed and identifying important information.
- (ii) Searching for linguistic clues such as function words, explicit phrases and expressions, verb forms, discourse markers (subordinators, connectives, exemplifiers etc), structural headings and sub-headings.
- (iii) Classifying and paraphrasing the context of discourse based on the linguistic cues.
- (iv) Assigning discourse functions to the overall information contained in segments of text as well as constituent elements of information in the segments.

By means of this procedure, it can be shown that discourse patterns may be made explicit by means of surface linguistic clues.

Second, the analysis of Thematic Progression involved whole texts from the ABS and JRV as well as systematically selected text extracts from the RAP. This procedure was adopted in order to limit the texts of the RAP to manageable lengths for investigating the kind of problems outlined in this study and ensure some form of comparability in the length of all three genres. In doing this, cognisance was taken of the need to ensure that the extracts selected were as representative as possible of RAP discourse. To this end, 120 paragraphs taken from the first two in the Introduction, Materials/Methods, Results and Discussion sections of the RAP were selected. The purpose of selecting paragraphs from the RAP was not to achieve absolute comparability in text length since the comparison being undertaken is not dependent so much on equality of text in terms of absolute length as it is on equivalence of subject matter and differences in purpose, medium and audience of discourse. In contrast to the RAP, all paragraphs in the ABS and JRV were selected for analysis. This was

done for the following reasons:

- (1) The ABS and JRV are short relative to the RAP. Any further reduction in text length might result in absurdity and failure to capture the objective of the study which is related to the effects of purpose, medium and audience on the organization of discourse across genres.
- (2) Most ABS texts do not have paragraph boundaries. Therefore, it would not be possible to select an equivalent number of paragraphs from them even if it were desirable to do so.
- (3) The paragraphs of the JRV texts are short compared to their RAP counterparts. Thus, even though it might be possible to select paragraphs of equal number to the RAP, they would still not have been equal in terms of length.
- (4) The ABS and the JRV do not have an explicitly demarcated organizational structure like the RAP. It would therefore not be easy determining which sections of the texts would be selected to ensure that the extracts are representative of JRV discourse. In view of these reasons, comparability was not based on the equivalence of texts in terms of length, but was achieved through the method of data selection which ensured that each triadic text group discussed identical subject matter.

The basic unit of analysis is an utterance, defined in Danes (1964: 229) as "one of all possible different minimal communicative units of the given language". Thus, it is possible to have an "utterance clause" and an "utterance sentence" or an utterance which combines two or more of each type.

The focus of analysis is the identification of utterance theme within the particular configuration "theme-rheme". This is in accordance with Danes conception of the role of theme in the organization of information in discourse. According to Danes

(1974: 113) "from the point of view of text organization, it is the theme which plays an important constructional role". The rheme, on the other hand, shows its significance as the conveyor of "new" information.

In the analysis, an utterance theme extends up to and includes the mass of information accumulated up to a certain point of text that is contained in the particular utterance following that point. The uppermost limit to which information may be deemed recoverable from "the mass of information accumulated up to a certain point in text" cannot be easily determined. Therefore, in principle it will be limited to information contained in a paragraph preceding the utterance being analysed. Of course, the exact limit to which thematic information may be deemed recoverable within text would also be constrained by factors such as paragraph length, topic continuity, topic shift and the nature of argumentation.

The analysis of **cohesion** is based on the same data as that used in the analysis of Thematic Progression. That is, systematically extracted texts from the RAP and whole texts from the ABS and JRV were used in the investigation of cohesion across the three genres. The same reasons which guided the decision to extract texts in one case and to use whole texts in the other in the analysis of TP also apply here. However, unlike in the analysis of TP, the basic unit for the analysis of cohesion is the sentence. In making the decision to base the analysis of cohesion on the sentence, it is recognized that there is hardly any agreement amongst linguists about what a sentence is.

Despite this fact, it is common knowledge that educated speakers of English (native and non-native speakers) have little or no problems with the identification of a "sentence" in written prose materials. For practical purposes, therefore, the sentence is identified in this study with that unit of meaning in written expository prose which begins with a capital letter following a fullstop and ends with a fullstop. The sentence is chosen as the basic unit of analysis, because cohesion is a semantic relation holding between sentences (Halliday and Hasan, 1976). In accordance with this, relations holding within sentences, such as those between clauses, propositions and utterances

are not considered in the analysis of cohesion in this study.

The model of analysis adopted is a modification of that suggested by Halliday and Hasan (1976: 340-355). It consists of the identification of cohesive ties according to (1) sentence number, (2) cohesive item, (3) cohesive type and (4) pre-supposed item. Not included are the number of cohesive ties between pairs of sentences, and the distance between identical ties. These were excluded because they would seem to be more relevant for investigating chains and continuity in text, which have been taken care of by the analysis of Thematic Progression. A tie is established between two items which occur inter-sententially, one being the presupposing and the other the presupposed item. At the end of the analysis of each text, the number of ties are counted and classified according to the five main cohesion types. The ties are not counted for the purpose of quantifying differences existing between genres in the number of ties they contain. In addition to counting, the items which are regularly used to achieve cohesion with respect to particular cohesion types are identified and their functions in the context of discourse in the texts determined. It is this latter procedure that is most relevant for characterizing variation due to cohesion in this study.

CHAPTER FOUR

VARIATION IN SCHEMATIC STRUCTURE

4.0 INTRODUCTION

In Chapter 2, an attempt was made at a review of possible approaches to the study of variation, particularly as they relate to patterns of discourse organization in scientific texts. Citing Swales' (1981a) seminal work on Article Introductions, it was suggested that the schema-based approach held great promise for the analysis of patterns of discourse organization in genres of scientific discourse. In accordance with this view, an attempt is made in this chapter to characterize the extent of variation which exists between parallel genres of written medical texts based on Swales' (1981a) schematic approach to the analysis of Article Introduction. The aim is to characterize the schematic structure of information in each of the genres with a view to:

- 1 Ascertaining how the pragmatic conditions of audience, purpose and medium influence the organization of information in the original or source genre - the RAP - into hierarchical information structures or schemata.

- 2 Ascertaining how under different pragmatic conditions of audience, purpose and medium, information is organized in the derived genres (ABS and JRV) into hierarchic schematic structures and how this relates to the incorporation of information from the source genre to the derived genres.

- 3 Ascertaining how the same pragmatic conditions as in 2 above influence the incorporation or information specific only to the derived genres within the framework of the schematic structures they manifest.

The schema-theoretic approach is only one of a number of proposition-based approaches to knowledge representation in cognitive psychology and artificial intelligence which have relevance to linguistic analysis of language in use. Its

relevance to linguistic analysis lies in the enormous potential it has for providing a viable framework for explaining how discourse in text is packaged into related, but distinct units of information which manifest characteristic linguistic forms. The linguistic forms suggest possible interpretations for arriving at a consistent description of text structure.

The schema approach including **frames**, **scripts** and **plans** differs from other cognitive theories of knowledge representation such as semantic networks and conceptual dependency which are primarily concerned with lexical and sentential level information. Unlike these other forms of representation, the schematic approach focusses on the organization of propositional and cognitive content. In other words, while the former are concerned with minute details, the latter is concerned with higher units of information. It is with these higher units of information that the study reported in this chapter is mainly concerned. The rationale for adopting the schema theory vis-a-vis other theories of information processing is succinctly put by Minsky (1975: 211), thus:

It seems to me that the ingredients of most theories both in artificial intelligence and psychology have been on the whole too minute, local and unstructured to account - either practically or phenomenologically - for the effectiveness of common sense thought. The 'chunks' of reasoning, language, memory and perception ought to be larger and more structured, and their factual and procedural contents must be more intimately connected in order to explain the apparent power and speed of mental activities.

In order that the discussion in this chapter may be put in perspective, we begin with an attempt at a brief discussion of the concept of schemata. This is followed by a discussion of the criteria for the identification of moves; a detailed outline of schematic structures in the RAP and a comparison of differences between the genres.

4.1 THE CONCEPT OF SCHEMATA

In cognitive psychology and artificial intelligence, the term schema is often thought to have been first used by Bartlett (1932). But, it has been suggested (Anderson and Pearson, 1984) that historical evidence for the use of the term ought to

be given to Gestalt psychologists, particularly to Wulf (1922/1938) who used the word in explaining the results to one of his experiments. Although the concept seems to have been introduced into psychology long ago, its full development as a model of knowledge representation is the result of much more recent researches into associated concepts such as "Frame" (Minsky, 1975), "Script" (Schank and Abelson, 1975), "Plan", (Schank and Abelson, 1977) as well as numerous studies on "Schema" by Rumelhart (1975, 1980), Rumelhart and Ortony (1977), Rumelhart and Norman (1978, 1987).

The term Schema and these associated concepts are not synonymous, but they are related in the sense that they all function to account for holistic or higher-order knowledge representation. Rumelhart and Norman (1987: 35) define schema as "data structures for representing the generic concepts stored in memory". Anderson and Pearson define it as "an abstract knowledge structure". A schema is abstract in the sense that it tends to summarize how knowledge is organized in a group of related cases or events which may differ with regards to details. It is structured in the sense that it indicates the relationships which exist between the component parts of a case or an event. Because they are structured, schemas have characteristic features by which they can be identified. These features are fully developed in Rumelhart (1980, 1981), Rumelhart and Ortony (1977). Some of the features include:

- Schemas have variables
- Schemas can embed one within the other
- Schemas represent knowledge at all levels of abstraction
- Schemas represent knowledge rather than definitions
- Schemas are active recognition devices whose processing is aimed at the evaluation of their goodness of fit to the data being processed.

Briefly, schemas are like "conceptual envelopes" that contain variables. The concepts that constitute a schema have got a part that is fixed, or as it were obligatory,

and which can be extended at all instances of the concept. In addition to the fixed part, there is also a variable or optional part. For instance, in determining the schema for a research article, the surface organizational features - Introduction, Materials and Methods, Results and Discussion - may be taken as fixed while the parts as in Swales 'Move' schema in article introductions may be regarded as an optional schema within the IMRD schema.

Secondly, a schema consists of sub-schemas which in turn may consist of sub-schemas. Thirdly, schemas exist to explain all levels of experience and at all levels of abstraction. That is, schemas exist for all manner of knowledge, from the grand to the very small. Finally, schemas are active as opposed to passive processes. Therefore, they are constantly in the process of "evaluating their goodness of fit, binding its variables, and sending messages to other schemas that indicate its current estimate of how well it accounts for the current data". (Rumelhart and Norman, 1987: 38). Rumelhart and Norman (op cit) distinguish between two data sources that a schema can use in evaluating its goodness of fit as:

- Information provided by the schema's sub-schemas on how well they account for their parts of the input (bottom-up information).
- Information from those schemas of which the schema is a constituent about the degree of certainty that they are relevant to structuring the input (top-down information).

The process of interpreting a schema and determining its goodness of fit would involve an interaction of both "bottom-up" and "top-down" information processing.

Further to the features discussed above, the distinction has also been made between "formal schemata" and "content schemata" (Carrel, 1983: 83). Formal schemata refers to the background knowledge which human language processors possess about the formal, rhetorical, organizational structures of different kinds of

texts. It includes knowledge on "information about, and expectations of, differences among structures - eg. differences in genre, differences in the structure of fables, simple stories, scientific texts, newspaper articles, poetry etc". Thus, it is related to the obligatory or fixed part of schemas. Content schemata, on the other hand, refers to the background knowledge about the content area of a text or text-types - eg. Abstracts, Research Articles, Science popularizations, Instruction manuals, Newspaper editorials etc. It is, therefore, related to the variable or optional part of schemas.

The discussion in this section points to the relevance of the concept of schema to the understanding of discourse. As Rumelhart (1980: 47) points out, "the process of understanding discourse is the process of finding a configuration of schemata that offers an adequate account of the passage in question". Clues from the passage would suggest possible interpretations (instantiations of schemata) that are then evaluated against the successive sentences of the passage until a consistent evaluation is discovered. One linguistic investigation which has applied this process to the analysis of the formal schemata of a genre of scientific discourse is Swales (1981a), validated by Crooks (1985). As with Swales (op cit), the study reported in this chapter is concerned with characterizing the formal schemata of texts in the corpus. However, the study is not about schema theory, but about the application of a schema-related model of discourse analysis, referred to as "genre-analysis" to the analysis of three genres of written medical discourse.

4.2 THE SURFACE ORGANIZATION OF INFORMATION IN THE GENRES

As a prelude to the analysis of schematic structures in the genres, it would be proper to examine their surface organizational features. To this end, an attempt is made in this section to compare the genres in terms of their visual and iconic characteristics. The discussion will begin with an examination of the surface characteristics of the RAP, which for the purpose of comparison is regarded as the

norm. It is followed by an examination of the features of the ABS and a comparison of these features with those in the RAP. Finally, the features of the JRV are examined and compared with the RAP and the ABS where necessary.

4.2.1 The Research Article Proper (RAP)

The RAP has a fixed format for the presentation of the information it contains. As has already been stated in Chapter 2, biomedical research papers are usually organized into four main sections (O'Connor et al, 1976), thus:

- Introduction
- Materials and Methods
- Results
- Discussion

This is the traditional IMRD format for the presentation of information in most scientific research papers. Thus, the format can be said to represent the "empirical repertoire" of most scientific investigations (Gilbert and Mulkey, 1984). As Regent (1985: 111) points out, the IMRD organizational format is recommended by all medical journals and followed by all writers of original articles. The findings of Adams-Smith (1984: 27) lend weight to this contention. Adams-Smith identifies three major types of medical journal papers:

- Clinical case reports
- Research papers
- Editorials.

She points out that of these three, the research paper is the only one with a standard organizational plan. In accordance with the views of Regent and Adams-Smith above, all the RAP texts examined in this study have the four-section standard format for

research papers in medical research journals. However, not all the texts use the same terminologies to designate information contained in all the sections. Noticeable changes were observed within the texts in the phrasing of the "Materials and Methods" section. For instance, in two texts (RAP 3 and RAP 6 obtained from *The Lancet* and *The British Medical Journal* respectively), the phrase "Patients and Methods" was used instead of the standard phrase "Materials and Methods". One text (RAP 8 from the *British Medical Journal*) uses the phrase "Subject and Methods" instead. The remaining twelve texts simply use the lexical item "Methods", omitting either "subjects", "patients" or "materials". There does not seem to be any identifiable criterion to suggest the appropriate situation in which one type of heading may be used in preference to the other. What these observations made above point to is that medical research papers do not necessarily designate information in the Methods section using exactly the same headings as those commonly found in most scientific research papers. Having said this, it would seem that author's perceptions of the established conventions of particular journals guided the labels they gave to the methodology section in the texts examined. Again, it does also appear that the nature of the researches themselves tended to influence the captions given to the methods section. Since most of the researches reported in the texts made use of human beings as "subjects", it is only reasonable that they should be identified either as "subjects" or "patients".

The RAP texts exhibit a number of typographical markers. In virtually all the texts, indentations are used to indicate new typographical paragraphs. Capital letters are used, especially in the initial word or phrase of a text to signal the commencement of discourse or highlight the main topic of discourse in the introduction section of the texts (see RAP.1, RAP.7, RAP.8, RAP.9, RAP.10, RAP.11, RAP.14, RAP.15 in Appendix I). The capital letters are usually the first lexical items of the first sentence in the Introduction sections of the texts. Thus, in modified subject NPs, the modifying articles, adverbials and adjectivals tend to be capitalized. In unmodified subject NPs, on the other hand, the capitalized lexical items tend to be the first lexical item of the

noun which functions as subject. The former is illustrated by the example from RAP.9 and RAP.14.

- (a) SINCE 1940, when there were over 12000 notifications in England and Wales, meningococcal disease has been much less common (RAP.9).
- (b) THE relation between dietary fatty acids and coronary heart disease (CHD) has been investigated since Sinclair raised the possibility that a deficiency of essential fatty acids (EFA) might contribute to atherosclerosis (RAP.14).

The latter is illustrated by these examples from RAP.15 and RAP.1.

- (a) CONGESTIVE heart failure is a common condition reported to affect 1 per cent of the population; with an annual incidence of approximately 3 per 1000 (RAP.15).
- (b) RECURRENT spontaneous abortion has been treated by means of immunisation with parental cells and cells from multiple unrelated donors (RAP.1).

The consistency with which capital letters are used with the initial lexical items of the first sentence of the texts suggests that those items which are so indicated reflect the author's intended emphasis. Therefore, they would seem to function to highlight the lexical item or the concept it represents to the reader.

Another typographical feature of the RAP is the use of bold prints to identify the major section headings - Introduction, Patients and Methods, Results, Discussion. The texts also contain a lot of sub-headings. 80% of the sub-headings in the texts are written in italics. Most of the sub-headings are in the "Patients and Methods" and the "Results" sections. For instance, the "Patients and Methods" section of RAP.I is

divided into:

- (a) Patients
- (b) Sequential trial
- (c) Cytotoxic antibody testing

The "Results" section of RAP.7 contains the sub-headings:

- (a) Epidemiology
- (b) Clinical Features
- (c) Microbiology
- (d) Management of the Outbreak.

It would appear that while it is almost always the case that the "Patients and Methods" section has to be divided into sub-headings, the use of sub-headings in the "Results" section is entirely at the discretion of the writer. This is supported by the observation that while 73.4% of the texts have sub-headings in the "Patients and Methods" section, only 33.4% of them have sub-headings in their "Results" section. Sub-headings are conspicuously absent in the "Introduction" section and marginally present in the "Discussion" section. For instance, only one of the texts - RAP.2 - has sub-headings in its "Discussion" section. This might mean that sub-headings are not significant features of the Discussion section of Medical Research papers.

The paragraphs in the RAP are relatively long. There are on average about five or six sentences in a typical RAP paragraph. This number of sentences per paragraph may not present a clear picture of the length of paragraphs in an RAP text. However, viewed against the background of a high mean sentence length in RAP texts - 25.84 words per sentence - the paragraphs can be said to be lengthy.

4.2.2 The Abstract (ABS)

An abstract has been described as a structured, concise, accurate and coherent

representation of information in the original document (Cremmins, 1982). This would seem to suggest that abstracts faithfully mirror the form and content of their originals. However, as the Abstracts in this study reveal, this is not necessarily the case. A look at the Abstracts shows that while they can be said to mirror the content of their originals, they do not in any way represent the surface form of these originals. A significant feature of difference between the RAP texts and the ABS texts in this study is the fact that the ABS texts are not marked by an explicit surface organizational format. That is, the ABS texts do not exhibit explicit section headings of the IMRD type found in the RAP. This notwithstanding, there appears to be a deliberate tendency for the Abstracts to follow closely the information content of their RAP equivalents.

A distinctive characteristic of abstracts in this study is that they can be explicitly identified. All the Abstracts examined in this study bear the heading "Abstract" or "Summary" written in bold print at the top of the texts. Thus, the title of the Abstract differs in this unique way from the main title of the whole text of which the Abstract is part. While the title of the whole text indicates the content of discourse in the text, that of the abstract functions to define the genre.

Just as Abstracts in this study do not have section headings, they also do not have sub-headings. The only means by which information is delineated on the surface of text is by typographical indentations which are used to indicate the commencement of physical paragraphs. However, as observations from the texts show, even this may not be a common feature of abstracts. A look at the abstracts in the study shows that 60% of them do not delineate text explicitly into paragraphs. The suggestion would seem to be that the occurrence of explicit paragraph structures in the ABS is determined by the length of argumentation and the amount of information from the original RAP texts they embody. For instance, RAP.6 and RAP.15 with 314 words and 305 words respectively have clearly demarcated paragraph structures. On the other hand, RAP.7 and RAP.3 with 108 words and 127 words respectively are not structured into paragraphs. The implication of this observation is that abstracts accompanying research

articles, particularly in medical journals rarely organize information into paragraphs. This, therefore, tends to be another area of difference between the ABS and its RAP original.

It has been pointed out that the Abstracts in this study tend to mirror the information contained in their original RAP texts. In other words, the abstracts do not contain any information that has no equivalents in their RAP originals. In ABS.2 for instance, the information contained in the first sentence below:

"The relation between use of oral contraceptives (OCs) by young women and their risk of breast cancer before 45 years of age was investigated by analysis of data from a population-based, case control study conducted in eight geographical regions of the United States".

relates to that in RAP2. below:

We have now investigated the relation between OC use and risk of breast cancer in women under 45 by analysing data from the Cancer and Steroid Hormone Study of the Centres for Disease Control (CDC) and the National Institute of Child Health Human Development.

The examples above point to some fundamental differences between the ABS and the RAP. First, is the change of tense from the active in the RAP to the passive in the ABS. Thus, the expression, "we have now investigated the relation between OC use and ..." with an agent and an active verb form from the RAP is rewritten into the expression, "The relation between use of oral contraceptives (OCs) by young women..." without an agent as subject but with a transformation from an active to a passive verb form. Secondly, the example indicates that there is no one-to-one correspondence between the information presented in the ABS and that in the RAP. For instance, the nominal group, "Cancer and steroid Hormone Study Centre for Disease Control (CDC)" in the RAP gets replaced by the much shorter phrase "Case Control Study" in the ABS. These examples would seem to suggest that the writing of an abstract is not necessarily a case of the direct transfer of text, but a process involving syntactic transformations and information reduction.

Earlier in this section, it was pointed out that the abstract is a concise

representation of information in the original document. The extent to which an abstract has to be concise to be adjudged to have mirrored sufficiently well information in the original is not exactly known. What is obvious from an examination of the texts in this study is that a major method for achieving conciseness in the ABS is the tendency to combine two or more units of information from the RAP into a single complex sentence. This is illustrated by the example from ABS.12 below:

The relation between the fatty-acid composition of adipose tissue and platelet coronary heart disease (CHD) was examined in a case-control study of new angina pectosis (AP) and first acute myocardial infarction (AMI)".

Two functional units of information may be identified in the example above. The first unit - "The relation between the fatty-acid composition ... was examined" identifies the objective of the study, which can be said to be an examination of "the relation between the fatty-acid composition of adipose tissue and platelet membranes and ... coronary heart disease". This information is directly related to that contained in the first sentence of the Introduction section of the RAP presented below:

The relation between dietary fatty-acids and coronary heart disease (CHD) has been investigated since Sinclair raised the possibility that a deficiency of essential fatty acids (EFA) might contribute to arteriosclerosis".

The second unit - "... in a case-control study of new angina pectosis (AP) and first acute myocardial infarction (AMI)" refers specifically to the method of investigation. It identifies very closely with information contained in the last sentence of the Introduction section of the RAP text shown below:

These findings of low adipose linoleic acid in relation to CHD between populations and within a population indicated the need for a formal case-control study of angina pectosis (AP) and acute myocardial infarction (AMI)

The examples above point to intuitively obvious methods used in the abstraction of information - extraction and combination. They may also serve to

account for the relatively higher, albeit insignificant mean sentence length in the ABS. Compared with the RAP which as we have seen has a mean sentence length of 25.84 words per sentence, the ABS has an average of 26.32 words per sentence. This difference, though marginal, may be due to the tendency in the ABS to extract as much relevant information from the RAP as possible and combine them into one complex sentence. There exist other methods for achieving conciseness in an ABS text. But, it would appear that the use of complex sentences is the only one that has direct influence on the surface organization of information in the texts because it tends to play a significant role in the conceptual structuring of information.

Despite the differences in the surface organization of information in the ABS and the RAP pointed out in this discussion, both genres share a number of things in common. First, they share a common general title. Second, they are written by the same authors. However, they are written with slightly different intentions and the information they contain would seem to be of interest to slightly varied audiences. Abstracts serve the primary purpose of assisting readers of scholarly and technical monographs and journals to decide whether they should consult the full text of the material that is abstracted. Other potential audiences of the abstract include editors and information specialists such as indexers, information-retrieval analysts and lexicographers.

The differences and correspondences in the surface organization of information between the ABS and the RAP pointed out in this section have implications for the deeper and more important issues of text structure to be examined later in this chapter and the chapters following it. The discussion would seem to suggest that there are more differences than correspondences between the ABS and the RAP at the level of the surface organization of information.

4.2.3 The Journalistic Reported Version (JRV)

The Journalistic Reported Version (JRV) represents the category of texts

referred to as the "popular-scientific-style" (Glaser, 1982: 76) and "the narrative of nature" (Myers, 1986: 3). It is designed to disseminate scientific knowledge "among a general audience composed of non-specialists"; people generally referred to as "intelligent laymen", "the uninitiated" or "people in all walks of life" (Glaser, 1982).

Like the ABS, but unlike the RAP, the JRV has no conventionally prescribed standard organizational format. Pallab Ghosh, editor of the "New Scientist" claims (personal communication), that a standard organizational format would stifle the creativity and imagination needed to get across the excitement of scientific discoveries, the creative process involved in making such discoveries and the controversy that surrounds them. Therefore, the JRV does not follow the IMRD format of the original RAP text in presenting information. For instance, it would seem that unlike the RAP which normally begins with an 'Introduction' in which some background information and the purpose of the study are almost invariably stated, the JRV text begins with what appears to be a summary of the issues to be discussed in the text. Below are examples to illustrate this form of surface difference between the JRV and the RAP.

THE RISK of getting breast cancer before age 45 does not alter significantly for women who started using oral contraceptives (OCs) before they were 20 and went on doing so for more than four years.

(JRV.2: *New Scientist*, 7 November 1985)

The example above is the first sentence of JRV.2. It represents a summary of the results of the research being reported while at the same time providing an expansion to the title of the text, which is "Pill use and the risk of breast cancer". Its purpose would seem to be to express the main message of the text and expand upon the message contained in its title.

A look at the first sentence of the RAP equivalent of the JRV example cited above suggests that it represents an attempt to review previous research as can be seen below:

In 1983, two studies suggested that the use of oral contraceptives (OCs) by young women increased their risk of breast cancer before 45 years.

(RAP.2: *The Lancet*, Nov 2 1985)

It is interesting to note that while the JRV and RAP texts above differ in functional terms, they tend to be closely related in their choice of core lexical items. Such lexical items as "risk of breast cancer", "oral contraceptives", "young women", "45 years" occur in both example. However, the difference in tense forms between both texts from the two genres - the JRV examples using active while the RAP uses passive verb form as well as the use of the time adverbial "In 1983" to indicate a temporal frame and the explicit lexeme "two studies" to refer to previous studies - would seem to account for the difference in discourse function which exists between them. Surface differences of this kind give strong indications of the possible influence of audience and purpose of discourse on the wider issues of discourse organization which will be discussed in this and later chapters of this thesis.

Another area of significant difference between the JRV and the RAP is in the way both genres present their titles. As Myers (1986: 8), citing Bazerman (1985) points out, titles are crucial, for they indicate what authors or editors think will arrest the eye of the typical reader skimming the title page. The difference in the way titles are written in the JRV and the RAP make an interesting study, particularly since parallel texts of both genres in this study are supposedly reporting on the same research event.

The first superficial observation to be made with regards to parallel texts in both genres is that they differ in terms of title. First, the JRV favours short, simple titles. For example:

JRV.3: Aspirin prevents hypertension in pregnancy.

JRV.6: Bad blood for cancer patients.

These contrast with the lengthier titles of their RAP equivalents.

RAP.3: Low-dose aspirin prevents pregnancy-induced hypertension

and pre-eclampsia in Angiotensin-sensitive primigravidae.

RAP.6: Association between transfusion of whole blood and recurrence of cancer.

The short, simple titles of the JRV are eye-catching and would seem to have the effect of arresting the attention of the reader, particularly those of the readership who may be classified as "laymen".

Related to differences of length of titles is the amount of information contained in the propositions which make up the titles. Titles of JRV texts examined in this study tend to focus only on one medical problem around which the discourse in the rest of the text is built. The medical problem usually forms the first element of the title. The problem is most often a disease or an ailment which is considered to be at the centre of current medical discussions and therefore of great interest to most people, including the lay reader. The entire propositions embodied in the titles tend to suggest exciting prospects that the dangers posed by these diseases may soon be a thing of the past. The title may also be couched in such a way as to attract attention by causing indignation amongst uninformed readers. Examples of the discussion so far are provided below:

JRV.1: Immunisation can present miscarriage.

JRV.5: Heart disease could be in decline.

JRV.9: Meningitis check "being hindered".

First, all three titles focus on one medical problem which could either be "Immunisation" or "Miscarriage" in the first, "Heart disease" in the second and "Meningitis" in the third title. These contrast sharply with titles of some texts in the RAP which tend to focus on more than one problem at a time. For example:

RAP.5: Infant mortality, childhood nutrition and ischaemic heart

disease in England and Wales.

RAP.2: Oral contraceptives and breast cancer in young women.

Second, as the JRV.5 and JRV.9 examples show, the medical problems or diseases are the first elements of their respective titles. And, as the title of JRV.9 shows, the title of a JRV text is capable of attracting attention by causing indignation amongst readers. The main attraction for a lay reader encountering such a title would be to find out who or what is hindering the check of meningitis, a highly dreaded disease that has claimed many lives in one part of the country (Gloucester and Stroud) and that is threatening to spread to other parts as well. This would be followed by indignation at why anyone would wish to "hinder" the check of such a dreadful disease. Of course, the author is conscious of the effect he wishes to achieve in his assumed audience and so, encloses the crucial aspect of the title in quotation marks. The use of quotation marks in that title has two functions: a reliability function and an attitude function (Weizman, 1984). First, it assures the reader of the accuracy of the quotation and provides an assurance of the author's reliability. Second, the use of quotation marks to mark out the phrase "being hindered" in that title would seem to suggest that the writer does not want to assume any responsibility for the opinion of the source from which the quotation is extracted.

A related development to the differences in terms of the amount of information contained in the titles is differences in levels of abstraction and complexity of the sentences which make up titles in both genres. Although titles in the JRV have been characterized as short and simple, they are nevertheless, clear and precise. Their simplicity derives from the fact that they address only one issue at a time. They, therefore, tend to make less use of coordinators and subordinators, but instead are straightforward cases of simple expressions with simple subjects and predicates. Titles in the RAP on the other hand tend to be pitched at a high level of abstraction, with complex propositions containing complex nominal expressions. The abstractness of the titles in the RAP derives from the nature of the subject matter and the way the title

is phrased. For example, the title of RAP.2, "Oral contraceptives and breast cancer in young women" is not a real representation of information contained in the main text, which may be taken to mean that oral contraceptives cause cancer in young women. But as a look at the text shows, the study is an investigation into the relation between OC use and risk of breast cancer in women. The idea of "a relationship" that is risky and capable of causing breast cancer is not explicitly expressed in the title of RAP.2 above. In contrast, the title of JRV.2 clearly expresses this idea in its title, which reads: "Pill use and risk of breast cancer". In this regard, it would seem that the title of the JRV text is more explicit, precise and less abstract than its RAP equivalent. On the other hand, titles of RAP texts tend to present complex and abstract information which, though objective, may not reflect precisely the nature of the argument inside the text. However, the RAP reader can cope with the abstract titles because he shares the same level of knowledge of the topic with the author.

Titles in the RAP are not only stated at higher levels of abstraction compared to their JRV equivalents, they are also expressed by means of abstract nominal forms. For example:

RAP.1: Controlled Trial of Treatment of Recurrent Spontaneous Abortion by Immunisation with Paternal Cells.

RAP.5: Infant mortality, Childhood nutrition and Ischaemic Heart Disease in England and Wales.

The abstract nominal forms in the titles above, some of which express verbal as opposed to physical action are: "Controlled trial", "Recurrent Spontaneous Abortion", "Infant Mortality", "Childhood nutrition", "Ischaemic Heart Disease". These are timeless entities, they take the events out of the concrete or specific domain and place them in the general, abstract world. Kress (1983: 129) points out that abstract nominal forms enable events to be "assimilated to a pre-existent paradigm, or ideology" which in the case of RAP texts can be said to be the conventions of the discipline of science or

the conventions of "the narrative of science" (Myers, 1986) - a form of narrative which as has been pointed out elsewhere in this thesis, emphasizes "in its syntax and vocabulary the conceptual structure of the discipline".

The JRV titles, in contrast, tend to make use of non-abstract nominal lexical items. For example:

JRV.2: Pill use and risk of breast cancer.

JRV.3: Aspirin prevents hypertension in pregnancy.

JRV.4: Epstein-Barr virus finds a new route to cancer.

JRV.6: Bad blood for cancer patients.

The non-abstract lexical items which are central to the titles are "Pill use", "Aspirin", "Epstein-Barr virus", "Bad blood". They are concrete entities which help to keep the titles within the specific domain of the real world. Each of the items occurs in association with a real life problem - "Pill use"/Breast cancer, Aspirin/pregnancy, Epstein-Barr virus/cancer, Bad blood/cancer patients. Thus, they enable events in the JRV to conform to paradigms of the real world.

Another mark of difference between titles in the JRV and those in the RAP is the tendency for titles in JRV texts to have verbs and for RAP titles to have verbless propositions. For example:

JRV.12: Why essential fatty acids are essential.

JRV.13: Peptides implicated in asthma attack.

JRV.14: The body fights back against cancer.

These contrast with titles in the RAP equivalents of the JRV texts above which are verbless propositions.

RAP.12: Linoleic and eicosapentaenoic acids in adipose tissue and

platelets and risk of coronary heart disease.

RAP.13: Detection of tissue Killikrein in the Bronchoalveolar Lavage Fluid of Asthmatic subjects.

RAP.14: A progress report on the treatment of 157 patients with advanced cancer using Lymphokine-activated killer and interleukin-2 or high-dose interleukin-2 alone.

The effect of the verblessness of the propositions in the titles of RAP texts is the abstractness and the tendency for complexity and abstractness discussed above, for without an active verb, it might be difficult to assign an accurate interpretation to the titles prior to the reading of the texts. On the contrary, the use of active verb forms in the JRV titles makes for an easy and quick interpretation of the titles and gives the reader an objective insight into the content of the text.

The difference in the surface organization of information in the JRV and the RAP is also reflected in the functions which titles tend to perform in the genres. One of such functions is to indicate the approach which the author(s) intend to adopt in the presentation of argumentation in the text. As has been discussed (Chapter 2), four possible approaches exist for the presentation of medical information in the media (Karpf, 1988: 9). In brief, they are:

- (1) The medical approach, celebrating medicines curative powers.
- (2) The consumer approach, criticising the doctor-patient relationship.
- (3) The look-after-yourself approach, appealing for changes in individual behaviour.
- (4) The environmental approach, stressing the social origins of illness.

Titles of JRV texts examined in this study would seem to function to indicate one or the other of these approaches. For instance, the title of JRV.10 "Virus theory in doubt for cervical cancer" appears to be indicating that the discourse in the text would

be an attempt to either criticize or reverse popularly held views about the relationship between virus theory and cervical cancer. This fact is attested to by the lexical item "in doubt" used in the title. A rewritten version of the title could be:

"The theory that cervical cancer is caused by a virus is in doubt".

Underlying this version is information from the text itself:

- (a) Revised version about links between the virus found in genital warts and cancer of the cervix are proposed in research results published today by the Medical Research Council's Institute of Virology, Glasgow.
- (b) The scientists whose work is published in the latest issue of the New England Journal of Medicine believe too much emphasis has been placed on the fashionable virus theory to account for the epidemic sweeping the western world.

Following from these examples, the title of JRV.10 would seem to be indicating that the discourse in the text would be tending towards the consumer approach in the sense that it might be aimed at pointing out to the uninformed reader about the contradictory nature of medical researches. Another title which indicates the consumer approach is that of JRV.7 - "Meningitis check" "being hindered" - which as we saw earlier on in this discussion is capable of causing indignation amongst readers.

The title of JRV.11 - "Cholesterol: still risky at low levels" - would seem to be functioning to indicate the "Look-after-yourself" approach. It indicates that the discourse in the text is likely to be concerned with evidence to warn people against the danger of cholesterol, even when it is at low levels in the blood. This is confirmed by information from the text which refers to the observations from the research on which the JRV text is based as "alarming", thus:

The Journal of the American Medical Association offered an alarming answer last week. Four out of five middle-aged men in America run the risk of dying prematurely from a coronary - including many

with cholesterol levels most doctors would consider only moderately elevated, JAMA reported in a special issue on the subject.

A large number of the titles of JRV texts would seem to function to indicate the medical approach. For example:

JRV.14: The body fights back against cancer.

JRV.1: Immunisation can prevent cancer.

JRV.13: Peptides implicated in asthma attack, etc.

All these celebrate medical breakthroughs and suggest to the readers that through the **ingenuity of medicine, a cure has been found to some of man's most depressing ailments.** Statements from the texts suggest this view: For example, the statement from JRV.13 reads:

Many people will have wheezed and spluttered their way through the recent cold spell, victims of respiratory diseases. They are probably yearning for the long, hot days of summer - but warm weather brings its share of hazards. Deep breaths of pollen-laden air can also result in problems. For thousands of asthma sufferers, pollen from grass and other plants triggers an allergic reaction ...

... Asthma sufferers can breathe a little easier now because a team of doctors in Baltimore may have identified the missing mediators .
(*NewScientist*, 26 February 1987)

In contrast, titles in all the RAP texts in this study would seem to function to indicate the medical approach. That is, they are primarily concerned with presenting information related to the objectives of the study in accordance with the conventions for the presentation of information in medical academic publications, and not with the controversy that might have surrounded either research into that specific problem or to controversies surrounding previous researches related to it. In other words, they are less concerned with "celebrating the curative power of medicine" in a sensational manner, but more with celebrating results of medical research and the conventions of research publications. This is a major modification to Karpf's (op cit) characterization of the "medical approach", which in her study is concerned with the media (radio and television) and not with academic publications.

Finally, the JRV texts examined in this study do not contain accompanying visuals such as graphs, tables or figures. They may, however, contain photographs or pictorial illustrations. But, even this latter fact does not seem to be a feature of the JRV texts in this study, for only one of the texts - JRV.14 does contain a photograph. The low occurrence of photographs or pictorial illustration in my data would seem to contradict Myers' (1986: 17) view that pictorial illustrations "are a large part of the popular magazines appeal to the casual reader". However, the contradiction does not in any way nullify Myers' claims. Rather it strengthens them by pointing out that variations can exist even within popularized texts based on form and subject-matter or discipline. Therefore, the relative contradiction may be due to the fact that Myers' popularizations are taken from the *Scientific American*, which tends to have pictorial illustrations, while the popularizations in this study are mainly from the *NewScientist*, which tends not to make use of pictorial illustrations.

In contrast to the popularized texts in this study, professional RAP texts embody graphs, tables and figures. This, is an important surface feature difference between the RAP and the JRV in this study.

An attempt has been made in this section to examine the differences between the genres in this study in terms of the superficial elements of discourse organization. The differences identified between the genres underline the fact that despite similarities in subject-matter, parallel genres can differ on the surface due to the differences that exist between them with regards to audience, purpose and medium of discourse.

4.3 THE SCHEMATIC ORGANIZATION OF INFORMATION IN THE GENRES

4.3.1 The Identification of Discourse 'Moves'

An interplay of linguistic and psychological principles is employed in this chapter to account for the organization of information in the three genres under investigation in this study. To this end, principles embodied in the cognitive psychology "schematic" models of information representation are employed to account

for the segmentation of information in the genres into functional units of discourse referred to in this study as "Moves".

The term 'Move' has been used in discourse analysis in connection with the structure of information in classroom interaction (Bellack, 1966; Sinclair and Coulthard, 1975), in the analysis of conversation to refer to the choices of speech acts (Edmonson, 1981) and in genre-analysis (Swales, 1981a) to characterize the structure of information in Article Introductions. As in Swales, (op cit) the word 'Move' is used here to refer to a unit of discourse structure which presents a uniform orientation, has specific structural characteristics and has clearly defined functions, as shown in the example below:

Having achieved good results with our preliminary attempts at immunization, we decided to carry out a controlled trial of treatment with paternal cells. We wanted to find out whether the success was truly immunological or a placebo effect of a positive treatment, such as has been found with other treatments for recurrent abortion. We also wanted to study the possibility that immunisation with paternal cells might produce subfertility or other unwanted side-effects. The trial started in February 1982, lasted 3 years. Participants were selected from women attending a recurrent-abortion clinic.
(RAP.1: *The Lancet*, 27 April 1985)

A "bottom-up" analysis in which the communicative function of information in text is established through the processing of its constituent words, sentences and propositions would seem to suggest that the text above functions to "Highlight Present Research". This contention raises the questions: (1) Is there linguistic evidence in the text which supports this interpretation? (2) If there is, what form does it take. These questions may be extended to all instances of 'Moves' identified in all the texts of the genres in this study. In view of questions such as these, an attempt is made in this section to present the principles and procedures adopted in the identification of 'Moves' in this chapter.

As indicated above, a "bottom-up" approach was adopted in determining the information content and boundary of moves. To this end, an attempt has been made to establish the characteristic linguistic features of each move by an examination of their

constituent words, sentences and propositions. Five main linguistic features tend to signal the functions of discourse in the moves and indicate their boundaries. They are: (1) Explicit prefacing and preparatory statement; (2) Explicit lexemes; (3) Discourse conjuncts; (4) Grammatical specifiers of content relations; (5) Summary statements.

4.3.1.1 Explicit Prefacing or Preparatory Expressions or Statements

Preparatory or prefacing expressions are signals which reveal information abstracted from content occurring later in the text. It uses the same words or paraphrasing to state information towards the beginning of a passage or paragraph that is developed more fully later in the text. For example,

- (a) RAP.1: The major aim of this trial - to show whether injection of paternal cells produced a result significantly better than the control placebo effect over the known placebo effect in this condition - has been achieved.
(RAP.1: *The Lancet*, April 27, 1985)

In the example above, the expression, "the major aim of this trial" acts as a preface or preparatory statement to the opening sentence of the paragraph, thus functioning to indicate that the discussion in that segment of text would be concerned with "Highlighting main research outcome".

- (b) RAP.2: The characteristics of cases and controls are shown in Table 1. Cases were older at diagnosis than controls were at selection, partly reflecting the increased sampling of controls in the 20-24 year age-group".
(RAP.2: *The Lancet*, 2 Nov, 1985)

The text above consists of the first and second sentences in the "Results" section of RAP.2. The first sentence acts as a preparatory statement indicating that the discussion in the unit of text to which it prefaces is concerned with issues related to the characterization of aspects of the data as the sentence following it shows.

- (c) JRV.2: The researchers also found no change in breast cancer risk in women who used high progestogen "potency" OCs before the age of 25 and went on with them for more than six years. The earlier study had also made this suggestion, which led our committee on safety of medicines two years ago to issue a list of pills and their potency, with advice to doctors to prescribe only

low progestogen OCs.

Commenting on these contradictory findings, which must confuse "scientists, family planners and general practitioners", The Lancet suggests that they may be reconcilable on the assumption that the earlier study in part reflects the fact that young women in California may have used the pill sooner than elsewhere in the United States.

(JRV.2: *New Scientist*, 7 Nov 1985)

In the example above, the discussion in the first paragraph is concerned with presenting the results of the research investigation. In the second paragraph, however, there is a shift in discourse from the presentation of results to a discussion of the results. An indication of this shift is provided by the explicit preparatory statement or expression: "Commenting on these contradictory findings ..." Reference to the suggestion made in the 'The Lancet' is but one of the views expressed in the discussion of the research results.

4.3.1.2 Explicit Lexemes

These are words that explicitly signal what the content of each move might be about.

For example:

- (a) RAP.6: The methods used to collect data on patients with cervical and prostate cancer were identical with those reported in our retrospective study of colonic and rectal tumours.
(RAP.6: *BMJ*, 30 Aug 1986)

The example above, taken from the first sentence of the "Patients and Methods" section of RAP.6 contains some explicit lexical clues which suggest that the unit of text following it will be concerned with presenting methods of data collection. Words such as "*Methods*", "*Collect*", and "*data*" supports this view. Furthermore, the sentence following it also confirms this view as the words "*retrieved*", "*Collecting*" in the text below suggests.

In brief, the names and identifying numbers of hospital patients seen between 1970 and 1982 with these diagnosis were retrieved from a computerized database and the medical record reviewed, collecting a variety of demographic, clinical and treatment data.

- (b) RAP.4: Furthermore, in frozen sections of fresh ectocervical epithelium, the basal and parabasal cells showed unequivocal reactivity with two independent antibodies (HB5 and anti-B2) specific for different epitopes of the C3d/EBV receptor molecule on B cells. These observations on ectocervix strongly suggested that HB5/anti-B2 staining of epithelial cells in tissue sections was detecting the expression of a functional EBV receptor.

The present report describes the extension of these studies to pharyngeal epithelia which may be directly exposed to the virus during its natural transmission by the oral route.

(RAP.4: *The Lancet*, 1 Feb 1986)

In the example above, two segments of discourse are identifiable. The first segment presents work that has been previously done on the subject-matter of investigation. It is signalled by the past tense verb form "showed", and by the expression "These observations ... strongly suggested that ..."

The second segment is concerned with presenting the "new" research. It is signalled by the explicit lexical items "present" and "report" modified by the definite article "the".

- (e) JRV.4: Although many more people are affected by npc than by Burkitt's lymphoma, until now no one has been able to find out how the virus gets into squamous epithelial cells.

New research by Dr Lawrence Young and his colleagues at Birmingham University's Department of Cancer Studies has shown how the virus enters these cells (*The Lancet*, Vol i, 1986, p240).

(JRV.4: *New Scientist*, 6 March 1986)

The text above contains two segments of information. The first segment presents a situation needing remediation. It begins with the adversative "Although" and is further signalled by the negative expression, "Until now no one has been able to find out ...". The second segment seems to highlight present research. It is explicitly signalled by the lexical item "New" in the expression, "New research by Dr Lawrence Young and his colleagues ... has shown ..."

4.3.1.3 Discourse Conjunctions

These are cohesive devices which function to relate linguistic items which occur in succession but are not related by other structural means to each other in

discourse (Halliday and Hasan, 1976: 223) . They tend to signal to the reader about relations in text.

For example:

RAP.7: Infections due to B.15 R meningococci are becoming increasingly common in the UK, but this is the largest outbreak due to this organism yet described in the country.

Several features of the outbreak are of importance. First, there is no evidence, after 41/2 years, that the outbreak is waning. In Norway, over 11-years period from 1974 to 1984 the average annual ...

Secondly, the age-distribution of cases is unusual by comparison with recent national data (Figure 2). Meningococcal disease usually affects children under the age of 5 years, a reflection of the low antibody levels in this age group.

Thirdly, throat swabbing revealed a very low carriage rate of B.15R meningococci in symptomless contacts.. Similar low carriage rates of this organism were found ...

There are two possible explanations for the low carriage rate. B.15 R and B:15:P1.16R strains may be of high virulence ...

(RAP.7: *The Lancet*, 6 September 1986)

The text above is taken from the "Discussion" section of RAP.7. It would seem to be "presenting the main research observations". An indication that the discourse in this segment of text constitutes a 'move' with a uniform orientation is provided by the discourse conjuncts, "First", "Secondly" and "Thirdly" which function to provide a sequential relation between ideas in all three paragraphs which they introduce.

- (b) RAP.10: Although the possibility exists that the HPV-16 DNA sequences detected in the paired internal control group could represent contamination of samples, care was taken to select tissue in as aseptic a manner as possible, and we have no evidence suggesting that cells that are mechanically disrupted during resection release amounts of DNA that can be detected in Southern analysis. In addition, the presence of HPV-16 genomes does not itself indicate that the tissue is histologically abnormal. Furthermore, the HPV-16 genome cannot itself convert a normal cell to a malignant one. The finding of HPV-16 genome, however, suggests that this control tissue may be predisposed to malignant conversion in a manner similar to that of disease recurrence in control tissue adjacent to anogenital warts in which HPV-16 genomes can be detected after laser excision.

(RAP.10: *The NEJM*, 23 Oct, 1986)

As in the previous example, indication that the discourse in the text above

presents a uniform orientation is provided first by the conjunction "Although" in the first sentence which signals a shift in discourse and the commencement of a new line of argumentation, then by the additive conjuncts "In addition" and "Furthermore" which indicate an anaphoric relationship between the sentences they introduce and the first sentence, and finally by the adversative conjunct. "However" which suggests an end to the argumentation. In accordance with this, the segment of text above can be said to represent a "constituent element" of a 'move' which may be termed "indicating positive observations". Within this paradigm, the "constituent element" would seem to be functioning to "account for the observations made".

4.3.1.4 Grammatical Specifiers of Content Relations

Grammatical specifiers of content relations are usually clefts, pseudo-clefts, interrogatives (rhetorical questions) and there constructions. They are sentence types which signal thematic prominence in text (Jones and Jones, 1985: 2).

For example:

- (a) JRV.1: All the children born in the trial were normal and only two were premature. Mowbray and his colleagues conclude that although aspects of the technique could be improved, immunisation with their patients lymphocytes is the best treatment for women selected. But how does it work? Immunologically, pregnancy is a paradox. On the one hand, the fetus might be considered an allograft tissue placed in intimate contact with one individual, the mother, which expresses genes derived from another individual, the father.

(JRV.1: *New Scientist*, 9 May 1985)

Two segments of information are represented in the example above. The first segment, made up of S_1 and S_2 represent an attempt to present the results of the research investigation. The second segment, made up of S_3 , S_4 , S_5 suggest an attempt to explain principles underlying the results. In conceptual terms, the results and explanation which follows are distinct units of information which have different functions. An indication of the distinctness of both units of information is provided by the demarcation created by the interrogative sentence (S_2), "But how does it work?"

which signals the termination of the previous line of argumentation and the commencement of a new unit of information.

- (b) ABS.12: The relation between the fatty-acid composition of adipose tissue and platelet membranes and the estimated risk of coronary heart disease (CHD) was examined in a case-control study of new angina pectoris (AP) and first acute myocardial infarction (AMI). There were progressive inverse relations between adipose linoleic acid and platelet-membrane eicosapentaenoic acid and the estimated relative risk of AP. These relations were statistically independent of each other and traditional CHD risk factors.
(ABS.12: *The Lancet*, 24 January 1987)

The text above represent two segments of information. In the first segment, an attempt is made to present the research investigation and the method adopted in the study. The second segment begins with the "there construction". It functions to present the main results of the research investigation. In this regard, the "there construction" is a grammatical signal which specifies a change in the information content of the discourse following after it.

4.3.1.5 Summary Statements

These are statements or expressions which signal the end of a 'Move' or the beginning of a concluding 'move' in a text by paraphrasing the same words or information already presented at the beginning of a 'move' or by means of an explicit concluding phrase.

For example:

- (a) ABS.13: We conclude that elevated amounts of tissue kallikrein and kinin are present in the bronchoalveolar spaces of asthmatic subjects. Kinin generation may contribute to the asthmatic response directly through edema formation and smooth muscle contraction and by augmenting release and/or production of performed (histamine) and secondary mediators such as leukotrienes and platelet-activating factor.
(ABS.13: *J of Clin Invest*, Jan 1987)

The example above represents the concluding section of ABS.13. It is a unit of information with a clearly defined function in which the phrase "We conclude that ..." signals the commencement of a concluding "move".

- (b) RAP.13: In summary, this study provides the first demonstration of kininogenase activity and elevated amounts of kinin in BAL fluid of asthmatic subjects. The predominant kininogenase possessed the structural, functional, and antigenic characteristics of a tissue kallikrein.
(RAP.13: *J of Clin Invest*, Jan 1987)

The text above is the last paragraph in the RAP text from which it is extracted. As in the example preceding it, there is an explicit indication that the information it contains is a summary of discourse in the text. The phrase "In summary" signals the function of discourse in that segment of text which can be said to be the presentation of research conclusions.

The approaches to the identification of discourse Moves discussed in this section are those in which identification is achieved by means of explicit linguistic clues. In a few cases, it may be possible that these linguistic clues are not present in a unit of text, yet that unit displays the uniform orientation characteristic of a move. In such situations, other less explicit linguistic features are examined. One of such linguistic features is the predominant tense form which the unit of text manifests. This may also contribute to providing the unit of text with a uniform orientation and enable the reader to determine the general orientation of information in the unit of text. Thus, in the example below, the predominant use of the past tense (passive) suggests that emphasis is on the action or activity being carried out rather than on the agent or the person who carried out the action. From the predominant tense used, it can be seen that the segment of text is describing the experimental procedure adopted in the research investigation.

RAP.9: Cross sections of the intracranial optic nerves were cut in thicknesses of 0.5mm and immersion-fixed (in 2 percent paraformaldehyde and 2 percent glutaraldehyde) for one week. The tissue was rinsed and immersed in 0.5 percent osmium tetroxide in the cold overnight, dehydrated through a graded series of ethanol solutions, and embedded in polybed. Sections that were 1um thick were stained in 1 percent p-phenylenediamine in methanol with use of the method of Sadun and Schaechter; this technique stains normal myelin and the lipid byproducts of degenerating nerve fibres.
(RAP.9: *The NEJM*, 21 Aug, 1986)

The ABS and the JRV which are relatively shorter in length and discourse content than the RAP provided far more difficulties in the identification of Moves, even

with all these linguistic features discussed in this section. A strategy adopted in the difficult circumstances where any of the approaches seemed not to provide a clue to the identification of Moves in the ABS and JRV is to match information in them with their equivalents in the RAP and to label them according to the function of the RAP move.

For example:

RAP.3: 4 women in the placebo group had PIH (1 mild, 3 severe) 7 pre-eclampsia, and 1 became eclamptic; in the aspirin group 2 women had mild PIH but pre-eclampsia did not occur.

ABS.3: In the placebo group PIH: pre-eclampsia and eclampsia developed in 4, 7 and 1 cases, respectively, whereas only 2 women in the aspirin group had mild PIH.

JRV.3: Only two of the women receiving aspirin had any symptoms of toxaemia - and then only PIH.

By matching information across the three genres in the example above, it can be seen that all three units of text are performing the same discourse function - that of presenting the research results.

4.4 DESCRIPTION OF MOVES

4.4.1 Outline of Canonical Moves

A detailed analysis of the schematic organization of information in the three genres was carried out following the procedures described in 4.4 above. The structure of information in the RAP was first analysed, and from it a set of canonical schematic structures and their constituent elements was identified. Following from this, the ABS and the JRV were then analysed for these canonical moves and constituent elements. Additions and omissions of schematic structural features identified in the RAP from the two derived genres were also noted. A discussion of these later observations would be presented in subsequent sections. In this section, however, an attempt is made to present an outline of the set of canonical schematic structures and constituent elements in the RAP which would form the basis for the characterization of differences between the RAP and the other two genres. The outline is also accompanied by exemplifications

from RAP texts.

The analysis of schematic structures reveals that there are as many as eleven discourse functions, representing eleven Moves in a typical RAP text, and that some of these eleven moves may be found in typical ABS and JRV texts. Below are tables indicating the number of moves identified in each text of each genre.

Table 2: The RAP: The Distribution of Moves per Text

Text	Move 1	Move 2	Move 3	Move 4	Move 5	Move 6	Move 7	Move 8	Move 9	Move 10	Move 11
1	-	+	+	+	+	+	+	-	+	+	+
2	-	+	+	+	+	+	+	-	+	+	-
3	+	+	+	+	+	+	+	+	+	+	+
4	+	+	+	+	+	-	+	-	+	+	+
5	+	+	+	+	+	+	+	-	+	+	+
6	-	+	+	+	+	+	+	-	+	+	+
7	+	+	+	+	+	-	+	+	+	+	+
8	+	+	+	+	+	+	+	-	+	+	+
9	+	+	+	+	+	-	+	+	+	+	+
10	-	+	+	+	+	-	+	-	+	+	+
11	-	+	+	+	+	-	+	-	+	+	+
12	-	+	+	+	+	+	+	+	+	+	+
13	-	+	+	+	+	+	+	-	+	+	+
14	-	+	+	+	+	-	+	+	+	+	+
15	+	+	+	+	+	+	+	+	+	+	+

Table 3: The ABS: The Distribution of Moves per Text

Text	Move 1	Move 2	Move 3	Move 4	Move 5	Move 6	Move 7	Move 8	Move 9	Move 10	Move 11
1	-	-	-	-	-	-	+	-	-	-	-
2	-	-	+	-	+	-	+	-	+	-	-
3	-	-	+	+	-	-	+	-	-	+	-
4	-	+	-	-	-	-	+	-	-	+	-
5	+	-	-	-	-	-	-	-	+	+	-
6	+	-	-	-	-	-	+	-	+	+	+
7	-	-	-	-	-	-	+	-	-	-	+
8	-	+	+	-	-	-	-	-	-	+	+
9	+	-	-	-	-	-	+	-	-	-	+
10	-	-	-	+	+	-	+	-	-	-	+
11	-	-	-	+	-	-	+	-	-	+	+
12	-	-	+	-	-	-	+	-	-	-	-
13	-	-	-	-	-	-	+	-	-	-	+
14	-	-	+	-	+	-	+	-	-	+	+
15	-	-	-	+	-	+	-	-	-	-	+

Table 4: The JRV: The Distribution of Moves per Text

Text	Move 1	Move 2	Move 3	Move 4	Move 5	Move 6	Move 7	Move 8	Move 9	Move 10	Move 11
1	+	+	+	-	+	-	+	-	-	-	+
2	-	-	+	-	+	-	+	-	+	-	+
3	+	+	+	+	-	-	+	-	+	+	-
4	+	+	+	-	-	-	-	-	-	+	+
5	+	-	+	-	+	-	-	-	-	+	+
6	+	-	+	-	-	-	+	-	-	+	+
7	-	-	-	-	-	-	+	-	-	+	+
8	-	-	+	-	-	-	+	-	+	+	-
9	+	-	+	-	+	-	+	-	-	+	-
10	-	-	-	-	-	-	-	-	+	+	-
11	+	+	+	+	-	-	-	-	+	+	-
12	+	-	-	-	+	-	+	-	+	-	+
13	+	-	-	-	+	-	+	-	-	+	+
14	+	+	-	-	+	-	+	-	+	+	-
15	+	-	-	-	-	-	+	-	+	+	+

The tables above provide information which suggests that all three genres tend to have a reasonable proportion of identical Moves. For this reason, we shall attempt for now to outline only the Moves in the RAP which are representative of the Moves in the other two genres and after that, in the next section, refer to the moves in the ABS and JRV in our discussion of differences in schematic structures.

In a typical RAP text, the author is likely to:

- Begin by providing the reader with some background details about the subject-matter of the research investigation.
- Conduct a brief review of previous studies related to the research he is reporting.
- Introduce the new research and state his objectives.
- Identify to the reader the source of his data and the method he adopted in collecting them.

- Discuss the methods adopted in the analysis of data.
- Discuss statistical procedures employed in the analysis of data.
- Discuss results which are consistent with the objectives of the study.
- Present results which deviate from expected outcomes.
- Discuss the main results and their significance.
- Explain and justify results and approaches adopted in achieving them.
- Provide a conclusion to the report and indicate implications.

These, in the main, represent the eleven discourse functions or moves identified in a typical RAP text. In discourse functional terms, they may be categorized as follows:

Table 5: Canonical RAP Moves

Move 1:	Presenting Background Information
Move 2:	Reviewing Related Research
Move 3:	Presenting New Research
Move 4:	Describing Data Collection Procedure
Move 5:	Describing Experimental Procedure
Move 6:	Describing Data-Analysis Procedure
Move 7:	Indicating Consistent Observations
Move 8:	Indicating Non-consistent Observations
Move 9:	Highlighting Overall Research Outcome
Move 10:	Explaining Specific Research Outcomes
Move 11:	Stating Research Conclusions

MOVE ONE is often, though not necessarily, an initiation move. To this end, it tends to provide information which seems to serve as background explanation to the topic of discourse and which tends to highlight the main research problem. Thus, the Move tends to contain two main "constituent elements".

1 Reference to established knowledge in the field.

2 Reference to main research problem.

The two main constituent elements are mutually inclusive. That is, both of them may occur as elements of Move 1 in the same text. However, the most frequently occurring element of the two is, "Reference to established knowledge".

The information contained in Move 1, often tends to be persuasive (Swales 1981a), but sometimes it tends to be anecdotal and didactic, presenting chronological accounts of events by means of temporal and locative adverbials as sentence elements. Furthermore, the information in Move 1 may indicate the commencement of discourse in text by presenting knowledge which in the context of the research can be regarded as holding true for a long period of time in the field.

Thus, Move 1 tends to be characterized by two main linguistic features:

- (a) The use of preparatory expressions or statements which have temporal adverbials or locative adverbials as sentence elements, eg.
 - (i) In England and Wales, death rates during the past hundred years have been consistently higher in the north and west of the country than in the south east (RAP.5).
 - (ii) Since 1940, when there were over 12000 notifications in England and Wales, Meningococcal disease has been much less common (RAP.7).
 - (iii) The Epstein-Barr virus (EB) is one of the five herpes viruses that attack humans. In Europe and the US infection can produce glandular fever (mononucleosis) (JRV. 4).

The temporal and locative adverbials in the examples above can be said to be scene-setting linguistic features whose function in Move 1 is to signal the commencement of information which would act as a background for understanding what the discourse in the text is going to be about.

- (b) The use of simple present tense forms, eg.
 - (i) Ischaemic heart disease - the narrowing and blocking of the arteries that supply blood to the muscles of the heart - is the biggest single cause of death in the western world in the 20th Century. (JRV.5).
 - (ii) Congestive heart failure is a common condition reported to affect 1 percent of

the population, with an annual incidence of approximately 3 per 1000. (RAP.15).

- (iii) Alzheimer's disease is a dementing disorder of unknown cause in which there is degeneration of neuronal subpopulations in the central nervous system.

All three sentences in the examples above make use of the simple present tense form, the verb to be. This is so because the items of information they contain are generalizations; the statements they make hold true for all times and in scientific terms, they can be described as theoretical statements. Therefore, within the context of discourse in the texts where they occur, they state what is already well known in the field; hence they are identified in this analysis as providing background information.

An example of how Move 1 is realized in text is provided below:

Table 6: Exemplification of Move 1

1	(1) The principal histopathological features of brains of patients with Alzheimer's disease include neuronal loss, neurofibrillary tangles, neuritic plaques and granulovacuolar degeneration. (2) These changes occur most frequently in the hippocampus and other limbic structures and in neurocortical association areas. (3) Primary neocortical areas such as motor and visual cortices are relatively spared. (4) The clinical diagnosis of Alzheimer's disease cannot always be made with certainty, and the neuropathological features are qualitatively not diagnostic of Alzheimer's disease, because similar histologic changes may occur in elderly patients who do not have dementia. (5) Memory deficits, aphasia, and apraxias are the most common signs of Alzheimer's disease, although defects in visual perception have also been described. (6) Difficulties in conducting ophthalmologic examinations of patients with dementia have largely precluded careful examination of their retinal and optic-nerve functions...
2	

(RAP.9: *The NEJM*, 21 Aug 1986)

Two moves can be identified in the sample text above - Moves 1 and 2. In addition to a change in information content from making reference to the problem in Move 1 to indicating a gap - "Difficulties in ... have largely precluded careful examination of ... " in Move 2, there is also a shift from the use of the simple present tense form in 1 to the past tense form in 2.

Move 1 in the sample text above contains the two constituent elements associated with the presentation of background information. The first three sentences

of the move make "reference to established knowledge in the field". The presentation of cognitive content in the move tends towards the didactic presentation of information, with the listing of attributes of patients with Alzheimer's disease in S1 and S5. There is also consistent use of present tense verb forms indicating that the propositions made in the sentences hold true and can be generalized.

MOVE 2 - Reviewing Related Research is next in the sequential order of discourse presentation in the texts. It contributes to the development of discourse in the texts in which it occurs by providing information against which the research being reported can be judged. That is, it attempts to place the research within the context of ongoing research in the field. Consequently, Move 2 tends to present two types of information, referred to as "constituent elements" of the Move. They are:

- 1 Reference to previous research.
- 2 Reference to limitations of previous research.

Both constituent elements tend to occur with equal regularity in RAP texts. The first constituent element - reference to previous research - is realized in the following ways:

- (a) Identification of the author by placing his name in subject position, eg.
 - (i) Thus, Pike et al reported that women who used OCs with high progestagen "potency" for more than 4 years before age 25 were over four times more likely to have breast cancer before age 37 than women who did not use these contraceptives before age 25. (RAP.2).
- (b) Identification of researcher by the use of common nouns, eg.
 - (ii) These investigators demonstrated that nasal installation of allergen to atopic individuals elevated the concentration of Kinins in Lavage fluid, correlating with symptoms, P-tosyl-L-arginine methyl ester (TAME) esterase activity, and histamine release. (RAP.13).
- (c) Generalized reference to previous research, eg.
 - (iii) A growing body of data shows that blood transfusions have immunological effects on patients and experimental animals beyond those of alloimmunisation

to blood cell antigens (RAP.6).

- (d) Reference to the product of research investigations, eg.
- (iv) The findings supporting this conclusion come from more than two score prospective epidemiologic studies in many countries throughout the world, from clinical investigations such as angiographic studies, and from postmortem investigations. (RAP.11).

The second constituent element - reference to limitations of previous research - tends to be realized in the following ways:

- (a) By offering a negative evaluation of previous research, eg.
- (v) Some studies have failed to find such associations or have found small differences that are not significant. Furthermore, almost all studies have been retrospective reviews of medical records, subject to many biases, raising the possibility that ... (RAP.6).
- (b) by indicating a gap in the previous researches (Swales 1981), eg:
- (vi) However, the possible prevention of PIH and pre-eclampsia in primigravidae by suppression of platelet TXA₂ production with low-dose aspirin has not been investigated (RAP.3).

As the examples above show, Move 2 tends to make use of five main linguistic forms:

- (a) Use of simple past tense verb forms to refer to single research events, eg. Thus, Pike et al reported that ... (example (i))
- (b) Use of present perfect tense form to refer to more than one previous research event whose results would seem to have some generality, eg. Some studies have failed to find ... (example (v)).
- (c) Use of present tense forms to refer to more than one previous research event whose results have implications for the research to be reported, eg. A growing body of data shows that ... (example (iii)). The findings supporting this conclusion come from ... (example (iv)).
- (d) Use of adversative adverbial conjuncts, eg. However, the possible prevention of PIH and ... (example (iv)).
- (e) Use of negative verb form, eg. However, the possible prevention of ... has

not been investigated (example (vii)). Some studies have failed to find ... (example (v)).

The text below illustrates the realisation of Move 2 in context.

Table 7: Exemplification of Move 2

(1) Difficulties in conducting ophthalmologic examinations, of patients with dementia have largely precluded careful evaluation of their retinal functions. (2) Cogan reported poor visual activity, alexia, agnosia, and spatial disorientation, which he attributed to abnormalities in the parietal cortex, in three patients with Alzheimer's disease. (3) Schlotterer et al similarly concluded that Alzheimer's disease affects central visual processing to a greater extent than it affects primary ocular functions. (4) Nissen et al observed elevated visual contrast-sensitivity thresholds in patients with Alzheimer's disease, suggestive of an optic-nerve abnormality.

(RAP.9: *The NEJM*, 21 August, 1986)

In the text above, S_1 makes reference to limitations of previous research. The rest of the text makes reference to the previous researches alluded to in the previous sentence. The underlined words and phrases indicate elements which tend to signal discourse in Move 2.

After providing a background to the study and reviewing related research, the author states the purpose of the new research in MOVE 3. In addition to stating research purposes Move 3 also tends to contain information which is related to the primary method of investigation adopted in the study. In some cases, reference may also be made to the sample data on which the study is based. However, the dominant "constituent element" in Move 3 is "Reference to Research Purpose". This dominant element is expressed in any of the following five ways or combinations thereof.

- (a) By means of nouns referring to the report/study in subject position and present tense form.
 - (i) The present report describes the extension of those studies to pharyngeal epithelia which may be directly exposed to ... (RAP.4).
 - (ii) In this study we describe a degeneration of retinal ganglion cells and of optic nerves in ... (RAP.9).

- (b) By means of the present perfect tense form + the adverbial "NOW".
- (iii) We have now investigated the relation between OC use and risk of breast cancer in women under 45 by analysing data from ... (RAP.2).
- (c) By the use of the third person plural personal pronoun we in subject position.
- (iv) Having achieved good results with our preliminary investigation, we decided to carry out a controlled trial of treatment with paternal cells (RAP.1).
- (d) By means of adverbial phrase as sentence elements:
- (v) In a randomised placebo-controlled double-blind trial we have now investigated the effect of low-dose aspirin, taken from ... (RAP.3).
- (e) By the use of explicit lexemes.
- (vi) The objective of the current investigation was to determine in a population-based study, whether levels of apolipoprotein B and ... (RAP.8).
- (vii) We wanted to find out whether the success was truly immunological or a placebo effect ... (RAP.1).

The text below is an example of

Table 8: Exemplification of Move 3

(1) Having achieved good results with our preliminary attempts at immunisation, we decided to carry out a controlled trial of treatment with paternal cells. (2) We wanted to find out whether the success was truly immunological or a placebo effect of a positive treatment, such as has been found with other treatments for recurrent abortion. (3) We also wanted to study the possibility that immunisation with paternal cells might produce subfertility or other unwanted side-effects. (4) The trial, started in February 1982, lasted 3 years. (5) Participants were selected from women attending a recurrent-abortion clinic.

(RAP.1: *The Lancet*, 22 April 1985)

The text above is overwhelmingly concerned with stating the purpose of the research investigation (S₁-S₃). The last two sentences, however, are concerned with general issues of research methodology. As usual, the underlined expressions indicate signals for identifying Move 3.

Thus far, an attempt has been made to establish the existence of three Moves in a typical RAP text. The three moves so far described represent the macro functional

units into which information may be structured in the "Introduction" section of an RAP text. It is possible to have all three moves in one text. But, occasionally, texts do have only two out of the three moves in their introduction section. The regularity with which the moves occurred in the texts analysed in this study is shown in Table 2 above. Based on their frequency of occurrence, Moves 2 and 3 with 100% occurrence are termed "Normally Required Moves". On the other hand, Move 1 with 40% occurrence is termed an "optional" move, where optionality is defined as a move that can occur but is not obliged to occur (Hasan, 1985: 62).

Having described the moves that make up the Introduction section, we shall now proceed to characterize the Moves in the "Methods" section of a typical RAP text.

MOVE 4 - Describing Data-Collection Procedure - represents the first in the order of presentation of information in the "Methods" section of an RAP text. As its functional label clearly suggests, the move is concerned with a discussion of all aspects of the process of data identification, selection and delimitation. In accordance with this, information contained in Move 4 may be summed up into three "constituent elements":

- (a) Indicating source of data, eg:
 - (i) HB5 (IgG2a) and anti-B2 (1gM) are mouse MAbs against the C3d/EBV receptor on human cells and were kindly provided in ascitic fluid form by Dr T Tedder (RAP.4).
 - (ii) In brief, the names and identifying numbers of hospital patients seen between 1970 and 1982 with these diagnoses were retrieved from a computerized data base (RAP.6).
- (b) Indicating data size, eg:
 - (i) The study population includes all children and young adults (up to age 26) living in Bogalusa, Louisiana (approximate population in 1980, 17,000). (RAP.8).
 - (ii) The final sample sizes for analyses involving a paternal or maternal myocardial infarction were 2416 and 2604 respectively (RAP.8).
- (c) Indicating criteria for data-selection,
 - (i) Criteria for inclusion in the trial were as follows:
 - 1 No detectable antibody against paternal lymphocytes.

2 No cause found for the abortions etc (RAP.1).

- (ii) Women were eligible if (1) they were 20-40 years of age at initial diagnosis (histologically confirmed) of breast cancer; (2) diagnosis was first made between 1 Dec 1980 and 31 Dec 1982 (RAP.2).

In accordance with the functions of the constituent elements, information in Move 4 tends to be identified with the following characteristic linguistic features:

- (a) Use of verb forms to indicate source of data:
- (i) The couples in the trial were taken from those referred to us for investigation of recurrent abortion (RAP.1).
- (ii) Samples of fossa of Rosenmuller, soft palate, oropharynx, cheek, tongue, nasal cavity, and trachea were obtained post mortem.
- (b) By means of conditional clauses to indicate criteria for data selection:
- (i) Patients were included in the study if they had undergone surgery at our institution, had a documented follow up of at least six months, had no evidence ... (RAP.6).
- (c) The use of past tense or present tense verb forms to indicate data size:
- (i) The study group consisted of 30 patients in district hospitals in Western Scotland (RAP.10).
- (ii) The study population includes all children and young adults (up to age 26) in Bogalusa ... (RAP.8).
- (d) By means of explicit lexemes:
- (i) The final sample size for analysis involving a paternal or maternal myocardial infarction were ... (RAP.8).
- (ii) The study population includes all children ... (RAP.8).
- (iii) Criteria for inclusion in the trial were ... (RAP.1).

Below is an example of Move 4, taken from RAP.12.

Table 9: Exemplification of Move 4

Move

- 4 The sampling frame was men aged 35-54 years listed on the Lothian Health Board Central Register with an address in the city of Edinburgh. A systematic sample of 6000 men (35-54) was drawn from the register and names were sent to general practitioners to check against practice records. Medical exclusions for serious disability or life-threatening illnesses were made by the general practitioners at this stage. If the general practitioner had no personal record of a patient registered against his name on the Central register, the man was still retained in the sample.

(RAP.12: *The Lancet*, 24 Jan, 1987)

MOVE 5 - Describing Experimental Procedure - tends to occur in those texts that involve laboratory tests or experimental work. It may also occur, albeit sparingly, in texts which are non-experimental, but are concerned with the analysis and description of already generated data, especially where some procedural or methodological processes are considered crucial to the overall observations made during the process of analysis. Move 5 presents in a logical and sequential manner the steps and procedures adopted during experimentation. It also tends to indicate the types of apparatus used in the research. In line with these, Move 5 tends to contain information which refers to the following "constituent elements":

- (a) Identification of main research apparatus, by:
- (i) Explicit lexemes (apparatus as grammatical subject in a passive construction) where apparatus is subject of the clause, eg.

A computer-generated list of pseudo-random numbers was used to allot the patient to receive either her own or her husband's lymphocytes. (RAP.1).

HB5 and Anti-B2 MABS were used down to ascitic fluid dilution of 1:2000 and 1:100, respectively, in phosphate-buffered saline (PB5) containing 10% normal goat serum; (RAP.4).

- (ii) Explicit lexemes (Apparatus as adjunct in a passive construction), eg.

Blood pressure was measured in the left upper arm with a standard sphygmomanometer (lower end of the mercury column at the level of the heart). (RAP.3).

- (b) Recounting experimental process, by:
- (i) The use of passive verb forms, eg. Strains of meningococci isolated were grouped and tested for sulphamide sensitivity at Gloucester PHL. (RAP.7).
- Blood pressure was measured every 5 mins until a stable diastolic pressure had been recorded for 15 mins (RAP.3).
- (ii) The use of temporal adverbials, eg. After 1h incubation at 37°C, the slides were washed in PBS and incubated with a 1:75 dilution of fluorescein-isothiocyanate-conjugated goat anti-mouse immunoglobulin (Sigma) for a further 1h at 37°C before final washing (RAP.4).

Cross-sections of the intracranial optic nerves were cut in thicknesses of 0.5mm and immersion-fixed (in 2 percent paraformaldehyde and 2 percent glutaraldehyde) for one week. The tissue was rinsed and immersed in 0.5 percent osmium tetroxide in the cold overnight, dehydrated through a graded series of ... (RAP.9).

- (c) Indicating criteria for success, by:

- (i) The use of explicit lexemes, eg.

More than 20 degenerated axonal profiles per 100 high power fields (63x objective) were considered to represent slight degeneration (+), more than 100 degenerated profiles per 100 high power fields indicated moderate degeneration (++) , and more than 500 degenerated profiles per 100 high-power fields indicated severe (+++) degeneration. (RAP.9)

A failure of treatment was an abortion before 28 weeks and a success was scored when the pregnancy had continued beyond 28 weeks. (RAP.1).

An indication of the realization of Move 5 in text is provided by the example text below:

Table 10: Exemplification of Move 5

The name, address, and date of birth of each man were entered on a computer file and the names were randomised into blocks of 500. A postal questionnaire was sent in batches of 500 every month for a year (April 1983 to April 1984). The questionnaire included the self-administered World Health Organisation chest pain questionnaire. If there was no reply to the first questionnaire within 2 weeks, a reminder letter and questionnaire were sent. The response rate to the postal questionnaire was 72%. Replies were divided into three groups. Group 1 consisted of men with no reported history of CHD who were negative to the WHO chest pain questionnaire; Group 2, men with no reported history of CHD who were positive to the WHO chest pain questionnaire; and Group 3 men with a reported history of CHD, either MI or AP. Men in Group 3 were not studied further.

(RAP.12: *The Lancet*, 24 January 1987).

MOVE 6 - Describing Data-Analysis Procedure - tends to occur mainly in texts which report on researches which adopt statistical, quantificational approaches to the analysis of data. Therefore they differ from Move 5, which is concerned with describing the practical processes of experimentation. However, both Moves do tend to occur simultaneously in the same text. Move 6 represents an attempt by the writer to identify the statistical tools used in the study and recount how they were applied to the data analysed. Also reported in Move 6 are any modifications to a statistical tool and definitions of measuring instruments as used in the context of the study. Thus, Move 6 tends to present information which indicates that it is made up of the following "constituent elements".

(a) Defining terminologies:

This tends to be indicated by means of explicit lexemes, eg:

- (i) The term "ever-use" of OCs is taken to mean use of any duration, for any reason, at any time in a woman's life. (RAP.2).
- (ii) Definitions of pregnancy outcome were decided before the start of the trial. PIH was defined as the presence of a diastolic blood pressure 95mm Hg on at least two occasions 6 or more hours apart. (RAP.3).

(b) Indicating process of data classification:

It is signalled by the use of Explicit lexemes, eg:

- (i) We have divided the causes of infant deaths into five groups according to Woolf's classification: congenital causes (Register General, 1921, Short List Nos 27, 28); bronchitis and pneumonia (18-19); infectious diseases (2-9, 13); diarrhoea (1, 22) and other. (RAP.5).
- (ii) Birth weights were classified according to the centiles of distribution of normal weight.

(c) Identifying analytical instrument/procedure. This is indicated by:

Use of passive verb forms, and by placing an analytical instrument either in initial or end positions in the sentences, eg.

- (i) Logistic regression was used to predict the probability of myocardial infarction in the parents on the basis of serum levels of the variables in their offspring (RAP.8).
- (ii) The variables of subscapular skinfold thickness, smoking behaviour, alcohol

consumption and the use of oral contraceptives were then controlled for by analysis of covariance. (RAP.8).

(d) Indicating modifications to instrument/procedure. This may be signalled by:

(i) Explicit lexeme, eg.

Initial adjustments were made with use of Z scores that were specific for race, sex and two-year age group (all Z-score calculations were based on groups of at least 37 children). (RAP.8).

Below is an example of Move 6 as realized in RAP.6.

Table 11: Exemplification of Move 6

The methods detailed in our previous paper on colonic and rectal cancer were used. We used X^2 statistics to compare the transfusion and control groups with respect to the various prognostic factors. We also used such statistics to compare the incidence of recurrence or death in the various groups defined by the variables under study. These latter analyses were followed by statistical techniques that allow adjustment for differences in prognostic factors. The prognostic factors studied included age, data of surgery, duration of surgery, packed cell volume on admission, kind and amount of blood components received, type of cancer, and stage of cancer. To analyse this last factor for patients with different types of cancer we devised an improvised classification for statistical analysis. "Early stage" included patients with cervical cancers of stage O and IA, colorectal cancers of stage A, and prostrate cancers of stage A₁, and A₂. "Intermediate stage" included patients with cervical cancer stages of 1B, 11A and 11B, colorectal and prostrate cancer stages B₁ and B₂. "Advanced stage" included patients with cervical cancer stages 111A and 111B, colorectal cancer stages C₁ and C₂, and prostrate cancer stage C. We appreciated that this process could not yield precise data on the effect of stage on recurrence of cancer, but wished to incorporate broadly representative information on stage in our study of prognostic factors that might be associated with transfusion and cancer recurrence. A multivariate stepwise logistic regression analysis was performed to determine which of the above prognostic factors were associated with whole blood transfusion.

When fitting proportional hazards models ...

(RAP.6: *British Medical Journal*, 30 August 1986)

MOVE 7 - Indicating Consistent Observations - is the initial Move in the "Results" section of an RAP text. It constitutes the major unit of information in the Results section. The move contains aspects of information which are concerned with stating the overall observations made during the process of analysis, reporting on all other significant observations which impinge on the objectives of the research and presenting information on visuals, such as tables, graphs and pictorials. It also

attempts to give an account of necessary procedural adjustments that were made before consistency was achieved in the observations made. In line with these, Move 7 may be summed up as consisting of the following "Constituent elements" or units of information:

(a) Highlighting overall observation, by:

The use of explicit preparatory expressions, eg:

- (i) At the beginning of the study the two groups were comparable with respect to age, body weight, diastolic blood pressure, and effective pressure (Table 1) (RAP.3).
- (ii) We initially thought that whole blood transfusions might be associated with greater risk of recurrence and death due to cancer than red blood cell transfusions. Table 1 shows data for incidence of recurrence and mortality for all of the cancer patients who had had transfusions (RAP.6).

By means of quantifying expressions, eg:

- (i) Both HB5 and anti-B2 gave unequivocal membrane-associated staining of cells in stratified squamous epithelia. (RAP.4).
- (ii) 65 cases of meningococcal infection were identified in the 4 1/2 years to March 31, 1986. Only half the cases (33/65) were formally notified; the remainder were ascertained through PHL records.

(b) Indicating specific observations, by

Reference to visuals, eg:

- (i) The characteristics of cases and controls are shown in Table 1. Cases were older than ...
- (ii) Relative risk of breast cancer, for every-user of OCs compared with never-users, by age at diagnosis is shown in Table 11. For the 5-year age intervals, from 25-44, there was no ...
- (iii) Table III shows our findings on the issue raised by Pike et al; the estimates of relative risk, by their approach and by the method of this study show no significant ...

By means of existential constructions, eg:

- (i) There was no difference in success rate between the 9 pregnancies in this group and those in women without a previous successful pregnancy (Table 11). (RAP.1).
- (ii) There were more premature deliveries and small-for-gestational-age babies in the placebo group, but these differences were not statistically significant.

By means of adverbial discourse conjuncts as preview or preparatory expressions, eg:

- (i) In the placebo group mean (± 50) value of Malondialdehyde production by platelets at 33-35 weeks gestation was 4.6 ± 1.1 nmol/10 platelets vs 0.51 ± 0.3 in the aspirin group. (RAP.3).
- (ii) In patients not being treated with vasodilators at the time of random assignment, the reduction in crude mortality was from 60 to 37 percent ($P < 0.02$). (RAP.15).

By the use of comparatives, eg:

- (i) In the aspirin group 5 women smoked 10-20 cigarettes/day by comparison with 4 women in the placebo group.
- (ii) The patients receiving three or fewer transfusions of red blood cells survived for longer periods than those receiving three or fewer transfusions of whole blood ($P = 0.047$ by Mantel-Cox).

By the use of past tense verb forms to report results, eg:

- (i) Of the 106 evaluable patients treated with LAK cells plus interleukin-2, 8 had total regression of all cancer (complete regression), 15 had more than a 50 percent tumour reduction (partial response), and ... (RAP.14).

By the use of passive verb forms to refer to tables, eg:

- (i) The correlations between infant mortality and the SMR's from the leading causes of death during 1968-78 are shown in Table 1. (RAP.5).
- (ii) See example i.2 above.
- (c) Accounting for the observations made. This tends to be signalled by:

The use of discourse conjuncts as sentence initial elements, eg:

- (i) However, because the levels of LDL cholesterol also tended to be elevated in these children, the ratio of LDL cholesterol to apolipoprotein B was not decreased among the children whose mothers reported myocardial infarction. (RAP.8).
- (ii) Because of the large number of children whose fathers reported a myocardial infarction, most of the subsequent analyses were restricted to paternal myocardial infarctions. (RAP.8).

Use of hedging verbs, eg:

- (i) This impression appears to be misleading, however, because of certain differences between the cases and controls of this age group that were related to their use of OCs. (RAP.2).
- (ii) Although none of the serum lipid levels or lipoprotein - cholesterol fractions were related to paternal disease, children whose fathers had had a myocardial infarction tended to have elevated levels of apolipoprotein B (an increase of ...) (RAP.8).

Move 7 is illustrated with the example below taken from RAP.3.

Table 12: Exemplification of Move 7

At the beginning of the study the two groups were comparable with respect to age, body weight, diastolic blood pressure, and effective pressor dose (Table 1). In the aspirin group 5 women smoked 10-20 cigarettes/day, by comparison with 4 women in the placebo group. Mean (\pm SD) dose of aspirin/day/kg body weight at 28 weeks' gestation was 0.8 ± 0.2 mg.

2 patients were excluded from the analysis after the Code had been broken. 1 woman in the aspirin group stopped taking her tablets after a week; she remained normotensive and delivered a healthy female infant of 2690g at 40 weeks. Another woman in the aspirin group had a high malondialdehyde value of 10.6 nmol/10 platelets at 34 weeks; she later disclosed that she had taken the tablets at first very irregularly and had stopped taking them altogether at 32 weeks gestation. She was admitted with severe PIH at 39 weeks and after induction of labour a healthy female infant of 3710g was delivered.

(RAP.3: *The Lancet*, 4 January 1986)

MOVE 8 - Indicating Non-consistent Observations - is concerned with the description of negative results. That is, incidences in which results which did not conform with the intended outcomes for applying a particular procedure were obtained in the study. Its occurrence in the RAP texts studied is low, thus, making it an "optional move". Its low level of occurrence (40%) in the corpus would seem to suggest that there is a tendency for researchers not to report negative aspects of their work. But more importantly, it lends credence to the fact that research articles do not reflect all that goes on in the research laboratory, but only those facts which the researcher considers important for his purpose (Knorr-Certina, 1981).

Another feature of Move 8 is that it does not always come immediately after Move 7. In some texts, RAP.3 and RAP.12, non-consistent observations are reported alongside their equivalent consistent ones. In RAP.11, on the other hand, it can be clearly identified as a segment of text at the end of portion of the 'Results' section immediately following the end of Move 7. Despite these few irregular instances, the identification of "non consistent observations" as a move in its own right is, no doubt, important considering the changes in information structure which a transition from a "positive" to a "negative" mode of reporting can introduce to overall meaning in the results section of an RAP text.

The information contained in Move 8 is normally concerned with just presenting non-consistent observations. Therefore, it does not contain additional "constituent elements".

Move 8 tends to be signalled by:

The use of negative verb phrase, eg.

- (i) However, quantification was not possible in this tissue because of the small sample size and the known considerable variations in numbers of ganglion cells from site to site.
- (ii) Thioflavine S staining (which is specific for amyloid) on sections of retina and optic nerve did not reveal neuritic plaques, neurofibrillary tangles, or amyloid angiopathy (RAP.9).

Use of negative qualifiers and noun phrases, eg.

- (i) Many of the residential retinal ganglion cells showed degenerative changes, including cell shrinkage or cell swelling with vacuolization (RAP.9).
- (ii) 1 genetically abnormal abortion was found, a trisomy 15 in a woman immunised with her husband's cells.

Apart from realisations in the forms indicated above, Move 8 may also be realised by the explicit reporting of negative results, as illustrated in text RAP.14 below:

Table 13: Exemplification of Move 8

Four patients died as a result of complications of treatment - one who received LAK cells plus interleukin-2 and three who received interleukin-2 alone. Two died of myocardial infarction, and two died of pulmonary insufficiency complicated by sepsis (sepsis due to infection in a central nervous cathreter in one patient and pulmonary insufficiency due to bronchial obstruction in tumour in another). One patient with extensive lung cancer died after a long hospitalization: his death was attributed to progressive neoplasia.

(RAP.14: *The NEJM*, 9 April 1987)

MOVE 9 - Highlighting Research Outcome - usually occurs as the first segment of information in the "Discussion" section of an RAP text, often corresponding with information continued in the first paragraph in that text. Its

function is to announce the attainment or otherwise of the main research objective through a restatement or confirmation of the main research result. To this end, it is usually brief, being made up in most cases of one complex sentence. Therefore, it tends to be made up of just the constituent information whose function is to highlight research outcome.

The Move is signalled by:

- (a) Explicit preparatory statements, eg:
 - (i) The result of this study suggests that OC use has no significant effect on the risk of breast cancer in women under 45 years of age (RAP.2).
 - (ii) Our results offer clear evidence of a distinctive histopathologic process in the retinas and optic nerves of patients with Alzheimers disease ... (RAP.8).
 - (iii) The present results show, as did our previous, smaller study, that HSU DNA sequences are not frequently retained in genital tumours, if southern hybridization is used as the criterion for measurement (RAP.10).
- (b) The use of explicit lexemes, eg:
 - (i) The major aim of this trial - to show whether injection of paternal cells produced a result significantly better than the control placebo effect over the known placebo effect in this condition - has been attained.

Below is a sample of Move 9 as realised in RAP.14.

Table 14: Exemplification of Move 9

This report presents our experience with the administration of 127 courses of LAK cells plus interleukin-2 in 108 patients and 53 courses of high-dose interleukin-2 alone in 49 patients, all of whom had metastatic cancer for which no effective systematic therapy was available.

(RAP.14: *The NEJM*, 9 April 1987)

MOVE 10 - Explaining Specific Research Outcome - constitutes the major Move of the "Discussion" section of an RAP text. It almost invariably follows after the research outcome has been highlighted in Move 9. Move 10 presents information which functions to further restate the main observations made in the study, indicate their significance, interpret and justify them by reference to research procedures adopted in the study and by contrasting them with similar observations made in related

studies. To this end, units of information which make up five "constituent elements" were identified in the move. They are:

(a) Stating a specific outcome, by:

Past tense verb forms:

- (i) Thirdly, throat swabbing revealed a very low carriage rate of B.15R meningococci in symptomless contacts (RAP.7).
- (ii) The current study which was free of all temporal uncertainty and selection biases indicated that adverse levels of apolipoproteins are evident long before clinical symptoms begin (RAP.8).

Other linguistic signals which are used to state main observations include:

- (1) Use of existential constructions (see examples in Move 7).
 - (2) Use of adverbial discourse conjuncts as preview or preparatory expressions (see Move 7).
 - (3) Use of comparatives (see Move 7).
- (b) Interpreting the outcome. This is signalled by:

Use of explicit lexemes (present tense verbs forms + that particle).

- (i) The association of ischaemic heart disease with neonatal mortality suggests that the childhood influences predisposing to it are related to nutrition during prenatal and early postnatal life (RAP.5).
- (ii) Thus, we can only speculate that the low dose of aspirin used in our trial prevented PIH and pre-eclampsia by suppressing thromboxane synthesis while leaving prostacyclic formation intact ... (RAP.3).
- (iii) This finding implies that certain HSV DNA sequences may have a function in the initiation of cellular events that are important for the transformed state.

(c) Indicating significance of the outcome. It is signalled by:

The use of explicit lexemes (nominal groups - adverbial modification + noun), eg.

- (i) These results are particularly important because they come in the wake of several observations suggesting that EBV infection of epithelia may be central to the events of primary injection and ... (RAP.4).
- (ii) The loss of retinal ganglion cells is of particular interest in that it is not accompanied by any filamentous accumulation in the retina ... (RAP.9).
- (iii) Thus, they lend powerful support - along with the data from the other quintile analyses and the decile analysis and data from other studies - to the conclusion that the relationship of serum cholesterol to CHD in the US is a graded one and not a threshold one (RAP.11).

(d) Contrasting present and previous outcomes. This is signalled by:

Explicit preparatory expressions/explicit lexemes, eg:

- (i) Previous morphometric studies of normal human optic nerves have shown variations in the number of retinal ganglion-cell axons in different persons (1.0 to 1.6 million) and even between the eyes of the same person. All the optic nerves in those studies had a compact arrangement of axons, in marked contrast to the optic nerves of the patients with Alzheimer's disease, which had sparse packing of axons and considerable glial replacement (RAP.9).
- (ii) Studies cited by Ferenzey et al and our results indicate the importance of close and frequent follow up of patients from whom malignant genital tissue has been excised (RAP.10).

By reference to author as subject, eg:

- (i) Ferenezy et al have discussed the difficulty of treating vulvar disease, with its extensive lesions and frequency of recurrence. Our study included 11 vulvar carcinomas, 9 of which (82 percent) contained HPV-16 genomes (RAP.10).

(e) Indicating limitations of outcomes:

Two types were identified in the data:

- 1 Indicating limitations of new research
- 2 Indicating limitations of previous research.

The first type is signalled by:

Explicit preparatory expressions, eg:

- (i) The major sources of error in studies of this type are bias and selection of cases compared with controls ... (RAP.2).
- (ii) We do not know the effect of immunisation on women with detectable anti-paternal antibody because these were excluded from the trial (RAP.1).
- (iii) The data do not support the assertion that this relation is entirely due to whole blood transfusions being ... (RAP.6).

Use of lexemes suggesting research constraints, eg:

- (i) It is difficult to quantify this in our group, who have had 3 to 8 previous spontaneous abortions (RAP.1).
- (ii) Our data was clearly unable to define conclusively the role of HPV-16 in genital neoplastic disease (RAP.10).

The second type, ie. "indicating limitations of previous research is signalled by:

Use of adverbial expressions + negative verb forms, eg.

- (i) Although one perspective study has evaluated the relative strengths of HDL cholesterol and apolipoprotein A-1 levels in predicting cardiovascular disease, the subjects were not fasting at the time of Venipuncture, and only 12 ... (RAP.8).
- (ii) Though some studies have reported a significant association between parental smoking and increased incidence of hospitalization of children for respiratory disease, they do not distinguish between the effect of passive exposure to cigarette smoke after birth and the effect ... (RAP.8).
- (iii) However, McCance et al did not mention the status of HPV-16 genomes in paired, matched internal control tissue.

MOVE 11 - Stating Research Conclusion - is the last move in an RAP text. It represents a writer's attempt to sum up his views on the possible contributions which the study makes to the field. To this end, it tends to contain information on the implications of the study. It also states the need for further studies. Thus, Move 11 is made up of two "constituent elements":

- (a) Indicating research implications.

This is signalled by the use of explicit lexemes, eg:

- (i) The practical implications of our data are that transfusions to patients with cancer should be performed only when clinically essential and that red cells should be used in preference to whole blood (RAP.6).
- (ii) The results of the current study support the concept that levels of apolipoprotein A-1 and apolipoprotein B may ... (RAP.8).

- (b) Prompting further research.

This is signalled by, explicit lexemes, eg:

- (i) The nature of such an effect deserves further study. It may be that ... (RAP.1).
- (ii) Further studies are required to support or refute this hypothesis (RAP.3).

In general terms, Move 11 is also signalled by words such as conclusion, summary, eg:

- (i) In summary, this study has shown that enalapril to patients with severe congestive heart failure is associated with a considerable reduction in mortality (RAP.15).
- (ii) In summary, this study provides the first demonstration of kininogenase activity and elevated amounts of ... (RAP.13).

The text below which is taken from RAP.1 shows Moves 9, 10 and 11

Table 15: Exemplification of Moves 10 and 11

9 The major aim of this trial - to show whether injection of paternal cells produced a result significantly better than the control placebo effect over the known placebo effect in this condition- has been attained.

10 Immunisation produced a much higher success rate (78%) and we suggest that it is the treatment of choice for the group of women we selected. We do not know the effect of immunisation on women with detectable anti-paternal antibody, because these were excluded from the trial. The success rate with our method of immunisation in terms of pregnancy outcome is high, but we cannot be certain always to produce detectable antibody with this regimen. These imperfect correlations may mean that abortion in some women is due to other causes, that the immunisation is sub-optimum in some, or that cytotoxic antibody is only an indirect index of effective immunisation. A better regimen of immunisation might be 2 or 3 separate doses of cells from smaller amounts of blood, and we intend to study this possibility.

The present study has shown a pregnancy success rate of 37% in the group treated with their own cells. Other studies have shown a strong placebo effect, but there is also a spontaneous success rate even after 3 abortions. It is difficult to quantify this in our group, who have had 3 to 8 previous spontaneous abortions.

We assume that any placebo effect is psychological rather than a result of receiving their own well-washed lymphocytes. The nature of such an effect deserves further study; it may be that a psychological effect is mediated hormonally and susceptible of study.

RAP.1: *The Lancet*, 27 April 1985

To recapitulate, I have presented an analysis of the schematic organization of information in a typical RAP text. The analysis yielded 11 functional units of schemata referred to in this study as 'Moves'; as well as 26 sub-moves or "constituent elements". The range of linguistic features which characterize these moves and constituent elements have also been identified. An illustration of how each move is realised in the content of an RAP text is also provided.

In conclusion, I present below an outline of the Moves and Constituent Elements as identified in the analysis.

Table 16: Outline of Moves and their Constituent Elements

<u>MOVE 1:</u>	<u>Presenting Background Information</u>
by	<ol style="list-style-type: none"> 1 Reference to established knowledge in the field 2 Reference to main research problem.
<u>MOVE 2:</u>	<u>Reviewing Related Research</u>
by	<ol style="list-style-type: none"> 1 Reference to previous research 2 Reference to limitations of previous research
<u>MOVE 3:</u>	<u>Presenting New Research</u>
by	<ol style="list-style-type: none"> 1 Reference to research purpose 2 Reference to main research procedure
<u>MOVE 4:</u>	<u>Describing Data-Collection Procedure</u>
by	<ol style="list-style-type: none"> 1 Indicating source of data 2 Indicating data size 3 Indicating criteria for data-collection
<u>MOVE 5:</u>	<u>Describing Experimental Procedure</u>
by	<ol style="list-style-type: none"> 1 Identification of main research apparatus 2 Recounting experimental process 3 Indicating criteria for success
<u>MOVE 6:</u>	<u>Describing Data-Analysis Procedure</u>
by	<ol style="list-style-type: none"> 1 Defining terminologies 2 Indicating process of data classification 3 Identifying analytical instrument/procedure 4 Indicating modifications to instrument/procedure
<u>MOVE 7:</u>	<u>Indicating Consistent Observations</u>
by	<ol style="list-style-type: none"> 1 Highlighting overall observation 2 Indicating specific observations 3 Accounting for observations made
<u>MOVE 8:</u>	<u>Indicating Non-Consistent Observations</u>
<u>MOVE 9:</u>	<u>Highlighting Overall Research Outcome</u>
<u>MOVE 10:</u>	<u>Explaining Specific Research Outcomes</u>
by	<ol style="list-style-type: none"> 1 Stating a specific outcome 2 Interpreting the outcome 3 Indicating significance of the outcome 4 Contrasting present and previous research outcomes 5 Indicating limitations of outcomes

Table 16 continued ...

MOVE 11:	<u>Stating Research Conclusions</u>	
by	1	Indicating research implications
	2	Prompting further research

4.5 DIFFERENCES IN THE SCHEMATIC ORGANIZATION OF INFORMATION

In the previous section, a set of canonical 'Moves' which are likely to occur albeit in different forms, in any of the three genres in this study (RAP, ABS, JRV) were outlined. The outline was based on an analysis of RAP texts, considered to be most representative of the moves embodied by all three genres.

In this section, an attempt is made to discuss the differences which exist between all three genres in the way information is schematized and organized. The discussion focusses on four main areas of difference:

- (1) Differences in the number of Moves realised
- (2) Differences in the constituent elements of Moves
- (3) Differences in the linguistic features of Moves
- (4) Differences in the order of presentation of Moves.

4.5.1 Differences in the Number of Moves

This is perhaps, the most obvious difference in the realisation of schematic structures that can be characterized in the genres. The reason for this contention lies in the fact that the notion of parallel texts (Hartman, 1980) as used in this study does not carry with it equality of text length and number of organizing elements between texts across the three genres. Therefore, the analysis shows that the RAP which tends to have longer texts and embody more cognitive content has more Moves than either the ABS or the JRV which are relatively shorter texts. Thus, as has been pointed out in the previous section, a typical RAP text consists of eleven Moves out of which some are "Normally Required Moves" and others that may be "optional" moves. The "normally

required moves" consist of 9 moves which have almost 100% occurrence in the RAP texts analysed (see Table 2). The Moves are:

- Move 2: Reviewing Related Research
- Move 3: Presenting New Research
- Move 4: Describing Data-Collection Procedure
- Move 5: Describing Experimental Procedure
- Move 6: Describing Data-Analysis Procedure
- Move 7: Indicating Consistent Observations
- Move 9: Highlighting Overall Research Outcome
- Move 10: Explaining Specific Research Outcomes
- Move 11: Stating Research Conclusions

The "Optional Moves", on the other hand, consist of two Moves with 40% occurrence in the texts which make up the RAP data. They are:

- Move 1: Presenting Background Information
- Move 8: Indicating Non-Consistent Observations

The analysis of ABS texts reveals that not more than nine 'moves' from the RAP may be identified in an ABS text. In reality, no text in the data had as many as nine moves. The maximum number of moves found in any one text is 5 (see Table 3). Out of the nine possible Moves, there appears to be only two "normally required Moves" and seven "optional" moves. The "normally required Moves" are:

- Move 7: Indicating Consistent Observations
- Move 11: Stating Research Conclusions

The optional Moves are:

- Move 1: Presenting Background Information
- Move 2: Reviewing Related Research
- Move 3: Presenting New Research
- Move 4: Describing Data-Collection Procedure
- Move 5: Describing Experimental Procedure
- Move 9: Highlighting Overall Research Outcome
- Move 10: Explaining Specific Research Outcomes.

Two Moves from the RAP were not identified in the ABS. They are:

- Move 6: Describing Data-Analysis Procedure
- Move 8: Indicating Non-Consistent Observations.

In terms of numbers of realised Moves, a total of nine moves were also identified in the JRV. But, no text in the data embodied all nine moves. The maximum number of Moves identified in any one text is seven (see Table 4). Out of the nine possible Moves which may be found in a JRV text, seven are regarded as "Normally required Moves", while two are identified as "Optional" Moves. The "Normally Required Moves" are:

- Move 1: Presenting Background Information
- Move 3: Presenting New Research
- Move 5: Describing Experimental Procedure
- Move 7: Indicating Consistent Observations
- Move 9: Highlighting Overall Research Outcome
- Move 10: Explaining Specific Research Outcome
- Move 11: Stating Research Conclusion.

The optional Moves are:

Move 2: Reviewing Related Research

Move 4: Describing Data-Collection Procedure

Moves 6 and 8 were not identified in the JRV.

The results obtained from the analysis so far are interesting, not so much because they present an index of the Moves which occur in each genre, but because they are capable of throwing some light on two issues which are of interest to this study, viz:

- 1 Determining what gets transferred at the global level when a research report designed for professionals/specialists is re-written for a less specialist reader.

- 2 Determining how what gets transferred is influenced by pragmatic factors of context and situation.

It would appear that the mapping of Moves from the RAP to either the ABS or the JRV is constrained by three information reduction strategies, namely: Omission, Combination and Construction.

Omission is the process whereby writers of ABS and JRV texts tend to leave out, not just propositions, but whole chunks of information or Moves from the RAP in the information transfer process. It relates to Van Dijk's (1980) macrostructure rule of "Deletion", except that deletion is concerned with the elimination of propositions rather than segments of text.

As we have seen, the ABS and the JRV each omit two out of the eleven moves said to occur in the RAP in their texts during the process information transfer. This is one cause of difference in discourse organization at the level of schematic structure between the RAP and these other genres. The extent of the difference in

discourse structure between the RAP and the other two genres suggests that omission is a major strategy adopted in the process of information transfer from the RAP to the JRV and the ABS.

Having established that Omission is a fundamental strategy of text reduction employed in the ABS and JRV, and therefore, responsible for the variation in the number of Moves identified in each genre, a pertinent question that needs to be asked is "What is omitted in the ABS and JRV and why is it omitted?" Simply put, both the ABS and the JRV omit Move 6 - Describing Data-Analysis Procedure, and Move 8 - Indicating Non-consistent Observations. But, in order to answer the latter part of the question, the Moves would need to be examined in context to determine what makes up their "constituent elements". The regularity of occurrence of the Moves in the RAP data would also have to be examined in order to establish their relative importance.

The analysis shows that Move 6 is made up of four constituent elements (see Section 4.5), viz:

- 1 Defining terminologies
- 2 Indicating process of data classification
- 3 Identifying analytical instrument/procedure
- 4 Indicating modifications to instrument/procedure.

An examination of the propositions which make up these constituent elements reveals that they tend to embody a lot of statistical terms, formulae and symbols, eg:

- (a) Definitions of pregnancy were decided before the start of the trial. PIH was defined as the presence of a diastolic blood pressure ≥ 95 mm Hg on at least two occasions, 6 or more hours apart.
- (b) Differences in relative frequency between placebo and treatment groups were analysed with the F test. Wilcoxon rank-sum test was used to analyse differences in measured variables between the two groups. A P value of ≤ 0.05 was reported as significant.

As in Van Dijk (1980: 46) macrostructure rule of Deletion, Omission involves

the deletion of Moves of the text base (RAP) which are too specific or localized and, therefore, irrelevant for the interpretation of global meaning or "which do not denote facts which may be subsumed as normal properties of a more global fact which is denoted by a macrostructure of the discourse". Thus, a possible explanation for the omission of Move 6 may lie in the fact that the information contained in the move is too specific, being primarily concerned with figures, formulae, symbols and a host of statistical details which can be conveniently left out without a loss in meaning to the global discourse. It might also be due to the difficulty which figures, symbols and formulae can cause in the information reduction process, because they cannot be easily paraphrased or generalized without a loss in meaning.

Move 8, which is the second move not realised in both the ABS and the JRV, is concerned with indicating non-consistent observations. It does not contain additional constituent elements and it has been described elsewhere in this study as a "peripheral" move due to its low occurrence in the RAP. As has been pointed out, its low occurrence would seem to confirm the fact that researchers tend not to report all that they observe in the laboratory in their reports. One aspect of research work which is hardly mentioned in reports is non-consistent or negative reports. Since non-consistent observations are hardly mentioned in research articles, it is not surprising, therefore, that they get omitted during the process of information transfer into a more condensed form. Related to this, might also be the fact that the ABS and the JRV are written for a slightly different audience and for a different purpose. The audience for which both genres are written need not necessarily be specialists in the field of medicine. The ABS audience might as we have seen, include information specialists, professional abstractors, science popularizers and students most of whom are interested in the positive results of the study for different reasons. The JRV audience, on the other hand, is made up mainly of students and scientists who are not necessarily researchers. It may also include laymen with interest in the results of the research being popularized. The purpose of popularizing scientific information for this class of audience is to acquaint them with latest research findings in the field. The

assumption is that the class of readers for which the JRV is written are likely to be more interested in the positive results than in the negative ones. This fact is supported by the observation that most JRV texts tend to begin with Move 9 which is concerned with "Highlighting Overall Research Outcome", which as the example below from JRV.2 shows is invariably positive.

- eg. The risk of getting breast cancer before age 45 does not alter significantly for women who started using oral contraceptives (OCs) before they were 20 and went on doing so for more than four years. This is the main conclusion of an American study - the largest of its kind to date.
(JRV.2: *New Scientist*, 7 Nov 1985)

The use of Omission as a strategy employed by mature writers for mapping text sentences to summary texts has been observed by Brown, Day and Jones (1983) and Sherrard (1986). The strategy is constrained by the need to tailor the discourse in the summary texts, in this case the ABS and the JRV, to the assumed "importance evaluation" of their audience. The need to tailor importance evaluation to readers' goals has been identified as a fundamental constraint in text summarizing (Fum, Guida and Tasso, 1986: 256). Similarly Kintsch and Van Dijk (1975, 1978), Van Dijk (1980) have also identified simple Omission as a common reduction strategy in text summarizing.

The second information reduction strategy which constrains variation in the number of moves between the RAP on the one hand and the ABS and JRV on the other is **combination**. This refers to the tendency for some moves in both the ABS and the JRV to combine information from two or more Moves in the RAP into one single Move. For example,

In a paired sequential double-blind trial of immunological treatment of recurrent spontaneous abortion successful outcome of the next pregnancy was significantly more common in women injected with purified lymphocytes prepared from their husband's blood than in those injected with their own lymphocytes.

(ABS.1: *The Lancet*, 27 April 1985).

The example above is taken from ABS.1. In it, there is a conflation of what

looks like information from two Moves into one Move. On the surface, it is not immediately clear which of the information from the two moves should take precedence over the other in terms of assigning a functional label to the "new" move. What is clear, though, is that the first part of the sentence ... "In a paired sequential double-blind trial ... recurrent spontaneous abortion" is an adverbial adjunct and that the latter part is the main clause. In functional terms, it is equally clear that the adverbial adjunct is functioning as the Theme (Danes, 1974; Halliday, 1979) of the sentence while the latter part is functioning as the Rheme. It has been observed (Firbas, 1974) that the theme carries a much lower communicative dynamism (CD) than the rheme in an utterance. That is, the rheme communicates more important information than the theme in an utterance. Following from this observation, therefore, the information contained in the latter part of the sentence in the example above is given precedence over information expressed by means of the adverbial adjunct in the assigning of a functional label to the "new" Move. Thus, information from two moves: MOVE 5: Describing Experimental Procedure, and MOVE 7: Indicating Consistent Observations - was assigned the functional label of MOVE 7 in the analysis of the sample text above.

The process of combining two or more moves to form a "new" one has some relationship with what Van Dijk (1980) refers to as the "Construction" rule in his theory of macrostructure. The rule involves the taking together of propositions and substituting them, as a joint sequence, by a proposition that denotes a global fact (Van Dijk, op cit: 48). Sherrard (1986: 331) has also observed that combination is a strategy employed during the process of text reduction.

The discussion so far would seem to suggest that there is nothing that can be said regarding possible differences between the ABS and the JRV in the number of Moves they both tend to select from the RAP, since according to the analysis both genres select nine Moves from the RAP.

However, as the analysis of the ABS shows, out of the nine possible Moves

which may be identified in an analysis of ABS texts, seven occur with a degree of regularity which suggests that they might be "optional" in the context of discourse in the ABS. And, out of the seven optional Moves, five, ie. Moves 1, 2, 4, 5 and 9 - which were found to occur in less than one-third of the texts in the data would appear not to be central to discourse in the ABS. On the other hand, three Moves - 7, 10, 11 - can be said with a fair amount of certainty to belong squarely to ABS discourse. In terms of number, this observation is clearly related to Urquhart's (1986) identification of four primary functions to which the discourse in Abstracts may be reduced (see Chapter 2, section 2.5). However, while Urquhart identifies "procedures" as one of the main functions, this study identifies "explaining research outcomes" as one of the functions which occur with a fairly high degree of frequency. The difference in observations may be due to differences in the labelling of functional categories. But more fundamentally, it may be due to the nature of Abstracts examined in both studies, for while this study is based on a close analysis of Abstracts by reference to their accompanying articles, Urquhart's study was based on Abstracts from Abstracting Journals without reference to their source articles. It is not unlikely that in the latter study, discourse functions were assigned to segments of text on the basis of linguistic signals alone following intuitions regarding how scientific articles are structured. Such a procedure is not likely to be very capable of distinguishing information which tends to explain results, usually from the Discussion section of the article, from that which only indicates results. In such an analysis, both instances are likely to be labelled as indicating research results. Furthermore, such an analysis is not also likely to distinguish between the function "Describing Data Collection Procedure" and "Describing Experimental Procedure". In an analysis which, as it were, takes the Abstract out of its context both functions are likely to be conflated into one, which in the case of Urquhart (op cit) is "Procedures". Such a conflation may be responsible for why Urquhart identifies "Procedures" as one of the four main functions in an Abstract.

In contrast to the ABS, only two of the nine Moves which may be identified in

a JRV text tend to occur with a frequency which suggests that they are "Optional" Moves. Those Moves are Move 2 and 4 which occurred in very few texts in the data. The others, eight in number, occur with a high degree of regularity, and can, on that basis be said to be normally required Moves in the JRV. Out of the two optional Moves, one (Move 4), whose occurrence is less than 15% in the data would appear to be an incidental Move in the genre.

Following from the discussion so far, it would appear that while the ABS and the JRV tend to select an equal number of Moves from the RAP, they vary in terms of the extent to which all of the Moves selected can be said to be normally required in the genres in which they occur. In this regard, it would appear that more RAP moves are normally required to occur to constitute a normal JRV text than are required for an ABS text. However, both genres would seem to be consistent in identifying Move 7 - Indicating Positive Results; Move 10 - Explaining Research Outcomes and Move 11 - Starting Research Conclusions as three Moves which are fundamental to discourse in their texts. But, while the ABS tends to favour Move 11 as the next most important Move (60%) after Move 7 (93%) which must be selected from the RAP to constitute discourse in texts in the genre, the JRV appears to favour Move 10 (80%) instead. This observation might be due to the assumptions which writers of either genre have regarding the information which best serves the interest of their respective audience.

4.5.2 Differences in the Constituent Elements of Moves

The most striking observation in the analysis of "Constituent Elements" of Moves is the variation in the number of those elements which texts in each genre are capable of realising. Table 16 shows the constituent elements which make up each of the Moves, in the RAP. As the analysis shows, twenty-eight constituent elements were identified in the RAP. In two of the Moves though (Moves 8 and 9), no distinction is made between information expressing the Move and that expressing a constituent element because their cognitive content is short and not capable of any further sub-divisions.

In contrast to the RAP, fourteen constituent elements were identified in the ABS. The constituent elements identified in the ABS are:

Table 17: Constituent Elements in the ABS

Move 1:	- Reference to Established Knowledge
Move 2:	- Reference to Previous Research
Move 3:	- Reference to Research Purpose
Move 4:	- Indicating Source of Data - Indicating Data Size
Move 5:	- Recounting Experimental Procedure
Move 7:	- Highlighting Overall Observation - Highlighting Special Observation
Move 9:	- Highlighting Overall Research Outcome
Move 10:	- Stating a Specific Outcome - Interpreting the Outcome - Indicating Significance of the Outcome
Move 11:	- Indicating Research Implications - Prompting Further Research

The analysis of the JRV, on the other hand, reveals that texts in the genre are capable of realising nineteen constituent elements. Some of the constituent elements occur in more than one move. The constituent elements identified in the JRV are:

Table 18: Constituent Elements in the JRV

Move 1:	- Reference to Established Knowledge - Reference to Main Research Problem - Stressing the Local Angle - Explaining Principles and Concepts
Move 2:	- Reference to Previous Research - Reference to Limitations of Previous Research
Move 3:	- Reference to Research Purpose
Move 4:	- Indicating Source of Data - Indicating Data Size
Move 5:	- Recounting Experimental Process
Move 7:	- Highlighting Overall Result - Indicating Specific Observations
Move 9:	- Highlighting Overall Research Outcome
Move 10:	- Stating a Specific Outcome - Explaining Principles and Concepts - Indicating Comments and Views - Indicating Significance of the Outcome - Contrasting Present and Previous Research Outcomes
Move 11:	- Indicating Research Implications - Prompting Further Research - Stressing the Local Angle

The results outlined above point to two obvious facts concerning the differences which exist between texts in the three genres with regards to the constituent elements they may contain. First, they point to the fact that the RAP contains more constituent elements than either the ABS or the JRV, and that the ABS has the least number of constituent elements of all three genres. This is not altogether an enlightening result as the difference may be due to differences in text length which exists between texts of each genre. But, second and more importantly, they serve to indicate the characteristic elements of text in each genre. This would appear to be a more revealing result, and one that is capable of throwing some light on the elements that constitute the differences within and across genres, and on factors responsible for

the differences.

A comparison of the constituent elements in the RAP vis-a-vis the ABS and the JRV shows that the differences between them is due to two main factors: (1) Deletion, and (2) Addition. In terms of the ABS, there is a tendency for the following RAP constituent elements not to feature in ABS texts:

Table 19: RAP Constituent Elements Deleted from the ABS

Move 1:	- Reference to Main Research Problem
Move 2:	- Reference to Limitations of Previous Research
Move 3:	- Reference to Main Research Procedure
Move 4:	- Indicating Criteria for Data-Collection
Move 5:	- Identification of Main Research Apparatus - Indicating Criteria for Success
Move 7:	- Accounting for the Observations Made
Move 10:	- Contrasting Present and Previous Research Outcomes - Indicating Limitations of Outcomes.

An examination of Constituent Elements in the JRV shows that the following elements from the RAP tend to be deleted when information is transferred to texts in the JRV genre:

Table 20: RAP Constituent Elements Deleted from the JRV

Move 3:	- Reference to Main Research Procedure
Move 4:	- Indicating Criteria for Data Collection
Move 5:	- Identifying Main Research Apparatus - Indicating Criteria for Success
Move 7:	- Accounting for Observations Made
Move 10:	- Interpreting the Outcome - Indicating Limitations of the Outcomes

In the ABS, these deleted RAP elements may be said to represent details that do not contribute to the purpose of discourse in the genre. As has been pointed out elsewhere in this study, the ABS is a highly condensed text; its primary purpose is to extract from the main research article the most important pieces of information which are capable of giving an insight into the research reported in the main text. In view of this, it is understandable that ABS writers tend to delete from ABS texts aspects of information which relate, for instance, to the indication of criteria for data-collection or to the contrasting of present and previous research outcomes. The same thing may be said for all the other deleted elements shown in Table 19 above.

In the JRV, the deleted constituent elements can also be said to represent details which do not contribute to the purpose of discourse in the genre. In addition, they can also represent elements which do not contribute to the construction of a theme or a central message in texts in the genre. The purpose of the JRV is to make the results of Medical research investigations available to a much wider audience. The process of popularization involves the choice of a theme, based on the overall research result, around which discourse in the texts is woven. For example, reporting on the topic "The Body fights back against Cancer", in JRV.14, the writer contends that: "Training the body's immune system to fight cancer is an increasingly promising therapy that is winning adherents" and supports his contention by reference to the research result being reported upon. His aim is to present the research as a breakthrough in cancer research. Therefore, he selects from the original research report aspects of information which furthers the theme or message in the light of his overall aim. Obviously, with a purpose and a theme to develop, JRV writers cannot select from all available information contained in the RAP. Therefore, it is understandable, considering the purpose of discourse and the fact that most themes of JRV texts are concerned with the overall research outcome, that aspects of information such as those already identified in Table 20 above should be deleted during the process of information transfer from the RAP to the JRV.

The second factor which contributes to variation in the constituent elements of

which texts in all three genres may be made up is "Addition". This factor is paramount in contributing to the differences in the number and form of constituent elements between the RAP and the JRV. The analysis of JRV texts shows that in addition to constituent elements from the RAP, they tend also to contain additional constituent elements which are unique to the genre. Such additional constituent elements include:

- 1 Stressing the local angle (Moves 1 and 11)
- 2 Explaining principles and concepts (Moves 1 and 10)
- 3 Indicating comments and views (Move 10)

The first added constituent element - Stressing the Local Angle - is a strategy often adopted in journalistic writings (Burkett, 1973; Newson and Wollert, 1988). It refers to the tendency by journalists to identify with their target audience by relating the topic of discourse to local experiences with which they are familiar. To this end, it relates to the observation by Farago (1976: 11) that effective science communication (popularization) "should be seen to be part of the reader's cultural heritage". It also relates to the contention by Funkhouser et al (1970) that "the lower the audience level of a publication, the more of its material should be given to practical applications". With regards to the JRV whose audience of students and non-practising scientists may be said to be at a lower level of expertise than the audience of the RAP, this tendency takes the form of relating the results of research investigations, especially those conducted in other countries, to problems, issues and developments in the country or environment of the majority of its target audience.

For instance, JRV.2 begins by indicating to its reader that it is reporting on the results of an "American Study". This is stated in its opening paragraph, thus:

The risk of getting breast cancer before age 45 does not alter significantly for women who started using ...
That is the main conclusion of an American study ... the largest of its kind today.

It concludes by indicating to its readers how the results of the "American Study" compare with developments in Britain, thus:

In Britain, far less than five percent of breast cases would have been attributable to OC use in 1980, compared with perhaps 10 percent in 1985 and 20 percent in 1990. In fact, the Royal College of General Practitioners found such an increased risk, though it was of marginal significance.

In the passages above, a contrast is drawn between the situation in America and that in Britain from where the "New Scientist" draws the majority of its readership. The aim of such a contrast, it would seem, is to point out to the readership that the risk of breast cancer due to pill use is becoming a growing problem in Britain, as in America where, according to the research results, it is a problem already.

The second addition in the JRV to the constituent elements derived from the RAP is "Explaining Principles and Concepts". It relates to what might be regarded as the didactic role of JRVs examined in this study. That is, the tendency for JRV writers to provide readers with basic instructions on principles and concepts underlying the research problem or the research results. Therefore, the constituent element tends to occur mostly in Move 1 and Move 10, although it may not be exclusive to them. Occasionally, some form of explanation accompanies information contained in other Moves, but this is not consistent in the texts examined.

The constituent element tends to be characterized in text by features such as the following:

- 1 Use of question and answer cues:
eg. In this trial, however, there were no adverse effects in either mothers or fetuses.

How does aspirin exert its beneficial effects?
Scientists are finding increasing evidence to suggest that normal blood flow in the mother and fetus depends on the right balance between the production of TXA₂, a vasoconstrictor, and prostacyclin (PGI₂), a vasodilator. In PIH and pre-eclampsia, the balance tips heavily in favour of TXA₂.

(JRV.3: *New Scientist*, 30 Jan 1986)

Myers (1986: 27) citing Morrison (1981) contends that questions and answer is "one of the most important powerful syntactical patterns of popular science texts".

From the example above, it would seem that the function of questions in popular science writings is to signal the commencement of an explanation to an aspect of the research result. This view is supported by the contention by Myers (op cit: 28) that:

"The effect of the question and answer structure of popular texts is to imply that in research, as in the undergraduate education, the questions are always given, and that, as in undergraduate education, the answers must surely follow".

This contention also supports the observation made earlier in this discussion regarding the didactic role of the constituent element. "Explaining principles and concepts".

2 The predominant use of present tense verb form

eg. In the control group, one woman suffered mild PIH, seven had pre-eclampsia, and one suffered eclampsia. Seven women required a caesarean delivery.

Aspirin crosses the placenta and can cause bleeding in newborn infants. It can also produce haemorrhaging by the mother during normal or caesarean delivery.

(JRV.3: *New Scientist*, 30 Jan 1986)

Two constituent elements are discernable in the example above. The first is "Indicating specific observations"; the second is "Explaining principles and concepts". Note the transition from the past tense verb form to the present tense form from one constituent element to the other.

The third constituent element which is exclusive to the JRV and is not found in either the RAP or the ABS is "Indicating comments and views". This refers to the tendency of JRV texts to contain comments and views obtained from either the authors of the research article being popularized or from independent sources, such as review articles or from other researchers on the subject. It may also refer to comments and views expressed either by the magazine or newspaper doing the popularization or by the popularizer himself.

Comments and views tend to be signalled in JRV texts by the following:

1 The use of reporting verbs, eg.

- (a) Young's team is now growing squamous epithelial cells in tissue cultures and infecting them with EB virus. They describe their results as "encouraging". When such infected cultures are established then, says Young, there are some exciting possibilities.
(JRV.4: *New Scientist*, 6 March 1986)
- (b) However, they did not check people with other forms of brain disease. Dr Peter Davis, an expert on Alzheimers disease, of Albert Einstein College of Medicine in New York, said that the significance of the work will not be clear until that is done.
(JRV.9: *The Times*, 23 August 1986)
- 2 By direct quotations, eg.
- He said: "The real issue here is not distinguishing Alzheimers from normals; it's picking out Alzheimers from any one of 40 other diseases".
(JRV.9: *The Times*, 23 August 1986)
- 3 By explicit lexemes, eg.
- (a) Commenting on these contradictory findings, which must confuse "scientists, family planners and general practitioners", the LANCET suggests that they may be reconcilable, on the assumption that the earlier study in part reflects the fact that ...
(JRV.2: *New Scientist*, 7 Nov 1985)
- (b) Dr Macnab and Dr Clements agree with the view that cervical cancer is a predominantly venereal disease and a vaccine against papilloma may be desirable.
(JRV. 10: *The Times*, Oct 23, 1986)

The discussion so far has indicated that the JRV differs from the ABS and the RAP because it tends to contain additional constituent elements aside from those which it shares with the other two genres. These additional constituent elements would seem to be consistent with the purpose and medium of discourse in the JRV, for unlike the RAP which is a highly specialised genre written for a highly sophisticated science audience, the JRV is a popular genre written for an audience with a lower level of science sophistication. This contention is in line with the observation by Funkhouser et al (1970) that the lower the audience level of a publication, the more it should rely on examples for exposition. The additional constituent elements which exist only in the JRV and are not found in the RAP and ABS would seem to confirm this assertion. Furthermore, it would also seem to be a strategy used to improve readability by improving on the levels of shared knowledge which exists between the writer and his audience. This contention is related to a similar one by Myers (1986: 30) that the

breaking down of compound and complex sentences into more simple sentences is a straightforward improvement of readability. In line with Myers, the new elements, especially those which relate to "Explaining principles and concepts" and those "Indicating comments and views" may be said to function to break down the issues contained in the research results to a much simpler level, thus making it easier for the JRV audience to read. The introduction of these additional constituent elements would also seem to be in line with the observation by Funkhouser et al (op cit: 220) that "a little non-science material incorporated into a piece of science writing may increase enjoyment and the tendency to seek further information and will not necessarily hurt information gain". While the constituent elements cannot exactly be said to represent non-science information, they, nevertheless, are not characteristic of information normally found in research articles. Thus, they may be said to be "non-research article information" and as such, are likely to have the same effect on the reader as non-science information. Therefore, following from Funkhouser et al, it would seem that these additional constituent elements function to "draw the reader in" as Myers (1986) puts it.

4.5.3 Differences in the Linguistic Features of Moves

An attempt has been made to characterize the linguistic features of Moves in the RAP (see section 5.2). In this sub-section an attempt is made to contrast the significant linguistic features observed in the ABS and JRV with those already identified in the RAP. The discussion does not take the form of a blow by blow account of differences between Moves across the genres. Rather, attention is focussed on only those Moves which manifest noticeable differences across the genres or on linguistic features which appear unique to texts in a particular genre which may not necessarily be associated with any one particular Move.

An examination of Moves in the RAP, ABS and JRV reveals that there are some differences between them in the type of linguistic features they tend to use to

convey information. The most noticeable linguistic differences across the three genres lie in the number of sentences that Moves in each genre tend to embody as well as the average length of sentences of moves in the genres. The average number of sentences per move, per genre is presented below (Table 21).

Table 21: Average Sentences per Move

Genre	No of Sentences	No of Moves	Average No of Sentences per Move
RAP	568	142	4.
ABS	101	50	2.02
JRV	281	78	3.6

As the results in Table 21 shows, on an average, there tends to be a greater number of sentences per move in the RAP than in either the ABS and the JRV. The ABS though, has the least number of sentences per Move of the three genres. It must be emphasised that the figures above are only an average of the number of sentences which a Move in any of the genres may contain. In reality, some Moves in the ABS and the JRV tend to contain only one sentence. In the ABS, almost all instances of Moves 1, 2, 3, 4, 5 and 11 tend to be made up of only one sentence, while Moves 7 and 10 tend to contain more than one sentence each. In the JRV, there is a tendency for each of Moves 1, 2, 3, 4 and 9 to contain one sentence. Although the majority of Moves in the RAP tend to contain more than one sentence, in some texts in the genre, Moves 3 and 9 were found to be made up of just one sentence.

The tendency for Moves in the RAP to have more sentences than those in either the JRV or the ABS would seem to indicate that the RAP communicates more information than do the other two genres. Similarly, information in Table 21 above also suggests that the JRV communicates more information than the ABS.

Table 22 shows the average length of sentences in each Move identified in texts in each genre.

Table 22: Average Length of Sentences per Move

Genre	No of Words	No of Sentences	Average Words per Sentence
RAP	14675 568	25.84	25.84
ABS	2658	101	26.32
JRV	6715	281	23.90

As the results in the table above show, the ABS has the highest average number of words per sentence. It is followed closely by the RAP. The JRV has the lowest average sentence length. This result would seem to be an indication of the level of complexity of sentences that may be found in texts in each genre. Thus, it may be inferred from the results that sentences in the ABS may be more complex than those in the JRV and that there might be little or no difference in the complexity of sentences between the ABS and the RAP. These observations relate to the contention by Myers (1986: 29) that a major syntactical revision which takes place when information is transferred from a professional text is the breaking of compound and complex sentences into more simple sentences. It is also supported by examples from texts in the genres as may be seen below:

- (a) In a randomised placebo-controlled double-blind trial, we have now investigated the effect of low-dose aspirin, taken daily from 28 weeks' gestation until delivery, on development of PIH and pre-eclampsia in normotensive primigravidae judged to be at risk of hypertension because of an increased blood-pressure response to intravenous angiotensin II.

(RAP.3: *The Lancet*, 4 January 1986)

- (b) The possibility of preventing pregnancy-induced hypertension (PIH) and pre-eclampsia in primigravidae by suppressing production of thromboxane A₂ with low-dose aspirin was investigated in a randomised, placebo-controlled, double blind trial.

(ABS.3: *The Lancet*, January 4, 1986)

- (c) A new study by H Wallenburg and his colleagues in the Department of Obstetrics and Gynaecology at the Erasmus University Medical School, Rotterdam, has shown that low doses of aspirin given in the last three months of pregnancy may prevent toxemia. Pregnancy-induced hypertension (PIH) is a feature of the toxemia. It includes eclampsia and pre-eclampsia.

(JRV.3: *New Scientist*, 30 January 1986)

All three texts in the example above present the same type of information. They all present information aimed at introducing the new research. But while the texts from the ABS and the RAP present their information in one complex sentence in which two functions - the purpose of the research and the method of investigation are combined, the JRV text, on the other hand, presents its information in three separate simple sentences. And unlike both the ABS and the RAP, the JRV only focusses on one aspect of the two functions contained in the other two genres. That is, making reference to an attained purpose in the study. As we have pointed out in section 4.6.2 above, Myers (op cit: 30) contends that this tendency to break down complex sentences

into more simple sentences is a strategy aimed at enhancing readability.

A number of other linguistic differences also exist between the three genres which may be illustrated using the sample texts above. First, there is a transformation from the active in Texts A (RAP) and C (JRV) to the passive in text B (ABS). However, the difference between the genres is as much a difference in form as it is of the function which the various forms of the verb perform in the context of discourse in the texts. In the RAP text which apparently uses the same voice as the JRV, the perfect tense functions to make reference to an action that went over a time in the past and is completed with the moment of speech or writing (Celce-Murcia and Lassen-Freeman, 1983: 64). In the JRV text, on the other hand, it functions to describe a state resulting from an action that went over a period of time in the past that got completed before the moment of writing. On the other hand, the passive is used in the ABS to present information which tends to emphasize and highlight the new research. The use of the verb in this context is not meant to indicate time of the situation referred to. Rather it functions to emphasize an important aspect of the research which the writer wishes to stress as what the study is about. The use of passive in this way relates to the observation by Lachowicz (1981: 107) that the passive voice is used in EST to emphasize what is important at the particular stage of research described, to the deliberate exclusion of everything else. In text B (ABS) above, the important aspect of research being emphasized is "The possibility of preventing pregnancy-induced hypertension (PIH) and pre-eclampsia in primigravidae by suppressing the production of thromboxane A₂ with low-dose aspirin".

Second, a transformation of tense was also observed in the way information is expressed in Move 4 in the RAP and that in its JRV equivalents. In the RAP,

information in Move 4 tends to be expressed by means of passive verb forms. In the JRV, on the other hand, it tends to be expressed by the use of active verb forms, particularly the past tense. The examples below illustrate the contrast which exists in the function of verbs used in Move 4 in the RAP and the JRV.

RAP Move 4: Describing Data-Collection Procedure

23 The trial was designed to include 46 angiotensin-11-sensitive primigravidae, in the treatment group and 23 in the placebo group. This number was chosen to be able to demonstrate ...

To collect the required number of angiotensin-11-sensitive women, healthy primigravidae attending our antenatal clinic with an uncomplicated pregnancy of 26 weeks' duration, were given written and oral information ...
(RAP.3: *The Lancet*, 4 January 1986)

In the example above, all the main verbs are in the passive. The function of the passive in the texts above would seem to be to highlight and emphasize important elements in the research methodology by placing those elements in subject position, which as Lackstrom (1977) claims in English is "typically the position of the most important idea". Of course, the passive may also be functioning to maximize objectivity and impersonality often associated with written scientific prose. Now, contrast the form of the verb in the example above with its equivalent in the JRV presented below.

JRV Move 4: Describing Data-Collection Procedure

The Rotterdam trial breaks new ground because all the subjects were women pregnant for the first time. Wallenburg and his team selected them using a test of angiotensin sensitivity (AST).
(JRV.3: *New Scientist*, 30 Jan 1986)

In contrast to the RAP text, the JRV text uses the past tense to describe data-collection procedure. The past tense verbs used in the JRV text would seem to function to enable the writer provide background information to the object of discourse in the sentence. As in the sentences in the JRV text above, past tense verbs in Move 4 in the JRV tend to be preceded by some background information. It has been pointed out (Comrie, 1985: 15) that in sentences such as those in the sample text, the simple

past may be interpreted as being part of the background to the discourse.

The use of questions provides a third area of difference, especially between the JRV, on the one hand, and the RAP and ABS on the other. Elsewhere in this discussion, it has been pointed out that questions are an important syntactic feature of popularized science texts. However, as the analysis shows, they are not used in the RAP and the ABS. Questions tend to be used in the JRV to signal the commencement of an explanation of the principles and concepts that underlie a research result. An illustration of this function has been provided in 4.5.2 above. Of course, the use of questions in the JRV may have psychological implications for the audience as well. Thus, as Glaser (1975: 192) points out, "the syntactic structure of a question placed at the beginning of a new passage may be an emotional appeal to the reader in that it arouses his curiosity and interest".

Finally, the genres also differ in the way the cognitive content in corresponding moves are presented. An example of equivalent information in identical Moves across the three genres will serve to illustrate the differences which exist between them.

- 1 Transfusion affects the immune response to renal transplantation and may be associated with recurrence of various human neoplasms.
(ABS.6: *Br Med Jnl*, 30 August 1986)
- 2 Most attempts to understand the biology of Epstein-Barr Virus (EBV) begin from a common premise; that is the extreme tropism which this agent appears to display for human B lymphocytes. Certainly, the virus preferentially infects B cells in-vitro and is associated with three lymphoproliferative diseases of B cell origin in-vivo namely, infectious mononucleosis (IM), Burkitts lymphoma, and lymphoma of the immunocompromised host. It has long been thought that such tropism reflects the existence of a B-cell specific surface receptor for the virus, and indeed recent studies with monoclonal antibodies (MABs) HB5 and anti-B₂ have shown unequivocally that the 140-145 kilodalton C3d receptor molecule CR2, expressed on B lymphocytes but not on any other haemopoietic cell, has EBV receptor function.
(RAP.4: *The Lancet*, 1 February 1986)
- 3 The EPSTEIN-BARR (EB) virus is one of the five herpes viruses that attack humans. In Europe and the US, infection can produce glandular fever (mononucleosis). In Central Africa, it can lead to Burkitts lymphoma, a facial cancer in children. In China and Southern Asia, EB virus plays a part in triggering the development of nasopharyngeal carcinoma (npc). In glandular fever and Burkitt lymphoma the virus infects one type of blood cell, the B lymphocyte. In npc, however, the virus infects not the lymphocyte but the

epithelial cells at the back of the mouth.
(JRV.4: *New Scientist*, 6 March 1986)

All three examples above belong to Move I. They attempt to provide some background information by reference to established knowledge in the field. Despite this similarity, a close examination of the information contained in each of the texts shows some contrast in the cognitive content they embody. For instance, the information contained in the ABS text is precise, highly abstract and not accompanied by any background details. Thus, it assumes a readership that is sophisticated and one that has a comparable level of shared knowledge with the author. An indication of this is the lack of any form of modification, whether definite or indefinite, to the subject of **the sentence and the use of the present tense verb to make reference to a fact that holds true** for a long period of time which the writer assumes the reader knows about. The precision and the lack of modifying information may also be due to the extreme pressure for space which most ABS writers face. The ABS falls within the category of texts referred to as documentation texts meant for special bibliographical purposes (Glaser, op cit: 197). As Glaser points out, an author's abstract preceding the main body of a paper is expected to be written "in accordance with the rules laid down by the UNESCO Conference in 1966" from which various international journals issue guidelines. Two of the rules are concerned with the length of documentation texts and the need to avoid redundancy. As Brown, Day and Jones (1983) point out, a strategy adopted by mature summarizers when they are pressed for space is to express "the essential gist of a large body of text in few words". Therefore, a combination of an assumed high level of shared knowledge between the writer and the reader and the constraint imposed on the writer by the extreme pressure for space would seem to contribute to the precise and abstract form of the sentence which provides background to the study in the text above.

In the RAP text, on the other hand, the reader is provided with some background information relating to the nature of the problem underlying the study. But, the amount of details provided is such that it is only sufficient for and accessible to the specialist in the field of medicine. It would appear, considering the amount of detail

provided in the text, that the writer is indicating that he is aware of the high level of knowledge which he shares with the reader, but still considers it necessary to, as it were, refresh the latter's memory of what they both already know. Hence, he begins the sentence with "vague language" - "Most attempts to ...", suggesting that "it is no use recounting those attempts because we already know what they are, but nevertheless, they are worth mentioning". This interpretation is supported by the contention that "vague language" can convey a judgement that "in a certain context, too great a degree of precision would be out of place" (Carter, 1987: 84). Vague language can also have psychological effects. As Carter (op cit: 83) points out, the primary purpose of vague language is to produce the specific effect of detaching the producer from the absolute truth of the proposition asserted. Thus, the writer in the RAP text above attempts to reassure his reader that though he is being imprecise, he nevertheless knows the details by means of the emotive words "certainly" and "indeed" which express his personal views on the subject. Further evidence of the lack of precision in the way information is presented in the RAP text above is provided by the extraposition, "it has been thought that ..." which also presupposes that the reader is aware of the line of thought being referred to.

Like the RAP, the JRV text also provides background information, albeit based on a different assumption and a different approach. It begins with an intensive equative construction - "The Epstein-Barr virus is one of the five herpes viruses that attack humans" - whose function appears to be that of helping the reader place the subject within the context of a much more dreadful reality - the Herpes virus - with all the implications of the fears which the herpes virus evokes in the layman. In other words, implicit in the equative sentence is the assumption that the reader is less likely to be familiar with what the Epstein-Barr virus is than he would be of the herpes virus. This would seem to suggest that the writer probably assumes that the reader's knowledge of the subject-matter of discourse is not too high. In recognition of this fact, the writer tends to adopt an approach to the presentation of information in later parts of that text which can be said to be didactic in that it attempts to explain concepts

which may be unfamiliar to the reader. Such concepts include the terms "mononucleosis" put in parenthesis and Burkitt's lymphoma which is placed in apposition.

So far, an attempt has been made to show that the three genres in this study vary considerably in the linguistic features which apparently equivalent Moves across the genres use to realise the discourse they embody. The results reported here would seem to indicate that the difference between the RAP and the ABS in this respect is slight and that both genres differ distinctly from the JRV in the linguistic features manifested in their Moves.

4.5.4 Differences in the Order of Presentation of Moves

A fundamental issue in accounting for the differences in the schematic structure of information across texts is the order in which the structures or schemata are presented.

With regards to the three genres analysed in this study, the differences may be characterized as existing between the professional genres (RAP, ABS) on the one hand and the popularized genre (JRV) on the other. This characterization derives from evidence from the analysis which indicates that the ABS and the RAP adopt identical strategies in the organization of the Moves they contain. In both genres, Moves are organized sequentially in a hierarchical manner. Thus, in the RAP which may have as many as eleven Moves, there is a sequential ordering of Moves from 1-11. The same sequence is followed in the ABS, even though it does not normally realise all the eleven Moves which may be found in the RAP. The ordering of Moves in the RAP may be represented by these eleven sentences selected from RAP.3 to represent the eleven Moves which may be found in texts in the genre.

Move 1. Pregnancy-induced hypertension (PIH) and pre-eclampsia appear to be associated with ...

Move 2. Thus, Crandon and Isherwood obtained a history of aspirin intake

- during pregnancy fromⁿ primigravidae and showed that ...
- Move 3. In a randomised placebo-controlled double-blind trial, we have now investigated the effect of ...
- Move 4. To collect the required number of angiotension-11-sensitive women, healthy primigravidae attending ...
- Move 5. The patients were placed inⁿ a quiet room in the left-lateral position. An intravenous catheter was ...
- Move 6. Differences in relative frequencies between placebo and treatment groups were analysed with the F test.
- Move 7. At the beginning of the study the two groups were comparable with respect to ... In the aspirin group 5 women smoked 10-20 cigarettes/day, by comparison ...
- Move 8. There was one intrauterine death in the aspirin group, in a 25-year-old woman with uncomplicated pregnancy who presented at 41.5 weeks amenorrhoea ...
- Move 9. Our results suggest that ingestion of low-dose aspirin from 28 weeks' gestation until delivery may prevent PIH and pre-eclampsia in primigravidae ...
- Move 10. Low dose aspirin as used in our study suppresses platelet TXA₂ synthesis, as indicated by the reduction of ... However we did not measure plasma concentrations or urinary ... Thus, we can speculate that the low-dose aspirin as used in our trial prevented ...
- Move 11. Further studies are required to support or refute this hypothesis.

As has been pointed out elsewhere in this discussion, nine moves of the eleven found in RAP texts were identified in the analysis of ABS texts. Despite this fact, the order of presentation of Moves follow the same pattern as that in the RAP. That is, the omission of some RAP moves do not lead to a re-organisation of the pattern in which information is normally presented in professional journal papers. Thus, as in the RAP, Moves in the ABS are organized to follow a sequential order of

hierarchical presentation of information. In propositional terms, the order of presentation of moves in the ABS may be represented by the following sentences selected from ABS.6.

Move 1. Transfusion affects immune response to renal transplantation and may be associated with ...

Move 7. Data from patients with clonic, rectal, cervical, and prostate tumours showed an association between transfusion of ... and ...

Move 9. These data support an association between transfusion and recurrence of cancer.

Move 10. They also suggest that some factors present in great amounts in whole blood, such as plasma, may contribute to ...

Move 11. Until the question raised by retrospective studies of cancer recurrence and transfusion can be answered by prospective interventional trials with washed red blood cells, red blood cells should be transfused to patients with cancer in preference to whole blood when clinically feasible.

A number of factors may account for the similarity in the ordering of Moves in both the ABS and the RAP. First is the question of authorship. Texts in both genres are normally written by the same authors. Therefore, it is possible to argue that the similarity in the pattern of information structure in both genres is influenced by authorship. But as evidence from Urquhart (1986) has shown, even abstracts written by professional abstractors follow the same logical order of presentations as those in main research articles. Therefore, this reason may not be altogether plausible. Second, it is also possible to argue that the similarity is due to the fact that both genres share the same title and therefore address the same topic. But even JRV texts, which as we shall see later, vary from the RAP and ABS in the order of presentation of information also addresses the same research topic as the other two. A third, and more plausible reason

for the similarity may be due to the fact that the cognitive content of texts in both genres represent an account of what Gilbert and Mulkey (1984) refer to as the "empiricist repertoire" and Myers (1986) as "the narrative of science" in which time is arranged "into a parallel series of simultaneous events" and which emphasizes the conceptual structure of the discipline in their syntax and vocabulary, and by extension, their patterns of discourse organization. Other reasons for the similarity might be the need for the ABS to make access to the RAP much easier to its readers. But whatever the reason for the similarity may be, the important fact which emerged from the analysis is that there is a very close relationship existing between the ABS and the RAP in the order in which Moves are presented.

While the ordering or sequencing of information in the ABS is closely related to that in the RAP, that in the JRV differs very strongly from the order in both genres. Unlike the other two genres which seem to have a rigid order of presentation of Moves, the JRV exhibits considerable flexibility in the ordering and sequencing of the Moves which it tends to be made up of. As has been pointed out elsewhere in this discussion, the lack of rigidity in the ordering of information in JRV texts has been explained as a strategy to overcome the stifling of imagination and creativity which a rigid ordering might impose on the writer. Therefore, information in the JRV does not follow the same sequential pattern as in the RAP.

A direct consequence of a lack of a rigid format for the organization of information in the JRV is the apparent lack of uniformity in which various writers organize information in texts in the genre. An obvious observation made in the analysis is that no two texts in the data organized their moves in an identical sequence. Despite the lack of uniformity, some regularity in the ordering of a few Moves in the JRV may be observed. For instance, it was observed that there is a tendency for Moves 1 and 9 to occur as the initial Moves in texts in which they occur and for Moves 10 and 11 to occur as the last two Moves in those texts which they occur together. In between these two poles may occur any of Moves 2, 3, 4, 5 and 7 in no clearly discernable order. However, it would appear that there is a tendency for Move 2 to

precede Move 3 and for Move 7 to come before Moves 4 and 5. Following from these observations, Moves in the JRV may be characterized as belonging to three main divisions - initial, medial and final moves. The initial moves consist of Moves 1, 9, 2 and 3. The medial moves include Moves 7, 4, 5 and the final Moves are Moves 10 and 11. This characterization is supported by evidence from the texts as well as evidence from research on the presentation of news in newspapers.

First evidence from the texts shows that most JRV texts would normally organize information to reflect:

- 1 A brief statement which functions to provide some background to the problem.
- 2 The presentation of the main research outcome.
- 3 An indication of the limitations of previous efforts at resolving the problem (this function is realized in very few texts).
- 4 Presentation of the researchers who conducted the study and what they set out to achieve.
- 5 Indicating some positive results.
- 6 Indicating the methods used in data collection (this function is realized in very few texts).
- 7 Describing some of the methods used during the process of experimentation.
- 8 Discussing and explaining specific research outcomes.
- 9 Stating the main conclusions of the research or report.

In functional terms, the order of presentation of information in JRV texts would seem to be in the order shown below. The Move labels on the left hand column are those assigned in the analysis of the RAP.

- Move 1: - Presenting Background Information
- Move 9: - Highlighting Overall Research Outcome
- Move 2: - Reviewing Related Research

- Move 3: - Presenting New Research
- Move 7: - Indicating Consistent Observations
- Move 4: - Describing Data-Collection Procedure
- Move 5: - Describing Experimental Procedure
- Move 10: - Explaining Specific Research Outcomes
- Move 11: - Stating Research Conclusions

Second, evidence from research on the structure of news in the press would seem to support the order of presentation reported above for the JRV. For instance, Van Dijk (1985: 82) contends that the writing strategy of news discourse production proceeds according to the following production principles:

- (a) Important consequences come first.
- (b) Details of an event or actor come after overall mentioning of the event or person.
- (c) Causes or conditions of events are mentioned after the event and its consequences.
- (d) Context and background information comes last.

Van Dijk's characterization of the strategy adopted in the presentation of news report in newspapers is almost similar to that posited above for the JRV, except that it places background information last. Even, then, Van Dijk's background information may be equated with the information contained in Move 10 in the JRV which functions to explain specific observations by, amongst other means, explaining principles and concepts underlying such observations.

Furthermore, according to Burkett (1973) most journalists tend to write according to the "DEE-System" of organization. The DEE stands for Description, Explanation and Evaluation. This will seem to relate to the three divisions - initial, medial and final moves - into which I contend that Moves in the JRV may be ordered.

Burkett further adds that "writers of the DEE persuasion must first identify the main theme or major thrust of the story for the reader clearly and plainly and the reader must encounter this statement of significance very soon in the story.

These studies provide evidence to support the tendency for Move 9 to occur almost as the initial Move in a JRV text. It is the move that best portrays the "important consequences" of the research report or "the main theme or major thrust" of the research report. They also support the occurrence of Moves 10 and 11 as the concluding Moves in the text. The medial moves would seem to conform with Van Dijk's details of events or actors, causes and conditions of events etc. Having said this, we are still left with the important question of why, for instance, do JRV texts have this kind of schematic structure and what effect they have on the reception of the message by the reader.

Despite the fact that JRV writers do not have any standard format to which they must conform, the pattern of presentation of Moves in JRV texts cannot be fortuitous and unmotivated. As with most texts, the order of presentation of Moves in JRV texts is constrained by social, cognitive and cultural factors. It is the result of the interaction between the social and professional routines of science journalists in institutional settings, on the one hand, and the need for an effective cognitive processing strategy of popularized science information by both the writers and the readers, on the other. Thus, as Burkett (1973: 178) points out, newspaper journalism once demanded a "five-w" beginning. The "five-w" beginning means packing the "who-what-when-where-and-why" into the first few sentences. But Burkett adds that now editors permit and even demand more. As a result newspaper journalists tend to begin their stories with a "lead" aimed at hooking and holding the reader. It has been pointed out (Funkhouser 1970, Burkett, op cit) that this need is conditioned by the fact that newspapers and magazines must compete for a reader's attention in the open information market. Therefore, any journalistic writing which fails to take this need into consideration stands the risk of losing the reader to numerous other competing attractions from other news stories within the newspaper or elsewhere.

JRV conforms to this characterization, for opening most of its stories with a lead which provides some background information and following it up with the overall research outcome, the writer is demonstrating his awareness of the need to provide the reader with the information he needs most to be hooked, held and drawn into reading the text. Again by also presenting information about main research results (Move 7) before those concerned with data-collection (Move 4) and research procedure (Move 5), the writer is also demonstrating the need to always project within the text information that will guarantee that the reader's attention is held and sustained.

In the course of this chapter, it has been demonstrated that of all three genres in this study, the RAP is the only one with a fixed surface organizational format. The analysis also revealed that all three genres have identifiable schematic structures and "constituent elements". However, considerable variation was observed in the number of schematic units and their constituents realized across the three genres. Considerable variation was also observed in the pattern in which schematic units are organized in text across the genres. The most obvious variation in this regard is between the RAP and the JRV. These variations in the distribution and organization of schematic structures in the genres was shown to relate to the pragmatic conditions of the interaction between writer and audience as well as that between writer, medium and purpose of discourse.

CHAPTER FIVE

VARIATION IN THEMATIC PROGRESSION PATTERNS

5.0 INTRODUCTION

In Chapter Four, an attempt was made to investigate the structure of information in the three genres in this study using Swales' (1981a) functional schematic approach to the analysis of Article Introductions. The aim was to determine the extent to which the three genres differ in their patterns of schematic structure.

Swales' "schematic structures" approach represents one major current of text analysis from which information structure may be characterized. It is concerned with characterizing more extended and global structures of text by presenting discourse elements as knowledge structures, which is in this study, built on the interaction of "Moves" and their "constituent elements". This is considered an appropriate starting point from which to probe the phenomenon of variation in information structure, for it provides an insight into the differences which may exist between texts at the global level of discourse organization. Whereas it is extremely beneficial to investigate the organization of information in the three genres from this perspective, it alone cannot provide all of the answers needed to account for differences in the way information is structured in the genres. If we truly are to understand the differences which exist between the three genres in this study in their patterns of information structure, the texts which make up the corpus will need to be subjected to further analysis using additional parameters for characterizing the structure of information in text.

Therefore, in this chapter, more questions are raised and answers to them are hopefully provided through an examination of the phenomenon of variation in information structure in the three genres at another level of text - the level of the progression of utterance themes in text. Apart from being motivated by the need to provide an in-depth understanding of how the genres differ in terms of information

structure, the study reported in this chapter is also necessitated by the fact that most genre-based studies of text have tended to concentrate almost exclusively on the analysis of schematic structures (see Swales 1981, Dudley-Evans, 1986, 1987, Crooks 1984, Cooper 1985 etc). So far, no study has to the best of my knowledge, in addition to investigating schematic structures, also attempted to characterize the structure of information in genres in terms of "Theme" and "Rheme", certainly none from an FSP (Functional Sentence Perspective) orientation, particularly from the standpoint of "Thematic Progression (TP)".

In view of the above facts, the discussion in this chapter is centred around the concept of "Theme", but particularly on the FSP notion of Thematic Progression (TP). It provides answers to the following questions:

- 1 What linguistic elements characteristically function as theme in each of the genres which may serve to distinguish one genre from another?
- 2 In what way is the progression of theme and rheme distributed across utterances in each of the genres?
- 3 Does the progression of theme and rheme differ in any respects across the three genres?
- 4 If so, what kind of contextual factors may account for the differences?

In accordance with these questions, this chapter discusses the concept of theme, reviews two schools of thought on the concept - the systemic-functional and the FSP models, with emphasis on the latter. Some FSP-based studies on TP are also examined and criteria used in identifying theme in the analysis outlined. TP patterns in the genres are exemplified after which a discussion of differences in TP patterns across the genres is reported.

5.1 THE THEME-RHEME CONCEPT

A lot has been written about the "Theme-Rheme" concept in linguistic literature. Some of these writings give the impression that linguists differ on the most appropriate term to describe the concept covered by the Theme-Rheme concept. For instance, some linguists use the term "old-new", "known-new", "topic-comment", "presupposition-focus" instead of the concept "Theme-Rheme".

These terminological differences notwithstanding, the term "theme" is generally used to imply one or all of the following:

- 1 The item(s) which occur(s) as the initial element in a clause (Halliday, 1967).
- 2 The elements which refer to something which has been mentioned and can easily be identified from information contained in the context of discourse (Danes 1970, 1974).
- 3 The item(s) which represent what the speaker wants to speak about, ie. the topic (Li and Thompson 1976, Keenan and Schiefflin 1976).

The "Theme-Rheme" concept is intuitively appealing as a functional paradigm for explaining how information in utterances is tailored or packaged by a writer to meet the assumed needs of the reader. It is this function that makes the notion relevant for explaining differences which exist in the way information is packaged in the three genres which constitute the data for this study.

5.2 APPROACHES TO THE CHARACTERIZATION OF THEME

A number of linguistic traditions exist which attempt to account for the nature of thematic relations which are present in discourse (Tagmemics, Systemic-Functional, FSP, Semantic theories). Each of these traditions defines different conceptions of theme and represents different approaches to the characterization of theme in

discourse. Thus, the appropriacy of any one theory to a linguistic investigation depends on the purpose of the investigation and the level of delicacy of the linguistic analysis. In view of this, the discussion in this section is centred on the Systemic-Functional and the FSP approaches, both of whose conceptions and methods for characterizing theme are relevant for investigating the kind of problems which this study is interested in. These two approaches have been characterized by Fries (1981) as the "separating approach" and the "combining approach" respectively. Each of them represents ideas from two linguistic traditions. The separating approach owes its origin to the Neo-Firthian tradition. Its exponent is Halliday along with linguists with whom he shares ideas of systemic-functional grammar. The combining approach, on the other hand, owes its origin to Mathesius, and is advocated by linguists of the Prague School.

5.2.1 The Systemic-Functional Approach

Halliday (1967) identifies two separate structures within the "Theme System" of English into which information contained in a clause may be organized. They are "Information structure" and "Thematic structure". According to Halliday (op cit) information structure is the distribution of discourse into information units. It consists of blocking the information units in a piece of discourse into two functional categories - "Given" and "New". These functions are realized phonologically by tonality, that is, the distribution of the message into tone groups. Thus, the message in a piece of discourse is characterized as consisting of an optional segment corresponding with the "pretonic" segment of an information unit. This segment conveys "Given" information. The other segment which is obligatory corresponds with the "tonic" segment of the information unit and conveys "New" information. "New" information is characterized "not in the sense that it cannot have been previously mentioned ... but in the sense that the speaker presents it as not being recoverable from the preceding discourse". Consequently, "Given" information is defined as "that which is known or

at least obvious in the given situation". From the foregoing, it would seem that the concept of information structure is not particularly suited to the analysis of written discourse since it depends on tonality for the characterization of "Given" and "New" information in discourse. In other words, the concept of information structure as conceived by Halliday is better suited to the analysis of spoken discourse.

With regard to thematic structure, Halliday (op cit: 211) points out that thematization is a choice which has its origin in the clause, not in the information unit, assigning to the clause "a structure in terms of the functions "Theme" and "Rheme". Distinguishing between information structure and thematic structure, Halliday (op cit: 212) contends that:

"Given" means 'what you were talking about' (or 'what I was talking about (or what I was talking about before)'), "theme" means 'what I am talking about (or what I am talking about now).

He points out that Information Structure is one aspect of the thematic organization of discourse, but that the functions "Given" and "New" are not the same as those of "theme" and "rheme".

According to Halliday (ibid: 212) "the theme is what comes first in the clause"; "what is being talked about, the point of departure for the clause as a message". Every other thing beside what comes first is assigned the function "rheme". Thus, in unmarked declarative, polar (Yes/No) interrogative and non-polar ('WH'-) interrogative, the element that occurs in first position "is the natural theme".

Halliday's approach leaves some questions unanswered. For instance, it is not quite clear what "point of departure" and "aboutness" really mean when elements other than the subject or object NP are fronted. Again, it is also not clear what constitutes the domain of "the point of departure". These led Lowe (1987: 6) to pose the following questions:

- In what sense is a fronted temporal or fronted gerundial or fronted past participial clause or an initial - if clause a point of departure?
- Is an initial element the point of departure for a clause and for the clause only? Or can it function as a point of departure for a stretch of discourse larger than the clause and if so for how far?

It is partly for the reason of questions such as these, particularly the second, that Halliday's approach was considered unsuitable for the kind of analysis reported in this chapter - a fact further underlined by the contention by Fries (1981: 4) when he states as follows:

No real argument has been brought forward to justify the statement that the theme or the beginning of a group, clause or sentence means 'the point of departure of the message expressed by that unit ... Because Halliday was not focusing on the meaning of theme, the examples he cites in his articles are often made-up ones and are rarely provided with context.

5.2.2 The Functional Sentence Perspective (FSP) Approach

In the tradition of the Prague School of Linguistics, as in Halliday's Firthian tradition, language exists to serve a communicative function aimed at satisfying various socio-psychological needs. Following in this tradition, a number of linguists of the Prague School proposed a functional approach to the analysis of sentences or utterances generated in a given language situation. The approach, commonly referred to as "**Functional Sentence Perspective (FSP)**" is a model for explaining how information is organized in a sentence or an utterance, discourse or text in accordance with how language functions in situational contexts.

Central to FSP is the theory of Thematization. The origin of FSP theory on thematization is often associated with Mathesius (1929, 1936). However, Prague

School linguists also point out that the basic principles underlying FSP conceptions of theme had been discussed by Weill (1884, 1887).

Weill identified two different orders that must interact in speech as "the syntactic march" and "the march of ideas". Writing on the need for shared knowledge in spoken discourse, Weill, cited in Maynard (1986: 78) states that it was necessary to lean on something present and known in order to reach out to something less present, or unknown". In addition, Weill identifies two aspects to the use of language in speech as "the point of departure" or the ground upon which the two intelligences meet" and "another part of discourse which forms the statement". Like Weill, Mathesius (1939) makes the distinction between "point of departure or known ("given information" and "statement" or new information). But unlike Weill, Mathesius characterizes the categories as "theme" and "rheme" respectively. Mathesius (cited in Danes 1974: 106 and Firbas 1964: 268) defines theme as "the starting point of the utterance which is known or at least obvious in the given situation and from which the speaker proceeds". Thus, theme is defined on the basis of two concepts: (a) information which is known or obvious in the situation, and (b) information from which the speaker proceeds. In other words, what determines the theme of an utterance is not necessarily the position of the element in the clause or utterance, but "how the information expressed in that sentence relates to the information already available in the linguistic and non-verbal contexts" (Fries, op cit). Therefore, unlike the systemic-functional approach, no distinction is made between the functions "Given" and "New" on the one hand and "theme" and "rheme" on the other. Instead, both groups of functions are said to define one and the same concept. Consequently, both Halliday's "Given/New" and "Theme/Rheme" are subsumed under the functions "Theme" and "Rheme".

Mathesius' conception of Theme has attracted some criticism. For instance, Firbas (1964: 208) cites Travnicek (1962: 163-171) as contending that "conveyance of known information is not the essential feature of the theme" arguing that to regard

Theme as such is to narrow the conception of theme. According to Firbas, Travnicek also criticises Mathesius for applying the narrow conception "even to sentences that cannot possibly convey any known information", hence he talks about "anticipatory themes" and about sentences that have no themes at all. Thus, Travnicek argues that an essential feature of theme must be its ability to be truly general in character and to remain invariably the same in all cases. He, therefore, defines themes as "the sentence element that links up directly with the object of thought, proceeds from it and opens the sentence thereby".

In view of Travnicek's criticisms of Mathesius' conception of theme, Firbas (1964) proposes the concept of **Communicative Dynamism** (CD) as a means of arriving at a truly general conception of theme and one that can be used to identify theme in all cases. By Communicative Dynamism, Firbas (op cit: 270) means the extent to which the sentence element contributes to the development of the communication, to which it "pushes the communication forward", as it were. Explaining, he states that:

It is obvious that elements conveying new, unknown information show higher degrees of CD than elements conveying known information. But even within a sentence section made up entirely of elements conveying new information, the degrees of CD are not the same (homogeneous).

Firbas (1964: 270)

Based on the above contentions, Firbas (op cit) defines theme as "the sentence element or elements carrying the lowest degree(s) of CD within the sentence" and rheme as "the sentence element(s) carrying the highest degree(s) of CD within the sentence". In addition to these two functional categories, Firbas also identifies a third function which he refers to as **Transition**. Transition carries an "Intermediate" degree of CD. Thus in Firbas characterization of the FSP of a sentence, there are three degrees of CD: Theme (with lowest CD); Transition (with intermediate CD) and Rheme (with highest CD). Firbas proposes the question test method for identifying

theme and rheme in a sentence. The question test method may be illustrated with this example from text RAP I below:

Recurrent spontaneous abortion has been treated by means of immunisation with paternal cells and cells from multiple unrelated donors.

The utterance above may be made in response to the question: "How has recurrent spontaneous abortion been treated?"

The relevant information needed to answer the question is: "by means of immunisation with paternal cells and cells from multiple unrelated donors". This information constitutes the "Rheme", while the rest of the information in the utterance is the "Theme". The purpose of the question test is to isolate the rheme. After the rheme has been isolated, what remains of the sentence would be identified as the theme and transition respectively. Jones (1977: 70) has pointed out that the question test while being a useful device for identifying theme and rheme in most sentences cannot always apply in all sentences.

Attempts have also been made to relate the FSP of sentences to the notions "Topic" and "Comment" (Dahl 1974, Sgall 1970, Sgall, Hajikova and Benisova 1973). Dahl (1974) defines "topic" as that part of the sentence "where we name or define a set or an individual", and "comment" as a propositional function "predicated" about this set or individual. Like Dahl, Sgall et al adopt the notion "topic" and "comment" instead of "theme" and "rheme". They also incorporated Firbas notion of CD in their description. Thus, they define topic as contextually bound elements known to the reader either "from the context, from the situation or from general conditions of the given utterance". They distinguish these contextually bound elements (topic) from the contextually non-bound elements (comment) or what they refer to as those elements not recoverable from the context of situation. Sgall et al also propose the question test method as a means of isolating "comment" and deducing "topic". They elaborate on the ordering of contextually non-bound elements and how they can be identified. They

also posit the ranking of elements into a hierarchy known as the "syntactic" ordering of participants' in which the elements are ranked into actor, time, place, manner, instrument, dative, object of type "about what", object (patient), direction, object complement, condition, purpose-cause etc.

These works by Firbas, Dahl and Sgall et al reported here are modifications of the earlier conception of "Theme" and "Rheme" by Mathesius. While they may differ from Mathesius' work in some respects, they share one thing in common with it; that is, they do not make a distinction between "Theme" and "Rheme" and "Given" and "New" as the systemic-functional approach does. Again, while they attempt to posit a means by which the FSP of all kinds of sentences may be arrived at, they do not discountenance the importance of the principle of "known" or recoverable information. As Firbas (1964: 272) points out:

"All this does not mean, however, that the criterion of known or unknown information could be dispensed with. Even if it does not cover the essential feature of the theme, it still renders much valuable service. On the one hand, it determines sentence elements that convey known information or information to be gathered from the verbal or situational context, ie. sentence elements through which a sentence becomes contextually dependent (and which are always thematic). On the other hand, it simultaneously determines sentence elements that convey new information and constitute the contextually independent section of a sentence ... The criterion of known or unknown information is further applicable in determining the degree of contextual dependence of a sentence as well as that of its contextual applicability (ie. the range within, or the extent to which a sentence form may become contextually dependent).

This observation by Firbas has been quoted at such great length because it lends powerful support to the decision in this study to base the analysis of the thematic organization of information primarily on the criterion of known or unknown information while also drawing insights from the criteria of "CD" and "Topic/Comment" both of which have been characterized by Jones (1977: 91) as being "very sentence bound". This takes us to another modification to Mathesius'

characterization of theme, the concept of Thematic Progression (Danes 1964, 1970, 1974) upon which the analysis in this chapter is based.

5.2.3 FSP and Danes' Theory of Thematic Progression

Unlike the FSP models of thematization discussed above which are very sentence bound, Danes' papers (1964, 1970, 1974) are attempts to show that the organization of information in whole texts, as opposed to just sentences, is determined by **Thematic Progression** (TP). By Thematic Progression Danes (1974: 114) means "the choice and ordering of utterance themes, their mutual concatenation and hierarchy, as well as their relationship to the hyperthemes of the superior text units (such as paragraph, chapter ...), to the whole of text, and to the situation".

Like other Prague School linguists, Danes (1970: 134) identifies two parts to the organization of information in an utterance as "Theme (something that one is talking about, TOPIC) and Rheme (what one says about it, COMMENT). He associates these functions with the related notions "Old, already known or given elements" functioning as "a starting point of the utterance" and the other which conveys "a new piece of information (being the core of the utterance). He, however, adds that the distinction between "Rheme/Theme" and "Given/New" was not necessary in determining TP since the two aspects would normally coincide in an analysis of the FP (Functional Perspective) of discourse. In addition, Danes (1974: 107) argues that the distinction "is an incomplete dichotomy" because the differentiation only concerns "the first members of the two pairs (ie. the known (given) piece of information vs theme), while the second members are identical, viz the core of the utterance or the rheme (what the speaker says about the theme). On the basis of this observation, Danes further argues that attempts to distinguish between both groups of functions and to argue as Halliday (1967) has done that "thematization is independent of what has gone before" is to

suggest that thematization is irrelevant in respect to the organization of text". But according to Danes (1974: 109),

Such a conclusion appears very doubtful in the light of the fact that the choice of the themes of particular utterances can hardly be fortuitous, unmotivated, and without any structural connection to the text.

Concluding, he asserts that:

In fact, even a superficial observation of texts shows that the choice and distribution of themes in the text reveal a certain patterning; this statement also corresponds to our intuitive expectations that the progression of the presentation of subject-matter must necessarily be governed by some regularities, must be patterned.

Pointing out that the Functional Perspective (FP) of discourse above the level of the sentence is neither dependent on lexical meanings nor on semantic realisations of surface syntactical functions, Danes (1970: 135) states that:

... the FP structure of utterances simply depends neither on the meanings of the particular lexical items contained in one utterance, nor on the semantic functions of the syntactic elements of the sentence (such as agent, action etc).

But, according to Danes (op cit):

... the FP should be interpreted in terms of the different degrees of actual communicative value of various elements of utterance in a particular act of communication. And as the two parts of utterance that constitute the framework for the distribution of this value might be termed, Theme (topic) and Rheme (comment), respectively, we might call the FP organization of utterance (and text) the "Thematic structure of the utterance" (in contradistinction to the semantic structure of the sentence).

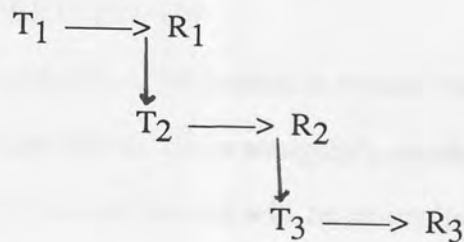
Based on the conceptions referred to above, Danes (1970) contends that the thematic organization extends "beyond the level of single utterances and the structure of the whole text", working in utterance sections as well as in utterance groups (paragraphs, section, chapter) or text (discourse) as a whole into a structure of

"Thematic Progression (TP)". He underlines his argument by identifying three main patterns of Thematic Progression which may occur in text as:

1 SIMPLE LINEAR TP (or TP with linear thematization of rhemes).

Danes (1974: 18) refers to it as the most elementary, basic TP. In this TP pattern, each rheme becomes the theme of the next utterance. It is represented thus:

Figure 1: Danes Simple Linear TP Pattern



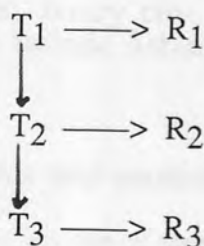
Danes illustrates this progression type with the example below:

The first of the antibiotics was discovered by Sir Alexander Fleming in 1928. He was busy at the time investigating a certain species of germ which is responsible for boils and other troubles.

2 TP WITH A CONTINUOUS (CONSTANT) THEME

This refers to the TP pattern in which the same Theme appears in sequence in a series of utterances, though not necessarily making use of identical wordings. Danes represents it thus:

Figure 2: Danes Constant TP Pattern



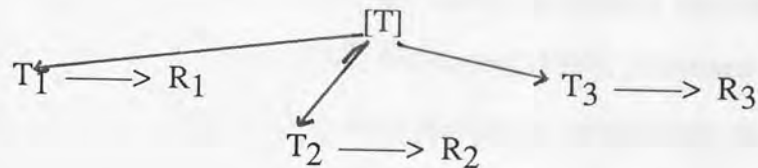
He illustrates the Continuous or constant TP pattern using the example below:

The Rousseauist especially feels an inner kinship with Prometheus and other Titans. He is fascinated by any form of insurgency ... He must show an elementary energy in his explosion against the established order and at the same time a boundless sympathy for the victims of it ... Further the Rousseauist is ever ready to discover beauty of soul in anyone who is under the reprobation of society.

3 TP WITH DERIVED T'S

Danes characterizes it as the pattern in which "the particular utterance themes are derived from a "hypertheme" (of a paragraph, or other text section) in which "the choice of the derived utterance themes will be controlled by various special (mostly extralinguistic) usage of the presentation of subject matter. It is represented thus:

Figure 3: Danes TP with a Derived Theme

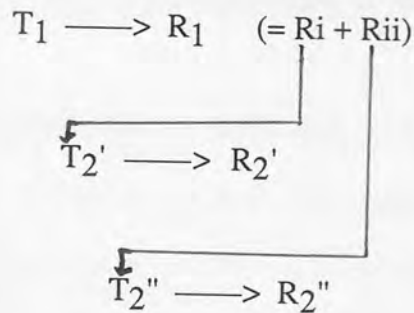


He exemplifies the TP pattern (TP with Derived Theme) using the short text below:

New Jersey is flat along the coast and southern portion; the north-western region is mountainous. The coastal climate is mild, but there is considerable cold in the mountain areas during the winter months. Summers are fairly hot. The leading industrial production includes chemicals, processed food, coal, petroleum, metals and electrical equipment. The most important cities are Newark, Jersey city, Paterson, Trenton, Camden. Vacation districts include Asbury Park, Lakewood, Cape May and others.

The three main TP patterns may combine in various forms to produce other patterns. Danes (1974: 120) identifies one of those TP patterns which may result from such combinations as THE EXPOSITION OF A SPLIT RHEME, represented thus:

Figure 4: Danes TP with Exposition of a Split Rheme



Danes illustrates this TP pattern with the example below:

All substances can be 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 ...

5.3 SOME TP-BASED ANALYSIS OF DISCOURSE

Some studies with a theoretical-descriptive focus have attempted to account for the structure of information in discourse based on Dane's theory of Thematic Progression (Dubois 1987, Giora 1983, Weissberg 1984, Maynard 1986, Curzon 1985, Taylor 1983). Other studies have tended to emphasize the pedagogical implications of Danes TP patterns (Vande Kopple 1983, Kraigher-Vesel 1985, Witte 1983). We shall limit the discussion in this section to the former.

Dubois (1987) attempts to reformulate the theory of Thematic Progression. Material for the study is taken from an academic meeting of the Federation of American Societies for Experimental Biology. The study involved the analysis of theme and thematic progression in 85 independent clauses. Theme was defined, following Halliday (1968) as "point of departure for the clause as message". Dubois investigated the clauses, first in terms of Danes three main patterns of TP - simple linear, constant and hypertheme TP patterns - and stated that these may be reformulated into two main types -

- (a) thematic - from a previous theme or themes
- (b) rhematic - from a previous rheme or rhemes

According to Dubois (op cit: 109), each of these main types can be simple, ie. from a single source. If this is the case, they may be "contiguous" or "gapped" so that repetitions of theme at any degree of givenness can be called "Constant thematic". On the other hand, progression can be "Multiple" (from more than one source), "with at least two sub-types (re)integration and separation". In this way, Dubois avoids Danes third pattern, the "Hypertheme" which she considers a case of constant TP pattern, albeit an indirect type.

Dubois' study is very insightful. The model it proposes may help to resolve practical problems associated with the analysis of TP in authentic discourse. However, there appears to be a methodological inconsistency in the study. By defining theme simply as point of departure of the clause as message" in accordance with Halliday (1968), while not emphasizing the criterion of givenness and recoverability of information, Dubois departs from an issue of fundamental importance to Danes theory of thematic progression.

Giora (1983) makes "an attempt to extend the notion of cohesion beyond the sentence level, by viewing it as a requirement of text for connectedness between segments larger than a sentence, such as paragraphs, or whole chapters". Giora adopts one of Danes (1974) TP patterns - the simple linear TP - and attempts to account for the cohesive relations which hold at the level of the paragraph and chapter. In the study which is based on verses or paragraphs from selected poems, Giora found that the concatenation of new discourse topics motivates text segments in a way that creates informational hierarchies.

Giora's work is a contribution to the theory of TP in that it demonstrates that the theory can apply at all levels of text, not just at the level of utterance. However, in

basing the analysis of text segmentation on the paragraph, Giora would seem to have over-simplified the concept of segmentation and segment cohesion in discourse.

In Weissberg (1984) an attempt is made to verify the existence of, and the extent to which Danes' three main TP types (the simple linear, constant and hypertheme) occur in published experimental research reports in Biology, Engineering and Agriculture. Weissberg examined 60 paragraphs taken from the Materials and Methods and Discussion sections of experimental research reports. The results section of the research reports was not included in the study. Weissberg found that:

"In general, research report writers regularly use topic development patterns based on the given/new contract in their published works".

And that:

"In those paragraphs where patterning did occur, the linear pattern was found to be the most common, being especially frequent in Introduction and Discussion paragraphs. The constant topic patterns occurred least frequently.

Weissberg's study is a useful contribution to the study of the place of TP on discourse production. However, it is doubtful to what extent his observations may be said to be representative of the strategy of discourse production in the genres he examined since it was based on selected paragraphs rather than whole texts and since the Results section was not included in the study. On the other hand, the study is significant in being one of the relatively early studies to investigate Danes' TP patterns in experimental research reports.

Maynard (1986) examines the interactional aspects of thematic progression in English casual conversations. The study adopts Danes' (1974) patterns of TP to account for the organization of information in conversations. In addition to established concepts such as "Theme" and "Rheme", Maynard introduces new notions such as "Thematic-turn", "Rhematic-turn", and "Stray-turn", otherwise T-turn, R-turn and S-turn respectively.

Following Danes (op cit), Maynard proposes "three patterns of interactional thematic progression patterns based on exchanges of turns" - thematic turn pattern with linear progression; thematic turn exchange with a constant theme; and thematic turn exchange with derived theme. Concluding, Maynard points out that what is unique in the study is that it takes into account thematic structure which is primarily an information-based organization principle unlike earlier studies of conversation organization which concentrates on turn-taking and speech act categories.

Maynard (1986) is certainly a contribution to research in Thematic Progression and conversation analysis in the innovations it introduces to both areas of linguistic enquiry. The insight which the study brings to bear on a vast array of naturally occurring linguistic data is no doubt useful to anyone interested in adapting Danes' model of TP to extensive data.

Up till this point, the discussion in this chapter has been concerned with reviewing relevant literature on the concept of theme and thematic progression. The discussion has been deliberately elaborate because I recognize the fact that though the concept of thematic progression is acknowledged as a unique contribution to the functional study of language and discourse, not much is known about its basic principles, because of lack of easy access to published works by linguists of the Prague School.

In view of this, the discussion so far serves two purposes. First, it puts the investigation to be reported in this chapter in perspective, especially in the context of existing studies in the area of thematic progression. Second, it provides the reader who has some interest in the subject-matter of discourse in this chapter in particular and the thesis in general with sufficient background knowledge with which to assimilate the discussion in the remaining sections of this chapter. Having dwelt, as it were, on general issues related to the concept of thematic progression, we shall proceed in the

next few sections to discuss issues which are specific to the analysis and results to be discussed later in this chapter.

5.4 THE VERBAL AND SITUATIONAL CONTEXT

It has been pointed out in this study (Chapter Three) that the data for the study reported in this chapter consist of representative paragraphs from the RAP as well as whole texts from the ABS and the JRV. The procedure adopted in constituting the data was also shown to be one that ensured representativity and comparability of texts across the three genres. It was also pointed out that the basic unit of analysis adopted in this study is the **utterance** defined by Danes (1964) as the minimal communicative unit of a given language. Thus, the utterance was said to relate to the clause or the sentence. However, this is not to suggest that the term utterance as used in this study is necessarily synonymous with either the clause or the sentence, at least in semantic terms. Unlike the clause and the sentence which are syntactic units, the utterance is conceived of as a functional unit. As such, it does not consist of an abstract structure or a static configuration with a pattern of distinct features. Rather it is a functional unit whose structure is dependent on the conditions of the act of communication as determined by "the general character and regularities of linear materialization and linear perception", as well as "by extra-linguistic content of the message, by the context and situation and by the attitude of the speaker towards the message and towards the addressee" (Danes, 1964). That structure is the division of the utterance into two portions: "the theme (or topic), conveying the known (given) elements, and the rheme (or comment), conveying the unknown (not given) elements of an utterance. It is on this structure that the thematic organization of utterances in this study is based.

Therefore, thematic organization is discussed in this chapter at the level of organization of utterances. The domain of this organization has to do with all that is connected with the "processural aspect of utterance". That is, "the dynamism of the relations between meanings of individual lexical items in the process of progressive

accumulation, as well as the dynamism of all other elements of utterance (semantic and grammatical too), arising out of the semantic and formal tension and of expectation in the linear progression of the making up of every utterance (Danes, op cit: 227).

In line with all the above, and as with Danes (1974: 112) the conception of theme adopted in this study is that which characterizes theme not only as "the point of departure" but also "as the opening element of the sentence that links up the utterance with the context and the situation, selecting from several possible connections one that becomes the starting point, from which the entire further utterance unfolds and in regard to which it is oriented". This definition emphasizes the criterion of givenness and recoverability of information in the determination of utterance theme. It ensures that the utterance is considered in relation to the mass of information accumulated up to the point in text preceding it. The amount of accumulated information from the preceding text which is deemed recoverable in the utterance is designated as theme, irrespective of its position in the structural configuration of the utterance (ie. whether at initial or end positions). Thus, this study recognizes two sequences to the thematic organization of information in the utterance: the 'OBJECTIVE sequence', in which the T precedes the R, and the 'SUBJECTIVE sequence', in which the R precedes the T as does Mathesius cited in Danes (1970: 135).

In adopting these positions, I am aware of the problems they raise for the identification of theme. One of them is the fact that the boundary between theme and rheme is often not very sharp. This problem is resolved in this study by applying other FSP theories on thematization such as Firbas (1964) theory of "Communicative Dynamism" and Firbas (1975, 1981) concept of "Appearance/Existence on the scene" in FSP. The other problem that may arise is the fact that some utterances may be made up entirely of information which is recoverable from preceding context. If an utterance should be made up wholly of recoverable information, elements of information in the utterance which forms the point of departure, or which opens the utterance and links it up with the context and the situation described in that utterance will be designated

theme. That is, initial elements in the utterance will be theme. This is in line with Travenicek's (1962) clarification of Mathesius' conception of theme. The third problem relates to the recoverability span of information which may be assigned thematic status in an utterance. With regards to this, Firbas (personal communication) points out that:

The problem of the length of text (the recoverability span) during which a piece of information may remain recoverable after its last occurrence in the text has not yet been definitely solved. It is, however, important to note that in analysing texts one finds that the retrievability spans are comparatively very short, so short in fact that the need of setting a limit to where the span in the preceding text can begin frequently seems to be unnecessary.

The fourth problem may arise with the identification of theme in initial sentences in text in which it may be argued have no preceding context from which information may be recovered; what Prince (1981: 235) refers to as "New discourse entities". Prince (op cit) points out that new discourse entities may be "Brand new" "akin to going out and buying a suckling pig"; they may also be "unused" in which case "the hearer may be assumed to have a corresponding entity in his/her own model and simply has to place it in (or copy it into) the discourse model". The initial utterances in the texts analysed in this study fall within the category of "unused" New entities - because their readers are either specialists or semi-specialists in the field of discourse. Therefore, the starting point of the utterance which links the utterance up with the readers assumed corresponding entity and from which the discourse proceeds is designated theme. As Dubois (1987: 108) points out:

... a theme is either recoverable or unrecoverable. On the one hand, if the theme is new, it derives from the speaker/writer's thoughts or surrounding physical or mental environment, and we can hardly say more about it; beyond the label. On the other hand, if the theme is not new, it has to have progressed from a preceding part of text.

The final problem relates to the identification of theme in "composed" and "condensed" utterances. Danes (1974: 116) defines a composed utterance as an utterance formed by the combination of two or more simple utterances into a single sentence frame. In syntactical terms, it would correspond to a paratactic construction. Danes (op cit) states that in composed utterances, "if the T's or R's of the two utterances are the same (from the semantic point of view), they will be mentioned only once". This interpretation relates to that by Fries (1981: 6) that since the order of the clause in paratactically related complexes cannot be changed, it could be said that there is no sentence level theme. Therefore, each of the independent clauses in the paratactic relationship will be considered as if they were two simple sentences which are independent of one another. Danes (op cit) defines condensed utterances as those utterances based on fusion. That is, where two simple utterances which get fused into a single utterance share a common FSP element, either by way of thematization or rhematization of one of the utterances. This would seem to correspond with a hypotactic construction. There isn't much problem with the identification of theme in condensed utterances. Since the two utterances in a condensed utterance share a common FSP-element, thematic status is assigned to the FSP-element which conveys the "point of departure" for the utterance as a whole.

In conclusion, every simple utterance is considered to have a Theme-Rheme (T-R) nexus, except, perhaps, imperatives. Composed utterances may have more than one T-R nexus. Condensed utterances are taken to have a T-R nexus. Exceptions to these guidelines, if any, are indicated wherever they are encountered in the analysis. The development of communication is measured by the pattern of progression from one T-R nexus to another. The process of progressing from one nexus to another is numbered from 1 to -nth progressions and the regularity of TP patterns measured in accordance with the number of progressions of any one TP type identified in a genre.

5.5 THE IDENTIFICATION OF UTTERANCE THEMES

In the previous section, we clarified issues related to the context and situation specific to the analysis of TP in this study. In this section, an account is given of the principles adopted in the identification of theme in utterances in the corpus.

Given the fact that theme is conceived in this study as information which is "Given", "known" or recoverable from preceding context, it becomes necessary that the grounds upon which certain utterance sections are designated as thematic and others as rhematic be explained. Theme in the model adopted here links an utterance to the utterance preceding it and to the context and situation of discourse in the text. Therefore, its identification is not highly dependent on surface structure syntactic descriptions. As has been pointed out elsewhere in this chapter, Danes (1970: 135) stresses that "the concept of FP should be identified neither with the semantic content and structure of the underlying sentence, nor with the amount of the selective information rendered by the parts of the sentence". Rather he contends that "the FP structure should be interpreted in terms of the different degrees of actual communicative value of various elements of utterance in a particular act of communication". This, contrasts with the views held in some other models of thematization, particularly Halliday (1967: 212) who not only identifies the theme of a clause as "what is being talked about, the point of departure of the clause as a message", but more specifically with the initial element in the clause. It is in line with this, that Brown and Yule (1983: 127) contend that "in many cases (often considered to be the neutral and unmarked cases) the theme of declarative sentences will be a noun phrase (the grammatical subject), that of interrogatives the interrogative word and that of imperatives the imperative form of the verb". While this is no doubt a very useful way of conceptualizing theme, particularly in single sentences taken out of context, it is doubtful whether such a conception would be useful in accounting adequately for the progression of thematic information from one sentence to another, particularly in naturally occurring language data as opposed to texts written for the purpose of

explicating the model. However, this is not to say that surface syntactic descriptions have no place at all in the model of thematization adopted in this study. After all, it is words that carry the information and the communicative value assigned to the theme and these words combine in a fixed syntactic form. Rather, it is just that these grammatical notions are not quite suitable for explaining the progression of information from sentence to sentence in discourse because theme is a functional notion associated with information communicated in discourse. It is in view of this that Danes (1974: 112) exhorts that:

We must not be content with a statement that certain sentence elements convey the known information (in contrast to others conveying the new one), but we ought to find out the principles exactly according to which this information and not another portion of the mass of known information has been selected. In other words, we must enquire into the principles of thematic choice and thematic progression.

Before proceeding to account for the principles exactly according to which theme was identified in utterances in this study, first "what is Danes' conception of Given information?" According to Danes (1974: 109):

- 1 Given or known is that information which is derivable or recoverable from the context, situation or common knowledge of the speaker.
- 2 The communicative structure of "givenness" assigned to particular sentence elements is a graded property.
- 3 "Givenness" depends on the portion of preceding text in relation to which the evaluation is carried out.

Conceding, on the basis of the above, that the contextual determination of theme is not a simple phenomenon, Danes suggests the following as some of the principles on which the identification of utterance theme may be based:

- (1) Identical wording
- (2) Synonymous expression
- (3) Paraphrase
- (4) Semantic inference

These four principles constitute, in the main, the means by which utterance theme is identified in this study. Each principle is explained with an accompanying example from the data.

1 Identical Wording

This refers to the repetition of identical lexical items, groups of lexical items or phrases from the preceding utterance(s) or context in an utterance at a particular point in text, thus serving to link the utterance at that point in text with the preceding one or with the context and situation.

Example:

We studied the effect of adoptive immunotherapy with lymphokine-activated killer (LAK) cells plus interleukin-2 or therapy with high-dose interleukin-2 alone in 153 patients with metastatic cancer for whom standard therapy had proved ineffective or no standard effective treatment was available. One hundred and eight patients were treated with 127 courses of LAK cells plus interleukin-2, and 49 patients were treated with courses of high dose interleukin-2 alone. (ABS 14:)

2 Synonymous Expression

An expression identified as theme is said to be synonymous with another in the preceding context or utterance if the information it communicates retains the meaning of the similar expression in the preceding context, albeit in different wordings:

Example

The neuronal degeneration is associated with various filamentous accumulations including actin (Hirano bodies), tubulin (granulovacuolar degeneration) and neurofilament

(neurofilarillary temples). The loss of retinal ganglion cells is of particular interest in that it is not accompanied by any filamentous accumulation in the retina. (RAP 9)

McCance et al found HPV-16 DNA in 92 percent of malignant cervical lesions. In their study only 18 percent of the colposcopically and histologically normal cervixes that were used as controls contain the HPV-16 genome. (RAP 12)

3 Semantic Inference

The theme of an utterance may be said to be related to information contained in a preceding utterance or context if the information which it communicates does not relate to information expressed in the previous utterance(s) by exact or slightly different wordings, but by direct or indirect association or by contrast.

Example

The researchers also found no change in breast cancer in women who use high progestogen "potency" OC's before the age of 25 and went on with them for more than six years. The earlier study had also made this suggestion which led our committee on safety of medicine two years ago to issue a list of pills and their potency, with advice to doctors to prescribe only low progestogen OC. (JRV 5)

4 Paraphrase

This refers to a situation in which the theme of an utterance communicates a part or all of the information contained in a previous or preceding context either by means of a single lexical item or group of words without losing the meaning content in the previous or preceding utterance.

Example

In a paired sequential double-blind trial of immunological treatment of recurrent spontaneous abortion, successful outcome of the next pregnancy was significantly more common in women injected with purified lymphocytes prepared from their husbands' blood than in those injected with their own lymphocytes. 17 of the 22 women given paternal cells had successful pregnancies, compared with 10 of 27 given their own cells.

5.5.1 Types of Utterance Themes in Texts

On the basis of the principles identified above and in keeping with Danes (1974), the themes identified in the texts which constitute the corpus for the analysis of TP fall into the following categories:

- 1 Simple themes
- 2 Multiple themes
- 3 Complex themes
- 4 Themes in exceptional utterances

Simple themes are found in simple utterances where the element that functions as theme is the first constituent of the utterance, eg.

- (a) Congestive heart failure is a common condition reported to affect 1 percent of the population (RAP.15).
- (b) Enalapril is an angiotension converting enzyme inhibitor (ACE 1). (JRV.15)
- (c) This paper presents the data on survival among the 253 randomized patients (RAP.15).

Multiple themes are those formed by the combination of two or more simple themes into a single thematic frame. In most cases the two or more simple themes are linked together by a conjunction. It is analogous to a "compound subject" in a declarative clause or sentence (see Danes, 1974: 117, example 16). Below are equivalent examples from the data in this study:

- (a) The purpose and nature of the trial was explained to all couples fulfilling the above criteria (RAP.1).
- (b) Memory deficits, aphasias and apraxias are the most common signs of Alzheimer's disease ... (RAP.9).

- (c) The name, address and date of birth of each man were entered on a computer file and the names were randomised into blocks of 500 (RAP.12).

Complex themes are those formed by the fusion of two or more subordinate elements to the main thematic element of the utterance. That main thematic element can function as a simple theme in a simple utterance. In a condensed utterance, the complex theme constitutes the common FP-element on which thematic status is assigned, (see Danes, 1974: 117, example 2a). Equivalent examples from the data are:

- (a) The estimated proportionate increase in risk of AP, independent of other CHD risk factors, was 1.2 (1.1-1.3) for 1% decrease in linoleic acid or a 0.1% decrease in eicosapentaenoic acid (ABS.12).
- (b) Many population studies, particularly the dietary data of the seven countries study, have answered the first part of the question, (RAP.12).
- (c) Ischaemic heart disease - the narrowing and blocking of the arteries that supply blood to the muscle of the heart - is the biggest single cause of death in the Western world in the 20th century. (JRV.5).

Also included in the category "complex themes" are themes which are formed by the fusion of such grammatical devices as adverbials, discourse adjuncts, relative pronouns and their clauses, nominalized expressions and conditional clauses to an NP to form a complex noun phrase, eg.

- (a) For each five year age group, the relationship between serum cholesterol and coronary heart disease (CHD) death rate was continuous, graded, and strong.
- (b) In age-specific analyses for the men aged 35 to 39, 40 to 44, 45 to 49, 50 to 54 and 55 to 57 at screening and in age-standardized analysis with stratification based on cigarette use and DBP, risk of CHD rose steadily for men in the

second, third, fourth and fifth quintiles, compared with the first quintile (RAP.11).

- (c) Using cultures of human ectocervical epithelium at 2 weeks post-explanation - an in-vitro model system in which small numbers of cells seem experimentally infectable with EBV - we recently identified by electron microscopy ... (RAP.4).
- (d) If their six-year age-standardized CHD mortality rate of only 1.6 per 1000 had prevailed for the whole cohort, the estimate is that CHD deaths would have numbered 560 instead of the total 2258 (RAP.11).
- (e) Last week, however, James Mowbray and colleagues at St Mary's Hospital in London described in The Lancet a treatment which has enabled 80% of the ... (JRV.1).
- (f) When a woman has her first baby she runs a much greater risk of toxæmia than in subsequent pregnancies. (JRV.3).
- (g) The scientists whose work is published in the latest issue of the "New England Journal of Medicine" believe too much emphasis has been placed on the fashionable virus theory to account for ... (JRV.10).

In all cases where these grammatical devices were positioned as initial elements in the utterance, as in **a-f** above, they were assigned thematic status. This is in accordance with the observation by Dubois (1987: 99) with conjunctive adverbials that their position as clause initial elements is "evidence of deliberate, if intuitive choice" which needs to be acknowledged by according to the expressions the status of theme. However, whether the NP's they precede are accorded thematic status as well depends, in this study, on whether they are "given" information or information recoverable from preceding context. In the case of relative clauses, they too are accorded thematic status, as in example "g" above, together with the NP preceding them.

5.5.2 The Identification of Theme in Some "Exceptional" Utterances

Considering the fact that our model of analysis is aimed at characterizing the progression of theme above the sentence level, it is necessary that we explain the principles underlying the identification of theme in certain constructions which seemed to disrupt the progression of themes due to their fixed word order. Utterances in this category are Existential Constructions, Extrapositions and Interrogatives. They are referred to as "exceptional" utterances in this study for want of a better term. Otherwise, they form part of what Dubois (1987) refers to as "marked" themes. However, since this term does not refer to the three utterance types identified above, it was not considered appropriate in this context. Besides, terms such as "marked theme" and "topic" have fixed meanings in other models of thematization (Halliday 1967; Givon 1983; Chafe 1976) which may not apply equally in an FSP model. Therefore, in order to account adequately for the progression of themes in contexts where these "exceptional" utterances occur, we would need to clarify the factors which constrain the organization of information they contain in accordance with the principles of FSP, with emphasis placed on the verbal and situational contexts in which the utterances occur.

5.5.2.1 Existential and Extraposition Utterance Themes

The identification of theme in existential and extraposition utterances deserve special mention because of the tendency they have for causing a break in the normal development of thematic progression in text. In traditional grammar, the existential there - construction is considered a "dummy-subject", a "place-minder" and a "slot-filler" for the more semantically significant would-be subject now in post-verbal position (Huddleston, 1984, Hudson, 1984, Quirk et al, 1972). Both Huddleston and Quirk et al also contend that the existential there-construction is a device for introducing new information into the universe of discourse. In particular, Quirk et al contend that the existential there-construction is a means of correcting the awkwardness which

would otherwise arise if new information occurred as the "initial element or theme of clause" which "typically contains given information", especially in sentences where the subject is an indefinite noun phrase. They point out that the awkwardness may be avoided by the introduction of "there" and the consequent postponement of the "notional subject" to a later, non-thematic position.

The account of the there-construction given by the FSP model is similar to those discussed above. Mathesius (Vachek 1975: 105) identified instances when the existential there-construction may be used in English in accordance with its FSP as:

- 1 When the subject is unknown or if it need not be expressed.
- 2 When the entire action including the subject belongs to rheme and the existential construction makes the expression of the entire action in the final section of the sentence possible in accordance with the Functional Sentence Perspective.

In a related development, Firbas (1975:) points out that the there-construction functions to express the "scene" while the phenomenon appearing or existing on the scene is conveyed by the (notional) subject. He explains that in FSP, "it is the context independent subject, expressing the phenomenon existing/appearing on the scene that under the circumstances carries the highest CD". In other words, the item "there" which serves to express the scene carries the lowest CD.

According to Firbas (1981: 48) an element is context-independent if in accordance with the immediate orientation of the communication it is non-derivable from the immediately relevant context. It is in this regard that Firbas (op cit) contends that a context-independent element is said to convey unknown, new information. Therefore in accordance with Firbas, the particle "there" is identified as theme in a there-construction in this study, as in the example below:

Since 1968, notification of meningococcal infection in England and Wales has been restricted to meningitis. There seems to be no logic behind the decision to exclude septicaemia, which only serves to confuse attempts at epidemiological investigation.

(ABS.7)

As with existential constructions, Extraposition involves shifting the subject from its normal position and placing it towards the end of the sentence in order to achieve end-focus or end-weight. According to Quirk et al (1972: 965).

"... the clausal subject is placed at the end of the sentence and the subject position filled by the anticipatory pronoun "it". The resulting sentence thus contains two subjects, which we may identify as the POSTPONED SUBJECT (the clause which is notionally the subject of the sentence) and the ANTICIPATORY (it)".

It appears that extraposition has not received any significant mention by the prominent FSP practitioners. However, if the characterization of "there-construction" by Firbas is extended to extrapositions, the pronoun "it" may be characterized as serving to express the scene while the phenomenon expressing existing/appearing on the scene is the context independent subject. In Firbas characterization, the former is low in CD while the latter is highest in CD. Consequently, in this study thematic status is conferred on the pronoun "it" in extraposition utterances. The rest of the utterance is analysed on the rheme. For example:

- (a) It is suggested that poor nutrition in early life increases susceptibility to the effects of an early diet (ABS.5).
- (b) It is a paradox that although the steep increase in ischaemic heart disease is now more common in poorer areas and lower income groups (RAP.5).
- (c) It seems likely, says Young, that the EB virus infects the squamous epithelium in the nose and throat of most - perhaps all - people (JRV.4).

5.5.2.2 Interrogative Utterance Themes

A very elaborate treatment of the FSP of English and Slavonic interrogative sentences has been made by Firbas (1976). In the paper, Firbas attempts to reconcile the views of Mathesius (1941, 1942) and Danes (1949) on the subject of FSP of interrogative utterances. He also proposes an approach which accommodates aspects of both views.

According to Firbas (op cit), Mathesius expressed the view that "the rheme of a pronominal question (a non-polar interrogative) is its initial interrogative word, ie. in English the Wh-element". Consequently, the rest of the question is the theme. Mathesius contends that the interrogative word stands for the unknown element which is to be disclosed as the rheme of the reply. He also gives the same interpretation to Yes/No questions, arguing that the finite verb in a Yes/No question expresses the rheme while the rest of the question is theme.

On the other hand, Danes (op cit) argues that the interrogative word in non-polar questions and the finite verb of Yes/No questions are not necessarily rhematic. He contends that in both types of questions elements other than those identified by Mathesius may carry the label "rheme".

In an attempt to reconcile both views, Firbas (op cit) contrasts the relation between the speaker and the listener in a question with that between speaker and listener in a declarative sentence. He points out that in the declarative sentence the speaker is in possession of some knowledge and is imparting it to the listener while in an interrogative sentence, "the speaker's primary concern is to obtain some knowledge from the listener". Thus, Firbas points out that a question performs two functions:

- (i) "It indicates the desire for knowledge on the part of the enquirer and appeals to the informant to satisfy the desire".

- (ii) "It imparts knowledge to the prospective informant in that it informs him of what the enquirer is interested in (what he is thinking of at the moment) and of the particular angle from which he wishes the intimated lack of knowledge to be satisfied".

According to Firbas, Mathesius seems to be overemphasising the first function while Danes has made the understanding of the second function possible. Instead he contends that the temporal and modal elements (TMES) in polar questions and the Wh-elements in non-polar questions "cannot be interpreted as rhematic in regard to the first function because it is a modal function which in unmarked usage "has to be looked upon as a concomitant phenomenon". With regard to the second function, he argues that the TMEs and Wh-element do not become rheme proper in unmarked usage, for despite the fact that they serve to indicate the angle from which the question is to be approached, "they do not ultimately specify this angle". Therefore, he contends that they can best be interpreted as transitional, "admitting that the Wh-element may come near or perhaps even occur in the periphery of the rheme". In this way, Firbas brings the issue of interrogative sentences within his theory of CD.

The interrogative utterances identified in this study do not seem to relate to the first function, but they relate partially to the second function. Since this study is not based on the theory of CD, the interpretation provided by Firbas for the second function, though revealing, will be disregarded in the analysis of data. Instead of designating the Wh-element in non-polar questions and the finite verb element in the polar questions as transitional elements, they are designated as rhematic in this study.

5.6 CROSS-GENRE COMPARISON

Having accounted for the basis upon which various categories of theme and TP patterns were identified in the corpus, the question that may now be asked is "how do

the genres compare with each other with regard to the way they organize information into identifiable TP patterns?" This question is addressed along three main dimensions:

- 1 The realization of TP patterns across genres
- 2 The relationship between TP patterns and text structure
- 3 The characteristic elements realizing theme across genres

5.6.1 The Realization of TP Patterns Across Genres

An obvious, but interesting result of this cross-genre comparison of TP is that all three genres make use of the Simple Linear and the Constant TP patterns in the organization and development of information. Below are brief instances of the use of these two TP patterns in each of the three genres.

THE SIMPLE LINEAR TP

The RAP:

(1) A growing body of data shows that blood transfusions have immunological effects on patients and experimental animals beyond those of alloimmunisation to blood cell antigens. (2) The best characterised clinical consequence is that of improved renal allograft survival in transfused patients compared with that in those not receiving homologous blood. (3) These findings have led to studies on whether transfusion at the time of cancer surgery might be associated with immunological modulation of the patient and unfavourable rates of tumour recurrence.

(RAP.6: *British Medical Journal*, 30 August 1986)

In the text above, the theme of the second utterance - "the best characterized consequence" - refers to information contained in the rheme of the preceding utterance, that is, that blood transfusions have immunological effects on patients. The link between both segments is provided by the relations between the lexical items "effects" and "consequence". Similarly, the theme of the third utterance, "these findings", relates to information in the rheme of the preceding utterance, the relation existing between both segments of text being that of paraphrase.

The ABS

(1) There was no significant increase or decrease in the risk of breast cancer for OC users according to age at first use or subsequent duration of use, even for women who first used OCs before the age of 20 yrs and who continued to use them for more than 4 years. (2) Risk was not altered significantly for women who used OCs with high progestogen "potency" before the age of 25 years (even when duration of use exceeds 6 yrs), for those who used OCs before first-term pregnancy (even when duration of use exceeded 4 years), or for OC users according to age at diagnosis. (3) These results suggest that the use of OCs by young women in the United States has no effect on the aggregate risk of breast cancer before 45 years of age.

(ABS.2: *The Lancet*, No 2, 1985)

In the text above, the theme of the second sentence - the lexical item "Risk" - refers to the expression "risk of breast cancer" in the rheme of the first utterance. The association between the rheme of the previous utterance and the theme of the second utterance is that of "identical wording". In the same way, the theme of the third utterance - "these results" relates to information contained in the whole of utterance 2, but specifically to information contained in the rheme of that utterance. Therefore, the theme of utterance three can be said to paraphrase the information contained in the rheme of utterance two.

The JRV

(1) Each woman's partner was asked to give a unit of blood, from which the lymphocytes were removed and washed. (2) These were used to immunise the woman by simple injection, and some weeks later blood samples were taken to find out if the woman had made antibodies to her partner's lymphocytes. (3) The trial of this procedure took account of the known tendency for recurrent miscarriage to improve slightly during any treatment.

(JRV.1: *New Scientist*, 9 May 1985).

In the text above, the theme of the second utterance, the demonstrative pronoun "these" refers to the information "a unit of blood, from which the lymphocytes were removed and washed" in the rheme of the first sentence. Similarly, the theme of the third utterance "the trial of this procedure" takes from information in the rheme of

utterance 2 which refers to the procedure by which women in the trial were immunised. In both cases, the association between theme and rheme is that of paraphrase.

TP WITH A CONTINUOUS OR CONSTANT THEME

The RAP:

(1) Selected characteristics of children, listed according to history of myocardial infarction in their parents are shown in Table 1. (2) Children whose fathers reported a myocardial infarction were most likely to be white, to smoke cigarettes, to be older, and to be obese, than were children whose fathers did not report a myocardial infarction. (3) In contrast, although children whose mothers reported a myocardial infarction tended to be older, no statistically significant differences relating to the disease in the mothers was observed.

(RAP.8: *The NEJM*, 18 Sept 1986)

In the text above, the themes of utterances two and three contain references to either children or myocardial infarction in their parents (fathers or mothers) both of which constitute the main elements in the theme of the first utterance. The association in the progression of theme from one utterance to the other is established mainly by means of identical wordings.

The ABS

(1) Data from patients with colonic, rectal, cervical and prostate tumours showed an association between transfusion of any amount of whole blood or large amounts of red blood cells at the time of surgery and later recurrence of cancer. (2) Recipients of one unit of whole blood had a significantly higher incidence of recurrence (45%) than recipients of a single unit of red cells (12%) ($P=0.03$). (3) Recipients of two units of whole blood also had a higher rate of recurrence (52%) than those receiving two units of red cells (23%) ($P=0.03$). (4) Recipients of any amount of whole blood had similar recurrence rates (38.52%).

(ABS.6: *British Medical Journal*, 30 August 1986)

In the text above, the themes of utterances two, three and four relate to the theme of the first utterance by semantic inference. The word "recipient" which is the subject of the utterances and the main thematic element in the utterances refer to an

equivalent item "patients" in the theme of utterance one. Thus, the TP is such that the themes of all four utterances in the text refer to the same entity.

The JRV:

- (1) The Epstein-Barr (EB) virus is one of the five herpes viruses that attack humans.
- (2) In Europe and the US infection can produce glandular fever (mononucleosis).
- (2) In Central Africa, it can lead to Burkitts Lymphoma, a facial cancer in children.
- (4) In China and Southeast Asia, EB virus plays a part in triggering the development of nasopharyngeal carcinoma (npc).
- (5) In npc, however, the virus infects only the lymphocytes but the epithelial cells at the back of the nose and throat.

(JRV.4: *The Lancet*, 1 February 1986)

In the text above, two elements - an adverbial and a noun - are designated "theme" in each of utterances 2, 3, 4 and 5. In utterance 5, though, the item "however" is conjunctive but not an adverbial. The noun is designated theme along with the adverbial element in each of those utterances because it refers to the item which is the topic of discourse in the entire text; that is, "the Epstein-Barr virus" either by synonymous expression or by inference. This, and the fact that the adverbial elements "In Europe and the US", "In Central Africa", "In China and Southeast Asia" function to set the scene for the more thematic nominal element make the progression of theme in the utterances constant or continuous.

The fact that the simple linear and constant TP patterns are features of all three genres in this study is one indication of the common grounds which the genres share in terms of discourse organization at the level of the progression of themes which is probably due to the relationship they share in subject-matter. It is also an indication that the two TP patterns are fundamental to discourse in the three genres. It may well be that they are equally fundamental to discourse in all genres, for Danes (1974) refers to them as basic patterns and Dubois (1987) refers to them as canonical types, from which other patterns are derived.

Another interesting and significant result obtained in the analysis is the observation that the **TP with a derived theme** was identified only in the RAP. No incidence of the organization of information by means of the derived TP pattern was observed in either the ABS or the JRV. And even in the RAP where it did occur, its use appears to be highly restricted as only two instances of its use were identified in the RAP texts analysed. The two instances were identified in the "Discussion Sections" of RAP.4 and RAP.7. Below is one instance of the Derived TP pattern identified in the RAP.

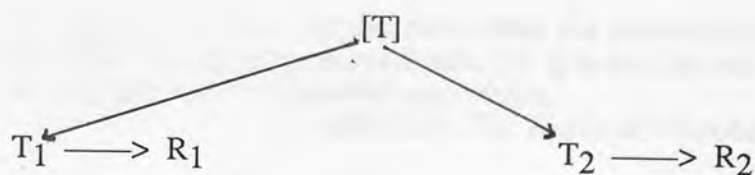
(1) The consistent presence of EBV genomes in the malignant cell of NPC, coupled to more recent evidence indicating a similar association of the virus with salivary gland, and tymphoepithelimas, has been difficult to reconcile with the accepted view of EBV as a strictly lymphotropic agent. (2) In this context, it has been argued that epithelial cells are not natural targets for EBV infection but that the virus may gain indirect access to such cells either through their fusion with a virus infected B cell or through their fusion with a virus infected B cell or through the formation of acquired envelopes from an unrelated epitheliotropic virus. (3) There are now firm grounds for believing that EBV infect epithelial cells directly. (4) Firstly, the evidence from work on the ectocervix model system strongly suggests that HBS and anti-B straining epithelial cells does indeed indicate their surface expression of a functional EBV receptor (Sixbey, JW et al, unpublished). (4) Secondly, as reported here, these two MABs, which are specific for quite distinct epitopes on the C3d/EBV receptor molecule of B cells, both show reactivity with cells throughout the epithelium lining the mouth and pharynx.

(RAP.4: *The Lancet*, 1 February 1986)

According to Danes (1974: 120) in TPs with a Derived Theme, "the particular utterance themes are derived from a hypertheme (of a paragraph, or other text section). In the context of the text above, the hypertheme is the heading to the section of the RAP in which it occurs, which is the "Discussion section". In the light of this, the first two utterances in that text may be regarded as generalizations which provide background to the main discourse in that section. The discussion of the main discourse of that section begins with utterance three which in some circumstances may be termed the "topic sentence" of the paragraph. Thus, it can be said to be derived much more

directly from the hypertheme of the text (Discussion) than either utterance one or two. Equally significant is the fact that utterance 3 is a there-construction. It has been observed elsewhere in this chapter that in a there-construction, the "there" particle is assigned the status of theme because as Firbas (1975, 1981) points out, it functions to express the scene and, therefore, carries the lowest CD. It is this scene or perspective that links the utterance with the preceding utterance and that provides the basis for processing other utterances following after it. Thus, in the context of the text above, the "there" particle is not only theme for the utterance in which it occurs, but it is also a derived theme, for it expresses the scene for the commencement of discourse relevant to the hypertheme of the text as well as providing a perspective from which subsequent discourse in the paragraph is to develop. In line with this contention, utterances four and five are developments from that perspective. The sequence of development from that perspective is signalled by the conjunctive adverbials "firstly" in utterance four and "secondly" in utterance five which combine with their respective NPs to form the theme of their respective utterances. Therefore, utterances 3, 4 and 5 have the TP pattern shown below:

Figure 5: Realization of Derived TP in the RAP



The final result obtained on the realization of TP patterns relates to the observation that there is a tendency for the **split-rheme** TP pattern to occur only in the RAP and the JRV. No indications of the use of split-rheme TP pattern was identified in ABS texts. Examples of the split-rheme TP as realized in the RAP and the JRV are provided below:

(1) Replies were divided into three groups. (2) Group 1 consisted of men with no reported history of CHD who were negative to the WHO chest pain questionnaire; (3) group 2 were men with no reported history of CHD who were positive to the WHO chest pain questionnaire; (4) and group 3 men with a reported history of CHD, either MI or AP.

(RAP.12: *The Lancet*, 24 January 1987)

The text above is taken from the methods section of text RAP.12. It is part of the description of the experimental procedure adopted in the research investigation. Utterance one is made up of a simple theme with a subject NP as theme and a rheme which quantifies the number of groups into which information contained in the theme is divided. Utterances 2, 3 and 4 are descriptions of the specification provided in the rheme of utterance one. By first specifying the number of groups into which replies are divided in the rheme of utterance one, the writer signals his intention to provide a detailed description of those groups in subsequent utterances. Since the information cannot be said to be complete until the three groups are fully described, utterances 2, 3 and 4 which provide the full descriptions are considered extensions of the rheme of utterance one. The themes of all three utterances make reference to the lexical item "three groups" in utterance one, with each theme representing an element of the group. Thus, the TP pattern is one of an exposition of a split-rheme.

(1) They drew three implications from the observations about the link; (2) the virus plays a part; (3) it is not the sole cause; (4) patients may need careful supervision.

(JRV.10: *The Times*, 23 October 1986)

In the text above, the theme of utterance 1, is the personal pronoun "they". The rest of the utterance is rheme. In the rheme, there is a specification of the number of implications which the researchers drew from their observations. These implications are further specified in utterances 2, 3 and 4. Thus, the information contained in utterances 2, 3 and 4 are extensions of the specification made in the rheme of utterance 1. In other words, the rheme of utterance 1 is split into three separate utterances,

giving a progression in which three separate utterances originate from a single rheme to form what is known as the "exposition of a split rheme".

To sum up, the observations reported so far show that in terms of the realization of TP patterns, only the RAP can be said to realize all four of Danes patterns; the JRV realizes three while the ABS is capable of realizing only two. This is not an altogether surprising result. It would seem to reflect the constraints imposed on texts in the different genres in terms of length and the scope for information development. In view of this, it is considered perfectly normal that the RAP with the greatest scope for the development of information should make use of all four patterns of TP, while the ABS on the other hand, with the least scope for the development of information should limit itself to the two basic TP patterns - the simple linear and the constant TP patterns. The JRV, whose scope for information development can be described as being mid-way between that in the ABS and the RAP gives an indication of the edge which it has over the former and its limitations in terms of information content against the latter by realizing three out of the four TP types.

5.6.2 The Distribution of TP Patterns

The observation presented in the previous sub-section reports on the TP patterns identified in the analysis of texts from the three genres. It does not purport to provide answers to issues related to the distribution of these TP patterns, nor does it account sufficiently well for the factors which influence the realization and ultimate distribution of TP patterns identified. Furthermore, it does not explain these issues in the context of the relationship which exists between TP patterns and text structure. An attempt is made in this sub-section to address these issues through an account of the distribution of TP patterns across the three genres. As usual, the approach adopted in presenting the observations made with regards to the distribution of TP patterns is qualitative rather than quantitative. Thematic Progression patterns are functional categories which combine in a number of complex ways in natural discourse. Problems are likely to arise over the number of progressions that one can regard as

constituting a unit for quantification, particularly with texts like the ABS and the JRV which tend to condense information and are likely to have very short progressions compared to the RAP which is likely to have much longer progressions. Therefore, the result on the distribution of TP patterns in this study is based on the observations made regarding common relationships which exist between TP patterns and text structure in texts across the three genres.

The analysis of the distribution of TP patterns shows that, though the occurrence of the simple linear and constant TP patterns is common to all three genres, the extent to which each genre employs either of the patterns to organize and develop information in its texts varies from one genre to another. Thus, it was observed that the primary method of utterance organization and development used in the JRV is the simple linear TP pattern. In contrast, the RAP tends to organize and develop information in utterances mainly by means of the constant or continuous TP pattern. The ABS, on the other hand, tends to make use of both patterns fairly evenly. As has already been pointed out elsewhere in this section, (5.5.1) very few instances of the Derived TP and the split-rheme TP patterns were identified. Therefore, their distribution in the genres is low relative to the simple linear and the constant TP patterns.

The distribution of these TP patterns across the three genres may be accounted for in two main ways. First, by reference to the relationship between TP and text structure. Second, by reference to the influence of social and contextual factors. With regards to the former, it has been observed (Fries, 1981: 6) that thematic progression correlates with the structure of text. By this Fries means that in a highly structured passage the themes of each utterance within the passage should reflect the structure of that passage. For instance, in argumentative and expository prose each utterance should follow logically from what has gone before. In accordance with Fries' contention, it is reasonable to assume that in parallel texts, equivalent units of information will tend to organize and develop information in their utterances by means

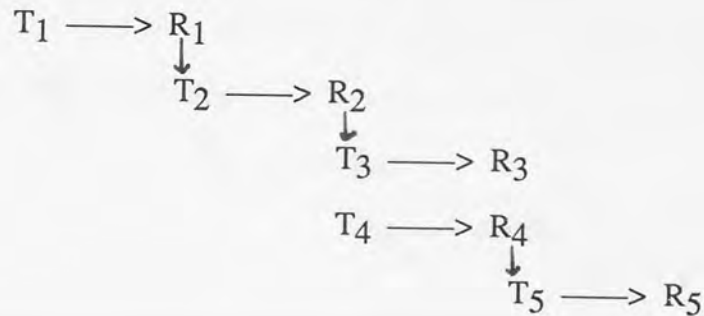
of identical TP patterns. Following this assumption, units of information identified in the analysis of schematic structures (Chapter 4) were examined across the three genres with a view to ascertaining the characteristic TP patterns they tend to use to organize and develop information in their utterances. The analysis revealed that there is no one-to-one correspondence between schematic units and TP patterns. That is, no schematic unit of text can be said to organize and develop information in utterances by means of only one TP pattern. In other words, information in schematic units of text tend to be organized and developed by more than one pattern of TP. However, despite the fact that there is no one-to-one correspondence in the relationship between TP patterns and schematic units, it was also observed that schematic units showed preference for one TP pattern over others. In line with this latter observation, schematic units were observed to belong to one of two groups - those which contain mainly expository or explanatory information and those which are made up mainly of descriptive and procedural information. In the former group are all schematic units in the Discussion section and one schematic unit in the Introduction section of the RAP and the equivalents in the JRV and the ABS. In the latter group are schematic units in the Methods and Results sections of the RAP and two units from the Introduction section as well as their equivalents from the JRV and the ABS.

On the basis of the above characterization it was observed that in all three genres the use of the simple linear TP pattern tends to predominate in schematic units involving explanation and exposition, as in the examples below:

(1) To our knowledge, tissue and plasma kallikrein have not been previously identified in vivo from human lung. (2) In vitro studies on human basophils (39,40) and lung tissue (41) have detected the presence of kininogenase(s) after anaphylactic challenge. (3) The kininogenase activity(s) in each case is distinguished from the BAL fluid activity by high molecular weight. (4) Similarly, although the IgE-mediated release of a kininogenase from human lung mast cells has been reported, this activity can also be distinguished from the tissue kallikrein in BAL fluids on the basis of both molecular weight (145,000) and the fact that the mast cell kininogenase generated kinin optimally at pH 5.5 (42). (5) The properties of this mast cell kininogenase

have led to the suggestion that it may be identical to tryptase, the major neutral protease from human lung mast cells.
(RAP.13: *J of Clin Invest*, January 1987)

Figure 6: Realization of Simple Linear TP Pattern in RAP



The text above has the progression pattern shown in Fig. 6. It is taken from the Discussion section of RAP.13. The lexical item "lung tissue" in the theme of utterance 2 relates to an equivalent item "human lung" in the rheme of utterance 1. Similarly, the item "kininogenase activity" in the theme of utterance 3 relates directly to the lexical item "kininogenase(s)" in the rheme of utterance 2. In contrast the lexical item "this activity" in the theme of utterance 4 refers to "kininogenase activity(s) in the theme of utterance 3, thus causing a shift in the TP. However, the simple linear TP is picked up again, with the lexical item "this mast cell" in the theme of utterance 5 referring to the phrase "mast cell kininogenase" in the rheme of utterance 4.

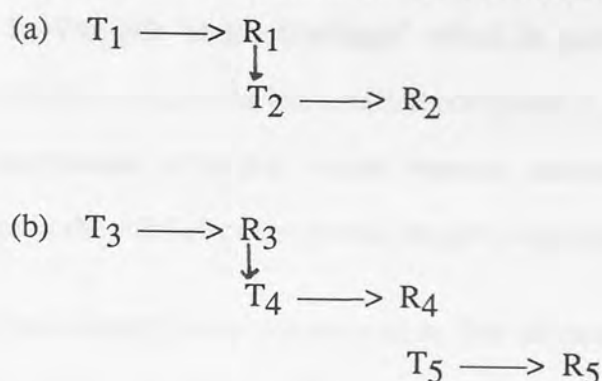
(1) Millions of people in the world suffer from the illness which in some unknown way affects brain cells. (2) Symptoms include memory loss, disorientation and changes in personality.

(3) But if the unusual type of nerve damage producing specific visual disturbances is a characteristic only of Alzheimer's victims, it could give doctors a way to separate the illness from other mental problems.

(4) That is important because even though there is no cure for Alzheimer's disease, it is sometimes confused with treatable diseases. (5) Accurate diagnosis may offer little hope at this point, even for those who are in early stages of the disease.

(JRV.9: *The Times*, 23 August 1986).

Figure 7: Realization of Simple Linear TP in the JRV



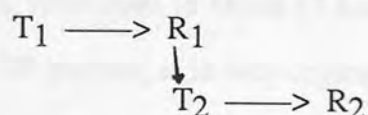
The text above has the progression pattern shown in Fig. 7. It represents a unit of text from the JRV, which attempts to provide background information to the reader about the nature of Alzheimer's disease. The progression of themes in the text follow the simple linear pattern with utterances 1 and 2 forming a single progression. The theme of utterance 2 "symptoms" refers to symptoms of "the illness" in the rheme of utterance 1. Utterance 3 begins a new progression. The theme of utterance 4, the demonstrative "that" refers to the entire information in the rheme of utterance 3. In the same way, the theme of utterance 5, "accurate diagnosis" relates thematically to the diagnosis of Alzheimer's disease and the need not to confuse it with treatable diseases which is stated in the rheme of utterance 4.

(1) Two monoclonal antibodies against different determinants on the EBV/C3d receptor of B cells were shown to react with pharyngeal epithelia in a cell differentiation-dependent manner.
 (2) These findings, which raise the possibility of direct virus entry into a naturally exposed epithelium, strengthen the evidence in favour of an epithelial reservoir of EBV infection in vivo and identify a means whereby the virus/epithelium interactions leading to nasopharyngeal carcinoma might be initiated.

(ABS.4: *The Lancet*, 1 February 1986)

The text above has a progression of:

Figure 8: Realization of Simple Linear TP in the ABS



As the TP pattern shows, the theme of utterance 2 is derived from the Rheme of utterance 1. The NP "**these findings**" which is part of the theme of utterance 2 refers, to a large extent, to the information contained in the rheme of utterance 1. The thematic progression in the text is short because, as observed in this study, the nature of discourse in the ABS does not permit longer progressions.

As has already been pointed out in this discussion, and as the sample texts above show, the simple linear TP pattern tends to be used extensively to develop information in schematic units in the Discussion section and part of the introduction section that attempts to present background to the study. This observation is consistent with that made by Weissberg (1984) to the effect that the simple linear TP occurs mainly in the Discussion and Introduction sections of experimental research papers. Having said this, the question that needs to be asked is, "How does this explain the higher tendency for utterances in JRV texts to be organized and developed according to the simple linear TP pattern?"

As the analysis of schematic structures (Chapter 4) shows, the JRV selects from the RAP only those schematic units of information which are directly relevant to the needs of its audience. Prominent amongst the schematic units which the JRV selects from the RAP is information contained in Moves 10 and 11 of the RAP, both of which may be found in the Discussion sections and which function to explain the results obtained in the study. It is equally worthy of note that in addition to selecting some **constituent elements** from these RAP moves, the JRV introduces new constituent elements which seek to explain principles and concepts underlying issues in the research being popularized. Therefore, the bulk of information in the JRV is explanatory or expository. This is further attested to by the fact that Move 11 is the most extensive of all the schematic units in texts in the genre. Therefore, if, as the analysis shows, utterances in Move 11 have a tendency to be organized following the simple linear TP pattern, it is only logical that in addition to its occurrence in other

Moves, the simple linear TP will tend to predominate over others in the organization and development of discourse in the JRV.

The high incidence of the simple linear TP in the JRV may also be explained as an information processing strategy which is consistent with the nature of the audience for which the JRV is written and the purpose of discourse in the genre. In this regard, a look at the examples given for the simple linear TP in this section as well as other instances across the three genres shows that utterances in a simple linear thematic progression tend to relate to each other mainly by means of "identical wordings" and by the use of "demonstratives", both of which make the identification of referents relatively easy. The ease with which referents can be identified is communicatively important for an audience, such as that in the JRV, whose knowledge of the subject-matter may not be very high. In the light of this contention, the predominant use of the simple linear TP pattern in the JRV can be said to represent a strategy employed by writers to enable readers without sufficient knowledge of the subject-matter to process specialist information with greater ease.

Related to this is the fact that the high incidence of the simple linear TP pattern in the JRV may be constrained by the level of shared knowledge which exists between the writer and the reader in texts in the genre which is reflected in the search by JRV writers for a viable "point of departure" for utterances in text. As Firbas (1981: 46) points out:

The speaker/writer and listener/reader's shared knowledge of the immediately relevant context forms part of their shared knowledge of the entire preceding verbal context and the entire relevant situation, all this knowledge in its turn forming part of their shared general knowledge and experience (cf. Firbas 1975a; 1979a, 31-2) ... Bound by the objectively existing context and the laws of interplay, the speaker/writer and the listener/reader can exchange, and in fact keep on exchanging their roles.

As has already been pointed out elsewhere in this study (Chapter 4), the JRV is written by someone other than the original writer of the research article being popularized. The writer may or may not be a scientist, and even if he is, he is not likely to be engaged in active research work. He is most likely to be a journalist with some knowledge of science.

The JRV audience likewise consists mainly of non-specialists; people whose knowledge of science is a little above 'A' level. They are most likely to be students in higher education or non-students with some interest in science. The nature of the writer and the audience of a JRV text suggests that the level of knowledge which they both have about the subject-matter is low. But as Dubois (1986) points out, the writer can resolve his doubts about issues in the subject-matter "by phoning one or more authors of the writing in question". This, plus the fact that he has the original text to work with adds significantly to his knowledge of the subject-matter. On the other hand, it reduces the level of shared knowledge which exists between him and the reader. This difference in shared knowledge tends to narrow down the range of options available to the writer for adopting a point of departure for the discourse in utterances in the text. In view of this, the writer is faced with the task of establishing common grounds, one that both he and the reader can share in the process of pushing the communication forward from one utterance to the other. In the context of FSP, the place to establish that common ground is in the theme, referred to by Mathesius (1982: 120) and Firbas (1987: 148) as the "foundation" for the utterance. In the face of lack of any meaningful shared knowledge between the writer and the reader on the subject-matter of discourse, the foundation for a subsequent utterance can only be established by selecting from the range of information contained in the rheme of the preceding context to which the reader and the writer can hold as shared knowledge. The simple linear TP pattern would seem to exist in the JRV to make such an arrangement feasible. On the other hand, it may also have direct influence on the extent to which the reader can retain information from the preceding context as he processes discourse in text.

From the point of view of FSP, the rheme carries the highest communicative value (Firbas 1964). It is defined by Mathesius (1939: 171) as "that which the speaker is communicating about". The combination of these two views points to the importance of the rheme in communicating information relating to what the writer is communicating about. Viewed in this context, the simple linear TP pattern in which information from the rheme is constantly made the theme of the subsequent utterance would seem to be a means of ensuring that information relating to what the writer is communicating about in the text is constantly made the foundation for subsequent utterances. From the point of view of the audience for which the JRV is written, this would seem to be a positive strategy; one that ensures that discourse for this class of audience is organized from the known to the unknown; that the reader's attention is constantly focussed on what the writer is communicating about. It is a strategy which may be capable of "drawing the reader in" (Myers 1986), thereby, sustaining his interest in the subject-matter of discourse and ensuring that he retains as much as possible information relating to what the writer is communicating about in the text. It is important that the JRV reader who is not a captive audience is hooked to the text in this way for he may decide to discontinue reading the text if the pattern in which information is organized makes it difficult for him to process the information being communicated.

A second significant observation made in this study with regards to the distribution of TP patterns is that the RAP tends to organize and develop information in utterances mainly by means of the constant TP pattern. This observation is also based partly on the relationship which exists between TP and text structure in RAP texts, and partly on the relationship between TP and authors assumptions of readers' knowledge.

The rationale for basing results on the distribution of TP patterns on the first relationship and the procedure adopted in assessing it has already been discussed. In that discussion, it was pointed out that schematic units of text tended to show preference for one TP pattern over others. Related to this was the observation that

schematic units which involve the description of a state, process, event or procedure tend to organize information mainly by means of the constant TP pattern. In the RAP, this involves schematic units in parts of the Introduction section and all of the Methods and Results sections as well as their equivalent forms in the JRV and the ABS. Specifically, they involve the following schematic units from the RAP:

- Move 2: Reviewing Related Research
- Move 3: Presenting New Research
- Move 4: Describing Data Collection Procedure
- Move 5: Describing Experimental Procedure
- Move 6: Describing Data Analysis Procedure
- Move 7: Indicating Consistent Observations
- Move 8: Indicating Non-consistent Observations

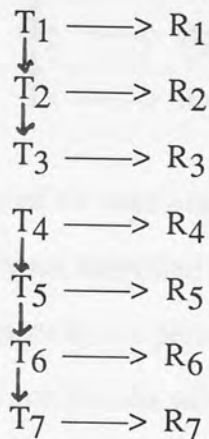
A few examples from texts across the three genres will serve to underline the discussion so far:

(1) The data collection, screening, and follow-up methods used in this research have been published. (2) In brief, from November 1973 to November 1975, 361 662 men aged 25 to 57 years were screened in 18 US cities at 22 MRFIT clinical centers. (3) Several methods of recruitment were used, including screening of employee, civic, and church groups, identification of men by house-to-house canvassing, and screening of respondents to mass media publicity. (4) Serum cholesterol levels were determined at one of 14 local laboratories using automated system of chemical analysis (Auto Analyzer II), with standardizations by the Lipid Standardization Program, Centers for Disease Control, Public Health Service, Atlanta. (5) Blood pressure was measured according to a standardized protocol by certified technicians, with first and ... (6) Three readings were taken with a standard stethoscope and mercury sphygmomanometer. (7) The average of the second and third readings was used in this report.

(RAP.11: *JAMA*, 28 November 1986)

The TP in the text above may be represented thus:

Figure 9: Realization of the Constant TP Pattern in the RAP



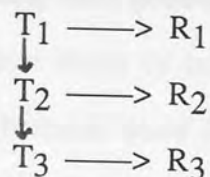
The text above is taken from the Methods section of RAP.11. It is a description of aspects of the procedure adopted in the research. Two distinct progressions are noticeable in the text. In the first progression which begins from utterance 1 and ends at utterance 3, the themes of all three utterances refer to the "Methods used in the research", with the theme of utterance 2 being related to the theme of utterance 1 by "Semantic inference". The second progression begins from utterance 4 and ends at utterance 7. The theme of all four utterances refer to methods adopted in determining "serum cholesterol levels", including "Blood pressure" measurements.

(1) In their study doctors examined the optic nerves which carry visual messages from the eyes deep into the brain. (2) They found specific damage to those nerves in Alzheimer's victims but not in healthy people. (3) However, they did not check people with other forms of brain disease.

(JRV.9: The Times, 23 August 1986)

The text above, has the progression:

Figure 10: Realization of the Constant TP Pattern in the JRV

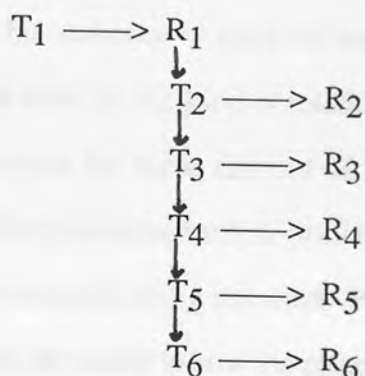


The theme of all three utterances in the text refer to the doctors who carried out the research. They are identified in the first utterance as "doctors", and in the theme of subsequent utterances by the personal pronoun "they". The text represents information from the Methods and Results section of its RAP equivalent.

(1) Between October 1981, and March 1986, there were 65 cases of meningococcal infection, about five times the expected number, in Gloucester Health District. (2) The cases, mainly in teenagers and young adults, were clustered in the Stroud district and in the Southern part of Gloucester city, and (3) most were caused by a sulphonamide-resistant group B type 15 meningococcus. (4) 2 patients died. (5) Only 57% of meningitis cases were formally notified. (6) 7 (11% patients had septicaemia without meningitis, not a notifiable disease. (ABS.7: *The Lancet*, 6 September 1986)

The TP for the text above is represented below.

Figure 11: Realization of the Constant TP Pattern in the ABS



The ABS text above contains information which is derived from the Results section of the RAP text which it accompanies. Apart from the theme of utterance 1, the themes of all the other utterances in the text refer to "cases of meningococcal infection in Gloucester Health District".

As has already been pointed out and duly illustrated by the examples above, the constant TP pattern tends to be used more frequently to organize and develop information in schematic units (see Chapter 4) in the Introduction, Methods and Results sections of an RAP text. It is also the predominant form of organization for utterances in schematic units in the JRV and ABS which may be regarded as being derived from the RAP sections referred to above. In quantitative terms, the constant TP pattern may be said to represent the predominant pattern for the organization of information in about 65% of schematic units in the RAP (see Table 2). This figure, which points to a high incidence of constant TP patterns in the RAP supports the observation that the distribution of the constant TP pattern is high in the RAP.

However, a more plausible explanation for the higher distribution of the constant TP pattern in the RAP might be found in the assumptions which RAP writers have concerning the knowledge they share with their audience on the subject-matter of discourse. Authors of RAP texts are specialists who have conducted investigations into medical problems which they wish to share with their colleagues in the medical profession. Therefore, their immediate audience is their colleagues in the same profession as themselves. Following from this, the relationship which exists between the writer and his audience in the RAP may be characterized as "specialist-specialist" relationship. In view of this kind of relationship, the writer is likely to assume that his audience possesses the same amount of knowledge as himself about aspects of the subject-matter of discourse, such as principles, concepts, procedures and tools that are relevant to the research being reported. Perhaps, what the audience may not share in great detail with the writer is how the principles, concepts, and experimental tools were applied by the researcher and what results were obtained in the process. This notwithstanding, the fact still remains that there is considerable amount of shared knowledge between the writer and his audience in an RAP situation. This is significant from an FSP point of view, for the knowledge which the writer shares with the reader determines the perspective from which information is presented in utterances. In the

context of FSP, the theme conveys known or shared information. That is, information which the speaker/writer considers well known to the hearer/reader and from which the rest of the discourse in the utterance proceeds. As has been pointed out, the writer and the reader of an RAP text have a considerable amount of shared knowledge in the areas of principles and concepts, experimental tools and procedures. They also share common knowledge about previous researches in the field. The question that needs to be asked is, "How do these account for the high incidence of the constant TP pattern in the RAP? A look at the themes of utterances which develop information by means of the constant TP pattern in the RAP shows that most of them are NP's which refer to names of objects, diseases, statistical tools, cited names, etc. In most cases, these get repeated as themes of subsequent utterances, thus leading to the realization of a constant TP pattern, as in the example below:

The name, address, and date of each man were entered on a computer file and the names were randomised into blocks of 500. A postal questionnaire was sent in batches of 500 every month for a year (April 1983 to April 1984). The questionnaire included the self-administered World Health Organisation chest pain questionnaire. If there was no reply to the first questionnaire within 2 weeks, a reminder letter and questionnaire were sent. The response rate to the postal questionnaire was 72%.

(RAP.12: *The Lancet*, 24 January 1987).

Therefore, following from the observations made in this discussion, it can be said that the high distribution of the constant TP pattern in the RAP is constrained by the assumptions which the writer has concerning the knowledge which he shares with his audience. Of course, the high occurrence of constant TP patterns in the RAP may also be constrained by the conceptual structure of the discipline which is reflected in the ordering of syntactic and vocabulary items in utterances. Thus, as Myers (1986) points out, professional research papers present a narrative which emphasizes in syntax and vocabulary the conceptual structure of the discipline. Therefore, in line with Myers (op cit) it can be contended that the high occurrence of the constant TP pattern in

the RAP is consistent with the conceptual structure of the discipline - a structure in which specimens, experimental tools and concepts are thematized.

Related to the level of shared knowledge between the writer and the reader is the need for ease of discourse processing. This has been discussed for the JRV in relation with the high distribution of the simple linear TP pattern. It appears that the same purpose is served by the high distribution of the constant TP pattern in the RAP. According to Danes (1970, 1974), the constant TP pattern presents a progression in which "one and the same theme appears in a series of utterances, though not necessarily in fully identical wordings". Danes (1970: 134) further points out that the theme contains "old, already known or given elements functioning as a starting point of the utterance". Related to Danes' characterization of theme is the contention by Chafe (1975: 211) that old information is the information which the speaker and hearer already share at the time the sentence is spoken, which constitutes a starting point based on concepts already in the air" to which the new information can be related. Following from these, and as has already been pointed out in the preceding paragraph, the choice of elements which may function as theme/old information would seem to be constrained by the level of knowledge which the writer shares with his reader about the subject-matter of discourse in the text. In view of this, it can be said that the manner in which the writer provides the reader with old information is a reflection of the writers' judgement of the readers' ability to process information in the discourse. In the context of the RAP, both the writer and the reader are specialists in the field of discourse. Kennedy and Bolitho (1984: 87) point out that "a scholarly paper is aimed at a specific readership of colleagues who will have a high degree of shared knowledge and assumptions about the particular topic". Therefore, in presenting information to conform with the knowledge he and the reader share about the topic, the writer does not need to constantly draw from the rheme of the preceding utterance as we saw in the case of the JRV. Rather, he is more likely to draw from the repertoire of knowledge which he shares with his audience. According to the principles of FSP, this shared

knowledge will always constitute the initial elements in the utterance and will provide the point of departure for the utterance as a message. If, as in the RAP, this element remains the same over a series of utterances, the ensuing progression will be of a constant TP pattern. The advantage of this to the RAP reader is that in a constant TP pattern in which there is the tendency for the same subject or topic to be thematized over long stretches of text, the task of information processing is made relatively easier because it ensures that the reader can devote greater attention to the information being communicated in the rheme. This contention is related to that made by Perfetti and Goldman (1974) to the effect that a thematized referent occurring as subject is a good aid to recall. In the RAP, the thematized elements which get repeated in subsequent utterances in a constant TP pattern (topic entity, specimens, materials, diseases etc) are features of old information or knowledge which the writer shares with his audience and which he brings to the fore in order to provide the utterance with a point of departure and the reader with material which can aid his recall of the new information contained in the rheme. It is on this basis that the contention that the organization and development of information in the RAP by means of the constant TP pattern is influenced by the audience and constrained by the need for ease of discourse processing is made in this study.

Finally, the RAP also differs from both the JRV and the ABS because it is the only genre which organizes information in utterances by means of the "TP with Derived Theme" pattern. In terms of the relationship which exists between TP and text structure, the TP with a Derived Theme tends to occur in paragraph structures in the RAP which involve itemization or serialization of events, sequences of events or the chronological presentation of events from a hypertheme as in narratives. To this end it is instructive that the only instances of the use of the TP with Derived Theme in the RAP corpus were identified in the Discussion section where it is possible to employ narrative and descriptive techniques in the process of exposition. This observation realates to that made by Holloway (1981: 208) to the effect that units of discourse that

are largely descriptive or narrative often repeat the overall theme of the passage (or some aspects of it) in the "old" section of each sentence. It is not clear what particular contextual factor constrains the distribution of the TP with Derived Theme. What appears obvious is that the extent to which it may be used to develop information in genres of the type examined in this study will depend on the extent to which the genres permit extensive discussions of subject-matter.

A significant observation made with regards to TP with Derived Themes is the fact that derived utterances may also embed within them a number of utterances with discernable TP patterns, as shown in the example below:

(1) Several features of the outbreak are of importance. (2) First, there is no evidence, after 4 years that the outbreak is waning. (3) In Norway, over the 11 year period from 1974 to 1984, the average annual incidence of meningococcal disease remained raised (74 cases/1000) with B.15R organisms predominating. (4) Recent localised outbreaks due to group B strains in England and Wales have lasted from 8 months to 3 years.

(5) Secondly, the age distribution of cases is unusual by comparison with recent national data (Fig 2). (6) Meningococcal disease usually affects children under the age of 5 years, a reflection of the low antibody levels in this group. (7) The high attack rate in teenagers and young adults by the B.15R strain of meningococcus has also been recorded in Norway but the absence of a peak in young children in the Gloucester outbreak is unexpected.

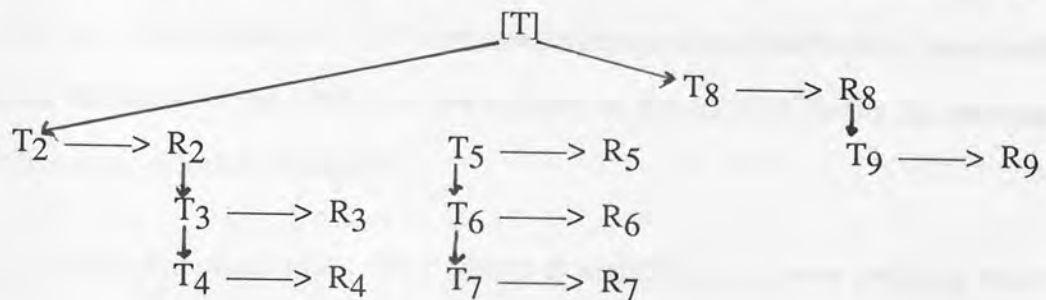
(8) Thirdly, throat swabbing revealed a very low carriage rate of B.15R meningococci in symptomless contacts. (9) Similar low carriage rates of this organism were found in an extensive swabbing investigation in Somerset after a minor outbreak of 1985-86 (Bowie C, Personal Communication).

(RAP.7: *The Lancet*, 6 September 1986)

In the example above, the theme of utterance 1 functions as the hypertheme from which the themes of utterance 2, 5 and 8 are derived. It is interesting to note that each of utterances 2, 5 and 8 precede a number of utterances which are integral parts of the derived TP complex. These intervening utterances are mainly exemplifications and justifications which develop the contentions made in those utterances which relate

directly to the hypertheme. Since the utterances do not cause a shift in argumentation or orientation of the discourse, and since they do not come at the end of the third derived T-R nexus, that is utterance 8, to form what may be regarded as a combination of the hypertheme and another TP pattern, they may best be characterised as embedded progressions within the Derived TP complex. This form of TP with Derived Theme is unusual considering that Danes' (1970, 1974) examples of the TP pattern did not make provisions for such irregular patterns. Consequently, it is referred to in this study as **Derived TP with embedding**. The pattern of progression for this form of Derived TP as in the text above is as follows:

Figure 12: Realization of the "Derived TP Pattern with Embedding"



The distribution of TP patterns in the ABS deserve some mention. Basically, there is nothing unique about the realization and distribution of TP patterns in the ABS. As has been pointed out, the ABS realizes only two TP patterns - the simple linear and the constant TP patterns. It was also observed that these two patterns are fairly evenly distributed in the organization of utterances in ABS texts. In addition to these, it may be added that the ABS is much closer to the RAP than the JRV in terms of the realization and distribution of TP pattern. Thus, schematic units which function to present previous or present research, present method of investigation and results tend to make more use of the constant TP pattern to organize information in utterances. On the other hand information related to discussion of results and conclusions tend to be organized following the simple linear TP pattern. However, as the analysis of schematic structures shows, much of the information in the ABS tends to be taken

from the Results and Discussion sections of the RAP. Very little information from the Introduction and Methods sections tend to be reported in ABS texts. The balance between the selection of information from the Results section and the Discussion section of the RAP may partly account for the balance in the distribution of the simple linear and constant TP patterns in the ABS.

5.6.4 The Characteristic Elements Which Realize Theme Across Genres

In the preceding discussion it was observed that there is a tendency for the constant TP pattern to be preferred as a means of organizing information in utterances in the RAP than in the JRV. This preference was partly attributed to the fact that the JRV is, as it were, a rewritten and condensed version of the RAP, hence, much of the original discourse of the RAP is either omitted or generalized during the process of rewriting and information transfer.

While the observation made above is significant, a more exciting result is provided in the variation which exists between the RAP and the ABS on the one hand and the JRV on the other in the elements which function to indicate the writer's perspective in those utterances and schematic units which are organized by means of the constant TP pattern.

As the analysis shows, RAP and ABS writers tend to indicate perspective, especially in those schematic units of text in which the constant TP is employed by means of syntactic elements which differ from those used by JRV writers.

Thus, in the RAP and to a large extent in the ABS, the theme of utterances in a constant TP pattern tend to consist more of a combination of any of the following:

- 1 Simple nominal elements
- 2 Complex nominal elements

- 3 Prepositional phrase adjuncts + nominal group elements
- 4 Adverbial phrase adjuncts + nominal group elements
- 5 Existential "There" elements.

Below, are examples of some of these syntactic elements as realized in texts in the corpus.

(1) Details of the immunotherapy administration are shown in Table 2. (2) In 142 of 180 treatment courses, the majority of doses were 100,000 U per kilogram (25 and 13 patients, respectively). (3) Regardless of dosage, interleukin-2 was continued until side effects precluded further administration, thus, (4) the number of doses that could be given before limiting toxicity was reached was inversely related to the number of units per dose. (5) In treatment courses in which most doses were given at 100,000, 300,000 or 10,000 U per kilogram a medium of 20, 27 or 68 total doses were administered respectively.
 (RAP.14: *NEJM*, April 1987)

In the text above, utterances 1 and 4 contain complex nominal phrases which function as subject and theme. On the other hand, utterances 2 and 5 consist of a combination of prepositional phrase adjuncts and nominal group elements as subject and theme. The theme of utterance 3 is an adverbial group adjunct plus an NP. The themes of utterances 1 and 2 are related by semantic inference. That is, it may be inferred from the context that "treatment courses" in utterance 2 presupposes "Immunotherapy administration" in utterance 1. All other themes in the text are related by identical wording - that is, by the repetition of the item "doses" or "dosage" in the context of information indicating the point of departure for the message in the utterances. These relationships enable information to be organised along a constant pattern of Thematic Progression in the text.

(1) Congestive heart failure is a common condition reported to affect 1 per cent of the population, with an annual incidence of approximately 3 per 1000. (2) The prognosis of congestive hear failure is poor: (3) annual mortality is in excess of 50 per cent. (4) Vasodilator therapy for congestive heart failure includes haemodynamic effects that are believed to be beneficial and (5) it has become widely accepted during the past decade.

The text above contains two simple NP themes - utterances 1 and 3, though utterance 3 may also be interpreted as complex nominal theme with an ellipited NP. The themes of utterances 2 and 4 are made up of complex nominals. The themes of utterances 1, 2, 3 and 4 are related by identical wording - that is, the repetition of the item "congestive heart failure". This item is inferred in utterance 3 where the theme can be said to be "annual mortality from congestive heart failure". Thus, the themes of utterances in the text present a constant thematic progression pattern of discourse organization.

(1) Some characteristics of subjects are shown in Table II. (2) There was no difference in the mean ages. (3) There was a substantially higher proportion of cigarette smokers in both groups of cases than in the controls. (4) Blood pressure did not differ significantly between AP cases and controls. (5) There was no difference in total cholesterol between AP cases and controls, but HDL cholesterol was significant. (RAP.12)

The text above contains five utterance themes, out of which three are existential "There" elements. Although the "there" elements do not follow each other sequentially in a manner suggesting that only they account for the realization of the constant TP pattern in the text, they, nevertheless, contribute significantly to making it possible for this to happen.

On very rare occasions, proper names and personal pronouns do contribute to the realization of constant TP patterns in RAP texts. However, because the tendency for them to occur is low, they are not considered significant features of TP in the RAP. The only instances in which they were identified in the RAP corpus are presented below:

(1) Difficulties in conducting ophthalmologic examination of patients with dementia have largely precluded careful evaluation of their retinal and optic-nerve functions. (2) Cogan reported poor visual activity, alexia, agnosia and spatial disorientation, which he attributed to abnormalities in the parietal cortex in three patients with Alzheimer's disease. (3) Schlotterer et al similarly concluded that Alzheimer's disease affects central visual processing to a greater extent than it affects primary ocular functions. (4) Nissen et al observed that elevated visual

contrast-sensitivity thresholds in patients with Alzheimer's disease is suggestive of optic nerve abnormality.

(RAP.9: *NEJM*, 12 August 1986)

In the text above, the theme of utterance 1 is not a proper name. The utterance, as it were, provides background information to what follows in the rest of the utterances in that text. Therefore, its theme is not necessarily related to the theme of utterance 2. But the themes of utterances 2, 3 and 4 are proper names. Although the themes of those utterances do not refer to the same person, they are, nevertheless, conceptually and semantically related in that they represent persons who play identical roles in the context of discourse in the text.

(1) Having achieved good results with our preliminary attempts at immunisation, we decided to carry out a controlled trial of treatment with paternal cells. (2) We wanted to find out whether the success was truly immunological or a placebo effect of a positive treatment, such as has been found with other treatments for recurrent abortion. (3) We also wanted to study the possibility that immunisation with paternal cells might produce subfertility or other unwanted side effects.

(RAP.I: *The Lancet*, 27 April 1987).

A constant thematic progression pattern is maintained in utterances in the text above by the repetition of the personal pronoun 'we' in thematic position in utterances 1, 2, and 3.

As has been pointed out, the JRV differs from either the RAP and the ABS in the elements which it frequently uses to indicate perspective in the utterance. This is not to suggest that the JRV does not make use of any of the items identified in the RAP and ABS. Rather, what it means is that while a number of the syntactic elements identified in the ABS and RAP do also occur as theme in JRV texts, they do not constitute the primary means by which authors indicate perspective in texts in the genre. Thus, JRV writers tend to indicate perspective primarily through the thematization of:

- 1 Proper names
- 2 Personal pronouns - they, their
- 3 Demonstrative pronouns - this, that
- 4 The pronoun "it"
- 5 Adverbial phrase + NP

Below are some examples of the realizations of these elements in text.

(1) Young and his colleague Donald Clark used monoclonal antibodies made against part of the receptor to search for the same receptors in the epithelial cells - and found them. (2) Young's team is now growing squamous epithelial cells in tissue cultures and infecting them with ER virus. (3) They describe their results as "encouraging". (4) When such infected cultures are established, then, says Young, there are some exciting possibilities. (5) Once they have worked out the detailed structure of the receptor it may be possible to develop inactive analogues to the 3cd molecule which could enter and block the receptor and so prevent infection.

(JRV.4: *New Scientist*, March 1986)

In this text, thematic progression is achieved by the constant thematization of proper names and personal pronouns. Thus, the theme of utterances 1, 2 and 4 contain information which make reference to either "Young" or his colleague Donald Clark or both. On the other hand, the elements which are significant for interpreting the themes of utterances 3 and 5 is the personal pronoun "they".

(1) In Britain, around 15,000 couples are affected by the particularly distressing form of infertility known as recurrent spontaneous abortion. (2) They usually have no special difficulty achieving pregnancy but (3) they repeatedly miscarry in the middle of stages of pregnancy. (4) The cause of this condition is not known and no therapy has proved consistently effective.

(5) Last week, however, James Mowbray and colleagues at St Mary's Hospital in London described in "The Lancet" a treatment which has enabled nearly 80 per cent of affected women to complete normal, successful pregnancy after having as many as eight consecutive miscarriages (Vol 1, p941).

(6) Each woman's partner was asked to give a unit of blood, from which the lymphocytes were removed and washed. (7) These were used to immunise the woman by simple injection,

and (8) some weeks later blood samples were taken to find out if the woman had made antibodies to her partner's lymphocytes. (9) The trial of this procedure took account of the known tendency for recurrent miscarriage to improve slightly during any treatment. (10) To control for this placebo effect, the women were randomly allocated to one of the two groups:
(JRV.1: *New Scientist*, 9 May 1985)

As can be seen in the text above, information is organised and developed from one utterance to the other by means of adverbial phrase adjunct + NP (utterance I, 8), personal pronoun (utterances 2 and 3), proper names (utterance 5), demonstrative pronouns (utterances 4, 7, 9 and 10) sometimes in association with complex nominal elements.

(1) The Epstein-Barr (EB) virus is one of five herpes viruses that attack humans. (2) In Europe and the US, infection can produce glandular fever (mononucleosis). (3) In Central Africa it can lead to Burkitt's lymphoma, a facial cancer in children. (4) In China and South-East Asia, EB virus plays a part in triggering the development of nasopharyngeal carcinoma (npc). (5) In glandular fever and Burkitt's lymphoma, the virus infects one type of blood cell and the B lymphocytes.
(JRV.4: *New Scientist*, 6 March 1986)

In the text above, a constant TP pattern of discourse organization is achieved by the repetition in subsequent utterance themes of a combination of adverbial phrase adjuncts and NPs which function as subject and theme. The adverbial phrase adjuncts set the scene and indicate the perspective from which the more thematic NP elements are presented.

These differences between the RAP and the JRV highlighted in the examples above raise one fundamental question: "Why do these genres which develop related information by means of the same TP pattern do so using different syntactic elements as theme?" Part of the reason for this variation may lie in the need to reconcile the problem which exists, especially in the RAP, between the requirements of the FSP which insist in most cases on the theme being placed before the rheme and the requirements of the conceptual structure of information in science and medical research reports which places emphasis on impersonality and the scientific activity in the

presentation of facts. The latter part of this contention is, as we have seen, in line with Myer's (1986) characterization of the professional article as representing the "narrative of science".

The most common means by which perspective is indicated in a constant TP pattern of discourse development in the RAP is by the use of NPs some of which are complex and others which are preceded by either prepositional or adverbial phrase adjuncts. In most cases these nominal elements tend to take the form of non-agent subjects functioning as theme, as shown in the example below:

- (1) Cross sections of the intercranial optic nerves were cut in thicknesses of 0.5 μ m and immersion-fixed (in 2 percent paraformaldehyde and 2 percent glutaraldehyde) for one week.
- (2) The tissue was raised and immersed in 0.5 percent osmium tetroxide in the cold overnight, dehydrated through a graded series of ethanol solutions and embedded in polybed. Sections that were 1 μ m thick were stained in 1 percent P-phenylenediamine in methanol with use of the method of Sadun and Schaechter. (RAP.9: *NEJM*, 21 August 1986)

In the text above, the non-agent subject themes function as the grammatical subjects of the passive constructions in which they are realized. Chafe (1975: 219) has pointed out that the semantic function of a passive verb is most easily explained if the passive is regarded as an inflectional unit "which can be added to an action-process (as well as an experimental) verb" such that the presence of the inflectional unit "allows an action-process or experimental process verb to exist without an accompanying agent or experience". Following from this, Chafe (op cit) sums up the function of passives in discourse as:

- 1 to allow an action-process or experimental process verb to occur without an agent or experience
- 2 to change the order of priorities for the distribution of new information.

This latter function is related to the observations made above concerning the realization of non-agent subjects as theme in RAP texts. Thus, evidence from the data, as supported by Myers' (op cit) characterization of the discourse of professional research articles and Chafe's (op cit) observations on the functions of the passive lend weight to the contention made earlier on in this discussion that the variation in the syntactic elements which the RAP and the JRV use to indicate perspective in a constant TP pattern of discourse organization is constrained by the need to reconcile the FSP of utterances in the RAP with the demands of the conceptual presentation of knowledge in science and medical research papers.

This problem of reconciliation is less marked for the JRV than it is for the RAP. This is probably because the JRV is essentially journalism, albeit with a difference - science journalism. As with other forms of journalism, the JRV is more of a narrative genre (Smith and Frawley, 1982) in having a participant line. It has been pointed out (Brown and Yule, 1983: 143) that "in literature the author frequently assigns the role of narrator to one of his characters and then manipulates the knowledge which the reader needs so that it can be plausibly known to and recountable by the narrator". The same strategy can be said to be at work, albeit in a more subtle form in the JRV. In the JRV, it takes the form of constant thematization of the names of the researcher(s) who conducted the research being popularized. Thus, in re-writing, the "narrative of science" of the RAP to what Myers (op cit: 2) refers to as the "narrative of nature" the JRV writer uses the names of the researchers who wrote the RAP article being popularized to create the participant line needed to put the story in perspective (see Example on page 259, JRV.4). As can be seen in that text, the researchers "Young and his colleague Donald Clark" are thematized in the first utterance. In the second utterance they are referred to as "Young's team" also in thematic position in the utterance. This is in line with Myer's (op cit) characterization of popularizing articles as "presenting a sequential narrative of nature". But in the popularized texts in this study, unlike in Myers' in which the plant or animal is the subject, authors of the texts

being popularized are the subjects and theme of utterances. The difference between both types of popularization in this respect is due partly to the role of the popularizers themselves and to the nature of the disciplines in which the popularizations are based.

Perhaps another reason why perspective in constant TP utterances in the RAP tend to differ from those in the JRV is the need for RAP writers to detach themselves from the text and the reader in accordance with established practices in the scientific community. This is achieved primarily by the thematization of complex nominals, especially those which function as non-agent subjects and by the re-location of thematic focus by means of existential construction (see examples on pages 256 to 258 above). The role of non-agent subject themes has already been discussed. In addition to that discussion it may be observed that most non-agent subject themes in the RAP tend to focus mainly on diseases, eg. "Congestive heart failure"; on specimens, eg. "Cross sections of interactional optic nerves ..."; experimental procedure, eg. "Details of immunotherapy administration ..." etc. These, would seem to be elements which tend to function to detach the writer from the subject-matter of discourse because they function to replace the agent in the utterances as a means of achieving what is usually regarded as "scientific objectivity". This contention is supported by observations made by Chafe (1982) with regards to the functions of the passive in discourse. According to Chafe (op cit: 45).

The detached quality of written language is manifested in devices which serve to distance the language from specific concrete states and events. Such a device in English is the passive voice, which suppresses the direct involvement of an agent in an action.

In a related development, Chafe (op cit: 46) also points out that nominalization is another device associated with detachment, contending that nominalization suppresses involvement in actions in favour of abstract verifications. This would also seem to apply to complex nominal themes observed in the RAP. Following from these observations, it can be said that the choice of elements which function as theme in

constant TP patterns in the RAP is constrained by the need for writers to detach themselves from the subject-matter of discourse and the audience.

In the JRV, on the other hand, the difference in the means by which perspective is indicated in a constant TP pattern of discourse development is influenced by the need for the writer to involve the reader in the discourse. This objective is achieved primarily by the thematization of personal names in these utterances (see example on page 259 JRV.4 above). Personal names in themselves do not necessarily involve the reader in the discourse. However, the way in which they are used tends, in some circumstances, to influence the reader's attitude towards information being communicated in the text. In the JRV, there is a tendency towards an informal reference to persons despite the fact that there may be more formal ways by which these persons can be identified. To this end, note the informal, almost casual reference to "Young and his colleague Donald Clark" or to "Young's team" etc in the example referred to above. Informalities such as these appear to be strategies for breaking down barriers between the writer and the reader on the one hand and between the reader and the text on the other. This contention is supported by Chafe's (op cit: 46) contention that a speaker's involvement with his or her audience is manifested by their use of first person references. In the case of the JRV, it is manifested by the use of personal names of researchers as theme in constant TP utterances. It must, however, be pointed out that the RAP also has aspects of this function as can be demonstrated by the use of the first person plural pronoun "we" in the Introduction section to highlight the new research (see example on page 258 above). However, in the RAP, unlike in the JRV, it is difficult to say whether the use of *we* as theme in circumstances identified in the example reflects "a semantically empty grammatical necessity or whether it refers to the team or to the "we" which expresses modality, not only by injecting actors across into the specific scene" (Dubois, 1987: 106). Therefore, it is not clear at this stage what effect the "we" pronoun in the RAP may have on the reader. This will be discussed in greater details in the next chapter on Cohesion.

In conclusion, this discussion of the differences in the realization and distribution of TP patterns across the three genres in this study has shown that there is a relationship between TP and text structure in the genres; that the genres differ in the choice and distribution of TP patterns as well as in the choice of formal syntactic patterns that realize them. These differences are shown to be constrained by contextual factors in the texts and by the contingencies of text production by the writer and information processing on the part of the reader.

CHAPTER SIX

VARIATION IN TEXTUAL COHESION

6.0 INTRODUCTION

Thus far, attempts have been made to characterize the extent of variation which exists between the three genres in this study in terms of the Schematic organization of information and the organization of information into identifiable patterns of Thematic Progression (TP). The analysis of Schematic structures represents an attempt to characterize the differences which exist between texts in the three genres at the level of the global organization of information. The analysis of Thematic Progression, on the other hand, is concerned with characterizing the differences which exist between texts in the genres at the level of the organization of the message contained in utterances in a way that indicates the writer's perspective. Both analyses represent complementary points in a continuum of related strategies which seek to account for the patterns of discourse organization in the three genres in this study. A third point in that continuum is represented by the account in this chapter of the extent of variation which exists between the genres in terms of what Halliday and Hasan (1976) refer to as "cohesion".

An account of the extent of variation in the cohesive properties of texts in the three genres is consistent with the overall purpose of this study as stated in Chapter 1. As Halliday and Hasan (op cit: 4) point out, the concept of cohesive ties makes it possible to investigate the question of the relationship which exists between "cohesion and the organization of written texts into sentences and paragraphs, and the possible differences among different genres." This contention underlines the observation made in the preceding paragraph regarding the common relationship which the analysis of schematic structures and thematic progression reported in Chapters 4 and 5 share with the analysis of cohesion reported in this chapter. But, more importantly, it points to the strong relationship which exists between the concept of Thematic Progression and

cohesion — a relationship which is also acknowledged by Halliday and Hasan (op cit: 27) when they state that "cohesion is closely related to information structure", adding that the two, indeed, overlap at one point.

In view of these relationships, this aspect of the study which dwells on the differences in textual cohesion across the three genres shares the same basic assumptions and raises the same fundamental questions as those in the chapters preceding it. However, as with these other chapters, the questions in this chapter focus on issues specific to it; that is, issues aimed at highlighting the extent of variation which exists between the three genres in terms of textual cohesion. To this end, the study in this chapter seeks to provide answers to the following general questions.

1. Do texts which dwell on the same subject-matter, in the same field of discourse, but for different audiences and published in different contexts differ in the use they make of the resources of textual cohesion?
2. If they do, to what extent are the observed differences in the use and function of the resources of cohesion across these related genres due primarily to the pragmatic influences of audience, medium and purpose of discourse?
3. What elements of cohesion does each genre characteristically employ and what functions do the elements perform?
4. Do the various patterns of cohesion correlate with the schematic structures identified in Chapter 4 for each genre?

In order to provide answers to these questions, attempts will be made to provide a brief account of the concept of Cohesion, review a few models of Cohesion, present the procedure for analysis, provide model analyses of texts from each of the genres and discuss the differences which exist between the three genres with respect to textual cohesion. The model of cohesion adopted in this inquiry is that proposed by Halliday and Hasan (1976) with modifications from Halliday (1985) and Hasan (1984).

6.1 THE CONCEPT OF COHESION

The Concept of Cohesion is best explained in relation to associated concepts such as Text, Texture and Genre. Halliday and Hasan (op cit: 1) refer to text as any passage, spoken or written of whatever length that forms a unified whole. In other words, any sample of discourse with an apparent meaning and a uniform orientation constitutes a text. It may be made up of a single sentence, group of sentences or even elements less than a sentence. According to Halliday and Hasan (op cit), a text has structure in the same way that a lexicogrammatical unit such as the sentence is structured. But they point out that, unlike the lexicogrammatical unit, the structure of a text is not governed by formal linguistic rules. Explaining, they contend that:

A text, as we have said, is not a structural Unit,
Whatever relation there is among the parts of a text —
the sentence, or paragraphs, or turns in a dialogue — it
is not the same as structure in the usual sense, the
relation which links the parts of a sentence or a clause.
(Halliday and Hasan, op cit: 6)

In a related development, Halliday and Hasan (op cit: 2) contend that what distinguishes a text from something that is not a text is texture. It has been pointed out (Mosenthal and Tierney, 1984: 24) that by texture Halliday and Hasan mean that quality of a text which establishes its wholeness or unity rather than its meaning.

Thus, it would seem that "text" and "texture" are created in linguistically identifiable ways. Thus, as Halliday and Hasan (op cit: 2) point out:

If a passage of English containing more than one sentence is perceived as a text, there will be certain linguistic features present in that passage which can be identified as contributing to its total unity and giving it texture.

(Halliday and Hasan, op cit: 2)

It is the relation which exists between these linguistic features as they combine to contribute to the unity and texture of a passage that Halliday and Hasan refer to as cohesion. They explain the role of these linguistic features in creating texture by means of this example:

Wash and core six cooking apples. Put them into a fireproof dish.

(Halliday and Hasan, op cit: 2)

In the example above, "them" in the second sentence presupposes "six cooking apples" in the first. The relationship created between the pronoun and its co-referent (Brown and Yule, 1983) is referred to as a relation of cohesion. In the example above, the presupposing item (them) and the pre-supposed item (six cooking apples) constitute what is known as a "tie". The notion of a tie is what enables patterns of cohesion in text to be identified. It can be used to characterize any segment of text by reference to the number and kind of ties it displays. The concept is also useful as a means of analysing the cohesive properties of text and providing a systematic account of its pattern of texture. And, as has already been pointed out in this discussion, Halliday and Hasan (op cit) contend that various types of questions can be investigated by means of the concept of a tie. Some of these questions are: the difference between speech and writing, the relationship between cohesion and the organization of written texts into sentences and paragraphs as well as the possible differences among different genres and different authors in the number and kinds of tie they typically employ. This observation by Halliday and Hasan points to the importance of the study of textual

cohesion in the characterization of the extent of variation which exists between related genres, thus, lending support to its inclusion as one of the variables in this study.

6.2 APPROACHES TO THE STUDY OF COHESION

Having provided a brief explanation of the concept of cohesion, an attempt will be made at an outline of a few approaches to the study of cohesion in various kinds of text. This is intended to provide insights into alternative conceptions of the notion of cohesion as a means of putting the discussion in this chapter in perspective as well as preparing for the application of Halliday and Hasan's model. Selected for discussion are the works of Grimes, Longacre, Gutwinski and de Beaugrande.

Grimes (1975, 1977) views text as consisting of two sub-systems — staging and cohesion — both of which interlock with content in text. Staging has to do with devices which the speaker has at his disposal in controlling the perspective from which he presents everything he says. To this end, it appears to be closely related to Halliday's (1967) notion of 'theme' or to the notion of 'topic'. Cohesion, on the other hand, is concerned with the way information mentioned in speech relates to information that is already available.

Grimes characterizes cohesion into "information blocks", "information centres" and "overlays". By information blocks Grimes means the package of information that results from the decision which the speaker makes regarding how much of the content of discourse he thinks the hearer can take in at a time. It corresponds with Halliday's (op cit) "Information Unit" and relates to the observation by Halliday that the speaker decides what quantity of information he thinks the hearer is capable of assimilating, and in expressing that quantity, decides on what part of it is likely to be least predictable to the hearer. Grimes points out that the package of information that constitutes information may or may not correspond to some easily recognized sub-string or content, explaining that in English, the extent of the information block is

signalled by a single intonation contour. For example, THIS/is the FIRST TIME/ we have EVER (Capital letters indicate intonationally prominent words; slashes indicate intonation contours.)

According to Grimes, several factors enter into the decision to block information in one way or another. First, is the textual factor, which is concerned with information that has already been given in the discourse; with anything the speaker feels is already clear and is not likely to be singled out by being put in a separate block. Second, is the contextual factor. This has to do with expressions that refer to the immediate situation of speech. The third factor involves decisions on the rate of "information injection" which the speaker wishes to establish. For example, "in English, the highest rate of information injection is that of telegraphic style in which the apparatus of cohesion is squeezed to the minimum." The natural tendency, Grimes points out, is for the speaker to help the hearer by making his information blocks shift when his rate of information injection is high.

Grimes further contends that each information block contains at least one "information center". An information center is characterized as that part of the block in which new information is concentrated. The rest of the block is said to contain predictable material. Echoing the views of Halliday (1967), Grimes states that in English the center of an intonational block is identified by intonational prominence.

"Overlays" involve putting together two or more planes, each of which constitutes a narration of the same sequence of events. The first plane is made up mainly of new information. The second plane and others that follow it begin the sequence over again. These latter planes consist partly of new information that is being given for the first time in that plane and partly of information that is being repeated piecemeal from an earlier plane. This repeated information, Grimes contends, has the special status of tying the whole overlay together.

Grimes' conception of cohesion, as stated here, seems to be closer to the concept of thematic progression (see chapter, 5) than it is to cohesion in Halliday and

Hasan's terms Its inclusion in this review is to indicate how close the concept of cohesion is to that of TP and how this closeness can lead to terminological differences in the way these concepts are represented in various models of discourse representation. Since we have already examined variation in TP patterns across the three genres, Grime's model of cohesion is not considered the most appropriate to use in the analysis reported in this chapter.

Longacre and Levinsohn (1977) propose a model of cohesion for analysing narrative texts. Three strands of cohesion are identified. They are an event-line, an agent-line and a repartee-line.

The event-line, also termed the "back-bone" of the story indicates successive events, times and places or a combination of these. Materials in the story may either be on the event-line or off it; they may be "progressions" or "digressions" from it. Aspects of discourse in narrative texts such as background materials, setting and collateral material are characterized as digressions from the back-bone of the story. Even within the back-bone itself, events are said to carry varying degrees of importance.

The agent-line or the "agent-action-axis" functions to track the participants through the discourse. As with background materials, minor participants are said to be off the agent-line. In stories which have extensive repartee, a line of repartee development may also be identified and classified. In addition to the agent, event and repartee lines, Longacre and Levinsohn also identify a fourth cohesion strand as "focal intention". It is concerned with questions such as, "Why is the discourse told in the first place? How was it elicited? What situation provoked its being given?" Clues to the focal intent of a discourse may occur almost exclusively at its beginning and end (bundled focal content) or it may crop up here and there throughout the discourse (scattered focal content).

The strands of cohesion identified above are supported by an identification of surface cohesive devices. The devices are:

1. The role of tense and aspect as indicators for identifying favoured mainline development of an event. Longacre and Levinsohn contend that the occurrence of a given tense/aspect tells us something about the classification of a text as to discourse type and provides cohesion throughout that text.
2. The use of particles and affixes such as those associated with nouns. They tend to mark the central character as opposed to other major participants, or major as opposed to minor participants.
3. Participant Anaphora. That is the identification of participants by name, by a common noun, by pronoun, by an affix or by zero.
4. Deictics, which function to help keep track of participants.
5. Lexical ties and paraphrase. They include relations such as synonymy, antonymy and items of the same semantic domain.
6. Summary and preview by means of generic predicates and substitutes on the one hand, and by giving in one part of a discourse what is to follow later, on the other.
7. Conjunctions and introducers which function to mark whether the subject of the clause which follows the conjunction is the same or different from the subject of the clause which precedes it.
8. Back-reference, involving the repetition of part or all of the preceding sentence at the onset of the next sentence. For example: "John took off in his Piper Club for Denver this morning. After taking off in the plane he"

The model of cohesion proposed by Longacre and Levinsohn (op.cit) raises a number of interesting questions which might be very relevant to the analysis of cohesion in the JRV, which as with most journalistic prose tends to exhibit some of the characteristics of narrative genres. However, since the model does not specify in great details the linguistic features against which the cohesion strategies it proposes can be determined, I consider it inadequate for our purposes in this research. Nevertheless, we shall draw useful insights from the issues discussed in aspects of the model in our interpretation of cohesion in the JRV.

de Beaugrande and Dressler (1981) propose a model of cohesion based on the notion of "continuity of occurrences". The model is aimed at showing how the concept of cohesion can be applied to abstract grammars such as semantic network models of information processing in text.

Thus, cohesion is characterized as the "sticking together" of syntax in communication. Pointing out that materials are placed in active storage or "working memory" where processing resources are distributed among elements of a presentation according to their importance, de Beaugrande and Dressler contend that surface structure elements are most predominantly maintained in "short-term" storage, while conceptual content is held in "long-term" storage. These cognitive factors are said to relate to the function of syntax in providing closely-knit patterns into which current materials of various sizes and complexity can be fitted. These patterns are based on well marked dependencies involving the Phrase, the Clause and the Sentence which can be utilized in a short span of time as processing resources. In addition to these, there are also devices for showing how already used structures and patterns can be re-used, modified or compacted in long-range stretches of text. These devices are RECURRENCE or straightforward repetition of elements, PARTIAL-RECURRENCE or the shifting of already used elements to different classes (e.g. from noun to verb), PARALLELISM or repeating a structure with new elements and PARAPHRASE or the repeating of content with different expressions. Others are PRO-FORMS or replacing the content-carrying elements with short place-holders of no independent content and ELLIPSIS or the repeating of a structure and its content but omitting some of the surface expressions.

de Beaugrande and Dressler also identify surface signals which serve to indicate the relationships which exist among events or situation in the textual world. These are the use of Tense, Aspect and Junction, as well as the ordering of expressions in accordance with their Functional Sentence Perspective and the indication of importance in spoken texts by means of intonation.

These strategies of cohesion are then related to Abstract Grammars such as the AUGMENTED TRANSITION NETWORK which consists of a configuration of NODES or "Grammar States" connected to LINKS or "Grammatical Dependencies". de Beaugrande and Dressler's model also reinforces the relatedness which exists

between FSP and cohesion. However, it appears that the model of text analysis it proposes is too all-embracing, straddling, as it were, a number of issues which impinge on textuality, of which "cohesion" in Halliday's "Systemic-Functional" grammar is only one. Thus, the aspect of the model which dwells on cohesion is not sufficiently detailed to warrant being used as a basis for an elaborate discussion of cohesion. In view of this, de Beaugrande and Dressler's model is not adopted in this study. Nevertheless, the model provides useful insights for approaching the issue of cohesion in texts.

Gutwinski (1976) proposes a model for accounting for cohesion in literary texts. The model is based on the stratificational approach to the organization of language by Gleason (1964) and Lamb (1966). Cohesion is defined simply as the relation obtaining among the sentences and clauses of a text. That is, relations which occur on the "grammatical stratum" and are signalled by certain grammatical and lexical features such as anaphora, substitution, and co-ordination, all of which reflect discourse structure on a higher "semiologic stratum". Gutwinski points out that these grammatical features do not in themselves constitute cohesion, but that they are, nevertheless, cohesive. He further points out that cohesion does not constitute discourse structure, but "it reflects indirectly, perhaps in part only, the underlying semiologic structure of text, that is, the discourse structure conceived at the semiologic stratum".

Underlying Gutwinski's characterization of cohesion is the notion of "order" as a cohesive device. This is based on the understanding that the order in which sentences follow one another in a text is an important feature in the recognition of a group of sentences as a text. Thus, as Gutwinski points out, the order in which clauses and sentences follow in a text is a cohesive factor which in combination with other cohesive factors indicates what kind of cohesive relations obtain among the sentences and clauses.

Following from these contentions, Gutwinski makes a list of cohesive features in literary texts as follows:

A. GRAMMATICAL

1. Anaphora and Cataphora

- a) pronouns
 - i) personal pronouns, e.g. he, him, she, it, they
 - ii) demonstrative pronouns: this, these, that, those
 - iii) relative pronouns: who, which, that, whom, whose
- b) determiners: the, this, these, that, those
- c) Personal possessives, e.g. his, its, their
- d) substitutes
 - i) verbal (do)
 - ii) nominal (one)
 - iii) partial
- e) adverbs, e.g. there, then
- f) submodifiers, e.g. such, so

2. Coordination and subordination

- a) connectors

3. Enation and agnation

- a) enate sentences
- b) agnate sentences

B. LEXICAL

1. Repetition of item
2. Occurrence of synonym or item formed on same root
3. Occurrence of item from same lexical set (co-occurrence group)

As we shall see later, Gutwinski's model of cohesion is closely modelled after that by Halliday and Hasan (1976). However, as Gutwinski points out, the two listings (i.e. his and Halliday and Hasan's) differ in the manner of classification. They also differ in the range of expressions they use to categorize some cohesive relations. In this regard, in addition to terms such as "anaphoric" and "cataphoric", Gutwinski

introduces newer ones such as "paraphoric". By paraphoric Gutwinski means a situation when "an item refers to something in another text, e.g. another work of art (say, an item in one of Shakespeare's plays refers to something very specific in another play of Shakespeare) expected to be known and related to the text at hand in a special way". Other new terms used by Gutwinski include "enation" and "agnation": these terms are used to avoid the risk of misunderstanding and imprecision involved in the general use of terms such as structural parallelism and structural similarity. According to Gleason (1965: 199) "two sentences can be said to be "enate" if they have identical structures, that is, if the elements, say words, at equivalent places in the sentence are of the same classes, and if constructions in which they occur are the same". To illustrate this, Gutwinski (op cit: 77) cites the following example:

.... As Caesar loved me, I weep for him; as he was fortunate, I rejoice at it; as he was valiant, I honour him; but as he was ambitious, I slew him. There is tears for his love; joy for his fortune; honour for his valour; and death for his ambition

(Shakespeare, *Julius Caesar*, III ii, 26-31)

Agnation, on the other hand, refers to relations which are opposite and complementary to enation. Again, according to Gleason (1965: 202) "Pairs of sentences with the same major vocabulary but with different structures (generally shown by differences in arrangement, in accompanying function words, or other structure markers) are agnate if the relation in structure is regular and systematic, that is, if it can be stated in terms of general rules". Gutwinski (op cit: 78) illustrates cohesion by agnation with this example:

There was nothing left for her but to sell the old family house. This she couldn't do.

Gutwinski argues that the use of the agnate sentence "This she couldn't do" instead of its basic form "She couldn't do this" is motivated by considerations which have to do

with appropriate transitions from sentence to sentence. He, therefore, contends that the agnate structure in the example above is a cohesive factor in the relations holding between the two sentences. In this regard, Gutwinski's model of cohesion puts considerable emphasis on thematic progression as cohesion.

So far, we have attempted to present possible approaches to the analysis of cohesion in text. The aim has been to point out the weaknesses inherent those approaches and to indicate why, owing to these weaknesses, the models are considered inadequate for the purposes of the study reported in this chapter. Therefore, they tend not to discuss both concepts with sufficient rigour and clarity to enable us adopt any of them as a basis for the kind of analysis reported in this chapter.

6.3 HALLIDAY AND HASAN (1976) MODEL OF COHESION

Halliday and Hasan's (1976) model of cohesion represents the most definitive statement on the subject existing at the moment. It is based on assumptions which underlie Halliday's Systemic-Functional Grammar. Basic to these assumptions are: (1) that the linguistic system consists of a phonological, a lexico-grammatical and a semantic component; (2) that the system also consists of an ideational (experiential and logical), an interpersonal and a textual component; (3) that all these components are mapped onto and realized as surface feature elements of text in the form of the lexicogrammatical component; (4) that the components of the linguistic system are organized by rank with the units at a lower rank forming the building blocks upon which the ranks above are built.

This idea of layering is also built into the textual component of the linguistic system. Based on the concept, the textual component is shown to consist of:

- a) the structure-generating systems — thematic and information systems
- b) the cohesive relations — made up of Reference, Ellipsis, Substitution, Conjunction and Lexical relations

Reference cohesion is concerned with the relation which holds between an element of text and something else by reference to which it is interpreted in the given instance. It consists of three types: Personal, Demonstrative and Comparative. Personal reference is that form of reference which is based on the idea of the role of the category of "person" in the speech situation. It may take the form of Speaker only (i.e. I, me, my, mine), Addressee with or without other person — (you, yours, your) etc. Demonstrative Reference is based on the idea of proximity.

For example: Near — this, these, here, now

Far — that, those, there, then.

Comparative Reference is based on indirect reference by means of identity or similarity.

For example: Identity — same, identical, equal

General Similarity — similar, additional, etc.

All three forms of reference are fundamental to text. That is they are basic text-forming properties of text. They serve to indicate that the interpretation of an item is to be sought somewhere else by relating presupposing and presupposed elements in text. Therefore, the relations they describe are essentially semantic ones.

Substitution is defined as the replacement of one item by another. It differs from reference in that it is a relation in the "wording" rather than in the meaning. As Halliday and Hasan (op cit: 89) point out, "Substitution is a relation between linguistic items, such as words, or phrases; whereas reference is a relation between meanings". In other words, reference is a semantic relation while substitution is a grammatical one. Halliday and Hasan identify three types of substitution as:

Nominal: one, ones, same

Verbal: do

Clausal: so, not

Ellipsis is very much like substitution. Halliday and Hasan (op cit: 143) refer to it as "substitution by zero". Both ellipsis and substitution, therefore, embody the same kind of relations between parts of a text — that is, a relation between words or groups or clauses. Despite this fact, they are nevertheless, distinct from each other. Ellipsis, according to Halliday and Hasan, is "something left unsaid" with the implication that what is left unsaid is understood, as in the example:

"Would you like to hear another verse?
I know twelve more."

In the example above, the Head noun verses is ellipped from the nominal group "twelve verses" in the second sentence. That Head noun is presupposed from the first sentence. Therefore, the relation which exists between the presupposing item in the first sentence and the presupposed ellipped item in the second is cohesive. Halliday and Hasan classify ellipsis into three: — nominal ellipsis, verbal ellipsis and clausal ellipsis.

Conjunction as a cohesive relation serves to specify the way in which what is to follow is "systematically connected to what has gone before". As Halliday and Hasan (op cit: 227) point out, "in describing conjunction as a cohesive device, we are focusing attention not on the semantic relations as such, as realized throughout the grammar of the language, but on one particular aspect of them, namely the function they have of relating to each other linguistic elements that occur in succession but are not related by other, structural means". Conjunctive cohesion elements fall under four main headings: additive, e.g. and; adversative, e.g. but; causal, e.g. so, and temporal, e.g. then. These general categories are further divided into finer categories.

Lexical cohesion is described as the cohesive effect which is achieved by the selection of vocabulary. It consists of two categories: reiteration and collocation.

According to Halliday and Hasan (op cit: 278) reiteration is "a form of lexical cohesion which involves the repetition of a lexical item". It may take the form of a repetition, a synonym or near-synonym, a superordinate or a general word. These are often accompanied by the reference item, the. Halliday and Hasan emphasize that two lexical occurrences need not have the same referent to be cohesive, pointing out that "many instances of cohesion are purely lexical", being a function of "the co-occurrence of lexical items, and not in any way dependent on the relation of reference. Collocation, on the other hand is concerned with the cohesive chains formed by items that have a strong tendency in the system to co-occur. It consists of synonyms and near-synonyms, e.g. climb — ascent; superordinates, e.g. boy — child; opposites, e.g. boy — girl; antonyms, e.g. love — hate and converses, e.g. order — obey etc. Therefore, the category of collocation overlaps with that of reiteration in some areas, leading to confusion in separating them as distinct relations in text. Both forms of lexical cohesion contribute to the structure of text. However, items which belong to the category of collocation are not easy to identify. Halliday and Hasan (1976: 284) refer to it as "the most problematical part of lexical cohesion". It is probably in view of this fact that Hasan (1984) proposes a new model of lexical cohesion which attempts to remove the problems associated with the identification of collocation items in the analysis of lexical cohesion.

The models of cohesion reviewed in this section have one thing in common. They are all concerned with accounting for the relations which hold between elements of text. As we have seen, most of them tend to conflate TP and cohesion. Of all the models, Halliday and Hasan (1976) appears to be more comprehensive and tends to display a much higher degree of stringency that distinguishes it as the most comprehensive theory of cohesion. Its underlying assumptions, methodological orientation and practical application are more suited to the kind of problems this study is interested in than any of the other models characterized in this discussion.

6.4 METHODOLOGICAL ISSUES

As has been pointed out, this study is based on Halliday and Hasan's (1976) concept of cohesion in English. This is not to suggest that the study is a completely unmodified version of Halliday and Hasan's cohesion model. Rather the model is adopted as a conceptual framework which can be modified in line with the purpose of an investigation and the data on which it is based. For instance, this study adopts Halliday and Hasan's characterization of cohesion as the relation which holds between sentences in text. In this study the term text is taken in its wider sense to include the context and the situation in which the texts are realized. In other words, cohesion is seen in this study as a dynamic and motivated relation linking the text to its communicative purpose and to its context of situation. This conception of cohesion contrasts with that which sees cohesion as a static relationship involving frequency counts of cohesion ties (e.g. Smith and Frawley, 1983).

In line with the above, all five of Halliday and Hasan's main categories of cohesion — Reference, Substitution, Ellipsis, Conjunction and Lexical — are examined in this study. However, they are examined primarily for the communicative function they perform in text and the role they play in distinguishing one genre from the other in terms of texture. Therefore, aspects of the sub-categories of each of these five main cohesion types which do not further the cause of this primary purpose are not discussed in this Thesis.

Following the conception of cohesion posited above, enabling categories are introduced to cater for that aspect of the characterization which relates cohesive items to the extra-linguistic context of situation. Thus, five phoric relations of reference cohesion are subsumed:

- a) Anaphora - reference backwards to some preceding items(s) in the discourse

- b) Cataphora - reference forwards to some item not previously mentioned in the discourse
- e) Exophora - reference to an element that lies in a situation entirely outside the text
- d) Homophora - reference to one's general knowledge
- e) Paraphora - reference to another text

"Homophora" and "paraphora" are terms taken from Gutwinski (1976). Examples of each of these relations, if any, will be provided in subsequent discussions on the analysis of cohesion in genres.

However, in accordance with Halliday and Hasan, three categories of reference cohesion are examined as shown below:

A. Personal Reference

- i) personal pronouns - he, him, she
- ii) demonstrative pronouns - this, these
- iii) relative pronouns - who, which
- iv) possessive pronouns - his, its
- v) proper names - John, James

B. Demonstrative Reference

- i) determiners
 - this, these
 - that, those
 - the
- ii) adverbial adjunct - here, there

C. Comparative Reference

- i) identity - same, identically
- ii) general similarity - similar, likewise, so, such
- iii) particular comparison - better, more

No major changes are made to Halliday and Hasan's categories of Substitution and Ellipsis. Therefore, the three main sub-categories of substitution and ellipsis identified in Halliday and Hasan are examined. They are: nominal, verbal and clausal substitution and ellipsis respectively. Examples of each of these will be provided whenever they are encountered in the data.

Conjunctive cohesion, the formal and logical means by which sentences are linked together in text, is examined based on modifications introduced in Halliday (1985a). The modified version is adopted because it appears more comprehensive than the original version in Halliday and Hasan (1976). Therefore, based on Halliday (1985a), three main types of conjunctive cohesion are examined in this study. The conjunctive cohesion types and their sub-categories are provided below:

A. Enhancement

- i) Spatio-temporal e.g. then, afterwards, finally, etc.
- ii) Cause-conditional e.g. so, therefore, consequently, etc.
- iii) Matter enhancement e.g. here, there, elsewhere, etc.

B. Extension

- i) Positive addition e.g. and, moreover, in addition, etc.
- ii) Adversatives e.g. but, yet, however, etc.
- iii) Variation e.g. on the contrary, alternatively, on the other hand, etc.

C. Elaboration

- i) Particularizing e.g. in particular, more especially, more importantly, etc.
- ii) Summatives e.g. briefly, in conclusion, in short, etc.
- iii) Exemplifying e.g. for example, for instance, thus, etc.
- iv) expository e.g. in other words, that is (to say), etc.

In terms of Lexical Cohesion, this study adopts the revision of lexical cohesive categories advanced in Hasan (1984). According to Hasan (op cit), "lexical cohesion appears to belong to two primary types: that mediated through 'general' lexical relations and that mediated through 'instantial' ones".

By the category of general lexical cohesive devices, Hasan means those devices "that are based upon semantic bonds which are supratextual, with a language wide validity". It includes relations which are based on identity of reference (repetition, synonymy), non-identity of reference (antonymy), and super-ordination of lexical items (hyponymy, meronymy).

Instantial lexical devices, on the other hand, refer to those "lexical relations that are text bound". As Hasan (op cit) points out, "their validity is an artefact of the text itself, and does not extend to the system". That is, their validity is attained only through the linguistic context of the utterance in which they occur. They involve relations which are based on referential identity (equivalence, naming, and semblance). Categories in the revised lexical cohesion relations by Hasan (1984) are shown in the table below:

Table 23: Hasan (1984) Revised Categories of Lexical Cohesion

<u>Categories of Lexical Cohesion</u>			
A. General			
i	repetition	e.g.	leave, leaving, left
ii	synonymy	e.g.	leave, depart
iii	antonymy	e.g.	leave, arrive
iv	hyponymy	e.g.	travel, leave (including co-hyponyms, leave, arrive)
v	meronymy	e.g.	hand, finger (including co-meronyms, finger, thumb)
B. Instantial			
i	equivalence	e.g.	the <u>sailor</u> was their <u>daddy</u>
ii	naming	e.g.	the <u>dog</u> was called <u>Toto</u>
iii	semblance	e.g.	the <u>deck</u> was like a <u>pool</u>

Hasan's revision of lexical cohesion categories is adopted in this study because in excluding collocation from the category of lexical cohesion and introducing new ones such as antonymy, hyponymy, meronymy, equivalence, etc, it presents more operational criteria for identifying lexical cohesion ties in text.

6.5 THE ANALYSIS OF COHESION IN SAMPLE TEXTS

In this section, an analysis of cohesion in representative text samples taken from the three genres is presented. The sample texts are examined for cohesion using categories outlined in Section 6.4 above. The aim is to provide a description of the main cohesive elements which distinguish one genre from the other. Since the genres are related in some respects, a number of their cohesive elements may coincide, overlap or repeat across the genres. However, since this is a comparative study of cohesion in genres, repetition is considered a necessary process of showing what is considered cohesive in each genre and why.

The account of cohesion presented by means of the sample texts is only an exemplification of the kind of analytic work to which texts in the corpus have been subjected. It is not supposed to be a complete analysis of all possible cohesion ties in the sample texts. Rather, it is meant to provide an insight into the procedure adopted in the identification of cohesive elements and the constitution of cohesive ties. It also provides a rational basis for making general statements regarding differences in cohesion across the three genres.

For the purpose of this study, any element (grammatical or lexical) or item which functions to hang two sentences together is considered a "tie". The second sentence of the adjacent pair is examined for the elements which function to link that sentence with the preceding sentence. This process is extended to groups of related sentences and to the entire text.

The analysis of each sample text is presented in stages. First, the sample text to be analysed is presented. The sentences in it are numbered serially from beginning

to end. Second, a tabular representation of the pattern of cohesion in the text is provided. The table indicates how the various cohesion types are realized in the sample text and how they function to organize related information. Information contained in the table includes: sentence number, cohesive item, cohesion type and the presupposed item to which the presupposing item or element relates. Finally, a discussion of the pattern of cohesion in the sample text is provided. The order of presentation and analysis of sample texts is RAP, then ABS and finally JRV.

6.5.1 Cohesion in the RAP

All the sample texts used in the analysis of cohesion in the RAP are taken from RAP.7 (*The Lancet*, September 6, 1988). The sample texts consist of four extracts from the Introduction, Methods, Results and Discussion sections of the RAP text. This method of selection is intended to satisfy the criteria of representativity discussed in Chapter 3. The order of presentation, analysis and discussion of the sample texts is Introduction, before Methods, then Results and Discussion.

The Introduction Section:

Passage:

¹Since 1940, when there were over 12 000 notifications in England and Wales, meningococcal disease has been much less common. ²Following the last peak in 1974, (1296 notifications) the annual number of notifications declined steadily until 1984, when 401 cases were notified. ³In 1985, 549 cases were notified and the rise has continued into the first quarter of 1986.

⁴During the first half of this century, group A organisms were responsible for major outbreaks of meningococcal disease in England and Wales. ⁵Since 1970 group B strains have predominated to an increasing degree and now account for about 60% of infections. ⁶The introduction by Fasch of a typing system for the protein surface antigens in the early 1970's increased the potential for epidemiological investigation. ⁷In 1974/75, 60% of group B isolates tested in England and Wales were type 2. ⁸The prevalence of this type has since declined while that of type 15 organisms has risen from 15% of all group B strains in 1979 to 56% in 1985 (Jones D M, personal communication).

⁹Meningococcal disease in Stroud, a rural district in Gloucestershire, conformed to the national pattern until 1982, when the isolation of a sulphonamide-resistant meningococcus marked the beginning of an outbreak. ¹⁰A surveillance scheme was set up to

record all cases of meningococcal disease in residents of Gloucester Health District (population 301 537) from October 1, 1981 (population figures listed in this paper are taken from the 1984 estimates of the Gloucestershire county planning department). ¹¹This report includes all cases recorded up to March 31, 1986.

Table 24:
Sample Analysis of Cohesion in Introduction Section of the RAP

<u>Sentence No.</u>	<u>Presupposing Item</u>	<u>Cohesion Type</u>	<u>Presupposed Item</u>
1	—	—	—
2			
3	notifications 549 cases	L	notifications 401 cases
4	notified meningococcal disease	L	notified meningococcal disease (S.1)
5	England and Wales group B	L	England and Wales (S.1) group A
6	1970's	L	1970
7	group B isolates England and Wales	L	group B strains England and Wales
8	<u>this</u> type since declined risen	R C L L	type 2 1974/75 declined (S.2) rise (S.2)
9	meningococcal disease national	L L	meningococcal disease England and Wales
10	meningococcal disease Gloucester Health District	L L	meningococcal disease Gloucester
11	This report all cases	R L	this paper cases (S.3)

Key: R = Reference cohesion, L = Lexical cohesion, C = Conjunction

The passage above represents the whole of the Introduction section of text RAP.7. The text is concerned with reporting research on the outbreak of meningococcal disease in Gloucester. Therefore, the topic of discourse in the text can be said to be "meningococcal disease", which functions appropriately as the subject of the opening sentence of the first paragraph. After its initial introduction into the discourse in Sentence 1, the item is subsequently repeated in Sentences 4, 9 and 10

either as subject or direct object. It is interesting to observe that the repetition of the lexical item "meningococcal disease" is not fortuitous, for it seems to occur only at strategic points in the discourse. For instance, it occurs as one of the items in the opening sentences in the first, second and third paragraphs of the text (Sentences 1, 4 and 9). It also occurs as the initial sentence (Sentence 10) in what, following the discussion in Chapter 4, has been characterized as a 'Move' (see Chapter 4 on Schematic structures). These sentences correspond with what in composition studies is referred to as "topic sentence". The occurrence of the lexical item as both grammatical subject and topic sentence in most of these sentences further confirms the initial assumption made in this discussion that meningococcal disease is the topic of discourse in the text. But more importantly, it points to the fact that the repetition of the item is a conscious cohesive strategy for linking segments of discourse in the text together.

Apart from meningococcal disease there are also other cases of explicit repetition of lexical items in the text. They are: England and Wales (S. 1, 4 and 7), Gloucester (S. 9 and 10), notifications (S.1 and 2) notified (S.2 and 3), and group B (S. 5 and 7). Two of these — England and Wales, and Gloucester — fall within the category of "naming" in Hasan's modified version of lexical cohesion. As with meningococcal disease, their repetition is not unmotivated. A look at the instances of use of the lexical items shows that they tend to occur as part of a complex noun phrase, usually as the head of a prepositional phrase. For example, "over 12 000 notifications in England and Wales (S.1), "major outbreaks of meningococcal disease in England and Wales" (S.4), "meningococcal disease in Stroud, a rural district in Gloucestershire" (S.9) etc. Apart from functioning to hang sentences in the text together, the items also function to limit the topic of discourse in the text to particular locales. The other explicitly repeated lexical items in the text also contribute to the development of aspects of the topic of discourse contained in identifiable segments of discourse in the text.

In addition to the explicit repetition of lexical items, there exist other less explicit forms of lexical cohesion in the text. For instance, a cohesive relationship exists between what may be classified as the antonyms "less" in the expression "much less common" and "peak" in the phrase "the last peak" (Sentences 1 and 2). The lexical items cohere by virtue of the contrast they draw between indicating a drop in the occurrence of meningococcal disease on the one hand and a rise in the incidence of the disease on the other. Another example is provided by the cohesive relationship which is achieved by the use of the lexical items "national" and "England and Wales" both of which are synonymous in the text.

Lexical cohesion, while being the most predominant form of cohesion in the text is not the only means by which sentences in the text are made to hang together. Equally observable in Table 24 are instances of Reference Cohesion. Of particular interest to this study is the homophoric reference to "this century" in Sentence 4. The item does not tie with any specific item in the text, but it can be said to tie with a reality which exists as part of the reader's general knowledge about time reference and by means of which he can interpret the discourse in that sentence. Therefore, from the point of view of the reader, the item can be said to be cohesive. Equally interesting is the cohesive relation which exists between the pre-supposing item "(this) report" and the item which it presupposes — "(this) paper" in brackets in Sentence 10. But since the text is a report about research that was carried out, the antecedent of the reference item "this report" can only be the text itself in the preceding sentence on which the entire research is based.

Finally, no instances of conjunction, ellipsis and substitution were identified in the text.

The Methods Section

Passage:

Case definition — ¹A patient was designated a case when *Neisseria Meningitidis* was isolated from the blood or cerebrospinal fluid (CSF), or gram-negative diplococci were seen in the CSF, or clinical signs of meningitis or septicaemia were accompanied by haemorrhagic rash.

Case details — ²Cases were identified from statutory notifications and Public Health Laboratory (PHL) records. ³In addition, informal enquiries were made of adjacent health districts. ⁴General practitioners (GPs) were asked to report all cases known to them. ⁵The Hospital Activity Analysis (HAA) system was checked for all deaths or discharges with a diagnosis of meningococcal infection (International Classification of Disease, 9th revision, Code 036). ⁶Mortality statistics produced by the Office of Population Censuses and Surveys were checked for deaths due to meningococcal infection. ⁷Clinical features, complications, and outcome were obtained from hospital records.

Microbiology — ⁸Strains of meningococci isolated were grouped and tested for sulphonamide sensitivity at Gloucester PHL. ⁹From March 1983, serogroup and sulphonamide sensitivity results were confirmed at the PHL's Meningococcal Reference Laboratory (MRL), Manchester, where strains were also serotyped.

Table 25:
Sample Analysis of Cohesion in the Methods Section of the RAP

<u>Sentence No.</u>	<u>Presupposing Item</u>	<u>Cohesion Type</u>	<u>Presupposed Item</u>
1	case	L	case (see Introduction
2	cases	L	case
3	In addition	C	(sentence 2)
4	(all) cases	L	cases
5	all deaths and discharges	L	all cases
	meningococcal (infection)	L	(signs of) meningitis (S.1)
6	mortality (statistics)	L	(all) deaths (and discharges)
	meningococcal infection	L	meningococcal disease
7	Hospital (records)	L	Hospital (Activity Analysis)
8	meningococci	L	meningococcal infection
	sulphonamide sensitivity	L	sulphonamide-resistant meningococcus
	(Gloucester) PHL	L	PHL (records)
9	sulphonamide sensitivity	L	sulphonamide sensitivity
	the PHLs	R	(Gloucester) PHL
	strains	L	strains of meningococci
	Laboratory	L	Laboratory

The passage above describes the method of investigation adopted in the research. It focuses primarily on the "cases" or subject on which the study is based. Of course, the "cases" referred to in the passage are those suffering from meningococcal disease. Thus, two lexical items which play significant cohesive roles, particularly within the first few sentences of the passage, are "case/cases" and "meningococcal infection/meningococci". The constant repetition of these lexical items across sentences in the passage represents the type of general lexical cohesion relation referred to as lexical repetition. The other lexical item which enters into a similar relationship to those above is "sulphonamide sensitivity". It is interesting to observe that these three items which have a tendency to be repeated across sentences in the passage are very closely related to the topic of discourse in the text. Other forms of general lexical cohesion also exist in the passage. For instance, the lexical item "all deaths and discharges" in Sentence 5 coheres with "all cases" in Sentence 4 where the items "deaths and discharges" can be said to cohere with the item "cases" in a synonymous relationship. Similarly, the lexical item "mortality statistics" would seem to cohere with "all deaths and discharges" where the lexical item "mortality" relates to "deaths" also as a near-synonym.

Other types of cohesion found in the passage are Conjunction and Reference. The item "In addition" is a cohesive device which functions to extend the methods by which cases were identified from Sentence 2 to Sentence 3 in a relation of "positive addition". The only instance of Reference Cohesion is that between the item "the PHL's" (Sentence 9) and its referent "Gloucester PHL" (Sentence 8) with the definite article "the" in the presupposing item functioning to make the referent specific and identifiable. No other cohesion types apart from the three discussed here exist in the passage.

The Results Section

Passage:

¹65 cases of meningococcal infection were identified in the 4¹/₂ years to March 31, 1986. ²Only half the cases (33/65) were formally notified; the remainder were ascertained through PHL records. ³If the national notification rate for meningococcal meningitis had applied in Gloucester Health District during this period the number of notified cases would have been 12. ⁴The HAA and GP enquiries yielded no additional cases. ⁵Between 3 and 5 cases were reported in most quarters, with a peak of 12 in the last quarter of 1983 (Fig.1) ⁶Over the 4¹/₂ year period the number of cases recorded in the third quarter of a year (1.8) averaged less than half that seen in each of the remaining three quarters (4.2, 4.0, 4.2); this seasonal pattern is similar to that seen nationally.

⁷Of the three local authority districts, only Stroud and Gloucester had attack rates higher than expected (Table 1). ⁸The disease seemed to travel slowly through discrete communities within a district. ⁹The town of Nailsworth (population 5453) four miles south of Stroud was affected in 1982 and 1983. ¹⁰In 1983 and 1984 Stroud itself (population 20642) was the main focus of cases while in 1985 and 1986 Stonehouse (population 6612), situated three miles west of Stroud, had the highest attack rate. ¹¹Cases occurred steadily, albeit at a lower rate in the southern half of Gloucester City, throughout the study period. ¹²The only large village in the Stroud district free of the disease was Painswick (population 3163), four miles north of Stroud.

Table 26:

Sample Analysis of Cohesion in the Results Section of the RAP

<u>Sentence No.</u>	<u>Presupposing Item</u>	<u>Cohesion Type</u>	<u>Presupposed Item</u>
1	65 cases	L	cases (methods)
	meningococcal infection	L	meningococcal infection (see Methods)
2	<u>the</u> cases	R	65 cases
	cases	L	cases
	notified	L	notified (see Introduction)
	PHL records	L	PHL (see Methods)
	<u>the</u> remainder	R	65 cases

/continued

(continued)

3	national	L	England and Wales (see Introduction)
	notification	L	notified
	meningococcal meningitis	L	meningococcal infection
	Gloucester Health District	L	Gloucester Health District (see Methods)
	<u>this</u> period	R	4 ¹ / ₂ years to March 31 1986
	period	L	4 ¹ / ₂ years to March 31 1986
	notified (cases)	L	notified
4	<u>the</u> HAA and GP enquiries	R	enquiries (see Methods)
	HAA and GP	L	HAA: General Practitioners
	additional	R	cases
	cases	L	cases
6	the 4 ¹ / ₂ year period	R	4 ¹ / ₂ years to March 31 1986
	cases	L	cases
	nationally	L	England and Wales
7	Stroud	L	Stroud
	Gloucester	L	Gloucester
8	<u>the</u> disease	R	meningococcal disease
	disease	L	meningococcal disease
	communities	L	Nailsworth, Stroud, Stonehouse (cataphoric)
	(a) district	L	(local authority) districts
9	Town	L	Communities
	(south of) Stroud	L	Stroud
10	Stroud	L	(south of) Stroud
	cases	L	cases
	highest	R	attack rate
11	cases	L	cases
	(southern half of) Gloucester city	L	Gloucester
12	large village	L	communities
	Stroud district	L	Stroud
	<u>the</u> disease	R	meningococcal disease
	north of Stroud	L	south of Stroud

As the analysis in Table 26 shows, the passage is a logical continuation of discourse in the previous two passages. Thus a number of items in the initial segment of discourse in the passage tend to cohere by reference to their equivalent forms in either the Introduction or the Methods sections. For instance, the initial use of the item "cases" in this passage coheres with the instances in which the item was used in the Methods section. Similarly, the item "notified" in Sentence 2 ties with its initial mentioning in the Introduction section. The items "HAA" and "GP" also contribute to cohesion in the text by taking the reader backwards to the Methods section where they were first used in the text. These instances point to the fact that in the RAP, cohesion can occur between adjacent sentences as well as between sentences mediated by explicit text boundaries.

As is the case with the two previous passages already examined, lexical cohesion continues to dominate as the main form of cohesion in the text. Very high in this respect is the tendency for lexical items to get repeated using exactly the same wordings. Lexical items which get repeated in this form tend to be those already identified in the analysis of the Introduction and Methods sections (notified, cases, meningococcal infection, Gloucester, Stroud). There is also a noticeable increase in the tendency for sentences in the passage to cohere by means of equivalent lexical items. For instance, the lexical items "period" (Sentence 3) and "4¹/₂ years to March 31, 1986" in (Sentence 1) can be said to be equivalent in the context of discourse in the passage. The same also applies to the items "town/communities", "(large) village/communities", "communities/Nailsworth, Stroud, Stonehouse". In Halliday and Hasan's (1976) categorization, the equivalent item "communities" falls within the category of "general nouns".

A significant result of the analysis is the tendency towards a greater use of reference cohesion ties in this passage than in the previous two. However, as the analysis in Table 26 shows, most instances of reference cohesion in the passage are based on the use of the definite article to indicate that the referent of the item it modifies

is not only specific, but exists somewhere in the passage. Finally, sentences in the passage do not seem to cohere by means of ellipsis or substitution.

The Discussion Section

Passage:

¹Infections due to B:15 meningococci are becoming increasingly common in the UK, but this is the largest outbreak due to this organism yet described in the Country.

²Several features of the outbreak are of importance. ³First, there is no evidence, after $4\frac{1}{2}$ years, that the outbreak is waning. ⁴In Norway, over the 11-year period from 1974 to 1984, the average annual incidence of meningococcal disease remained raised (7.4 cases/100 000), with B:15R organisms predominating. ⁵Recent localised outbreaks due to group B strains in England and Wales have lasted from 8 months to $3\frac{1}{2}$ years.

⁶Secondly, the age-distribution of cases is unusual by comparison with recent national data (Fig.2). ⁷Meningococcal disease usually affects children under the age of 5 years, a reflection of the low antibody levels in this age-group. ⁸The high attack rate in teenagers and young adults by the B:15R strain of meningococcus has also been recorded in Norway but the absence of a peak in young children in the Gloucestershire outbreak is unexpected.

⁹Thirdly, throat swabbing revealed a very low carriage rate of B:15R meningococci in symptomless contacts. ¹⁰Similar low carriage rates of this organism were found in an extensive swabbing investigation in Somerset after a minor outbreak in 1985-86 (Bowie C, personal communication) and during investigation of a school outbreak in London (Cann K, unpublished).

The passage above is an attempt to discuss observations made in the Results section. The observations are discussed in relation to information already presented in the Introduction and Methods sections. Therefore, it is not surprising that the first few sentences in the passage cohere with sentences in the three previous sections already examined, in particular with sentences in the Introduction section. Thus lexical items such as "UK", "outbreak", " $4\frac{1}{2}$ years", "B:15R meningococcus" enable the text to hang together by relating them to instances in the preceding sections where they were first mentioned.

As in the other passages, lexical cohesion remains the dominant form of cohesion in this passage. Again lexical repetition appears to be the greatest form of

Table 27:
Sample Analysis of Cohesion in the Discussion Section of the RAP

<u>Sentence No.</u>	<u>Presupposing Item</u>	<u>Cohesion Type</u>	<u>Presupposed Item</u>
1	B:15 meningococci UK	L	meningococci (see Introduction)
		L	England and Wales (see Introduction)
2	this outbreak outbreak	R	this outbreak (see Introduction)
		L	outbreak (see Introduction)
3	First 4 ¹ / ₂ years	L	the largest outbreak
		C	several features
4	the outbreak outbreak	L	4 ¹ / ₂ years to March 31 1986 (see Results)
		R	meningococcal disease
5	meningococcal disease B:15R organism	L	outbreak
		L	meningococcal disease
6	outbreaks group B strains	L	B:15R meningococci
		L	outbreak
7	England and Wales secondly	L	group B strains (see Intro.)
		L	England and Wales (see Intro.)
8	cases recent	C	several features
		L	cases (see Results, Methods)
9	national meningococcal disease	L	recent
		L	UK, England and Wales
10	B:15R strain of meningococcus also	L	meningococcal disease
		C	B:15 R organism (sentence 7)
11	Norway young children	L	Norway
		L	young adults
12	Gloucestershire outbreak thirdly	L	Gloucester
		C	several features
13	B:15R meningococci low carriage rate	L	B:15R strain of meningococcus
		L	low carriage rate
14	this organism organism	L	meningococcus
		L	meningococcus
15	swabbing investigation (a minor) outbreak	L	throat swabbing
		L	(the largest) outbreak
16	similar	R	low carriage rate

cohesion employed in the passage. Items which are repeated in their exact wordings include most of those which have played similar roles in the preceding passages. They are the items "outbreak", "4¹/₂ years", "meningococcal disease", "England and Wales", "cases" etc. In addition to these, there are other items which are repeated because of the significance they have for the discourse in this section of the text. For instance, items such as "B:15R meningococci", "Norway" and "swabbing" get repeated because they are central to the argument being made in the Discussion section. The analysis also shows that as in the Results section, equivalent lexical items continue to be used cohesively, e.g. "B:15R organism/B:15R meningococci/B:15R strains"; "UK/England and Wales, nationally", etc. But the unique development with regards to lexical cohesion in the passages examined is the tendency for some sentences in this passage to cohere by means of antonyms, e.g. "a minor outbreak/the largest outbreak"; "young children/young adults".

Other cohesion types identified in the passage are Reference and Conjunction. In terms of Reference Cohesion, there is the use of the demonstrative pronoun "this" as an item to link sentences within the passage as well as across identifiable sections in the text. An instance of the latter is the demonstrative pronoun "this" in Sentence 1 which coheres with the item "this outbreak" in the Introduction section. An instance of the former is the presupposing item "this organism" which refers anaphorically to the presupposed item "B:15R meningococcus" (Sentence 8). Three Conjunction cohesion items are identified in the passage. All three items are identical in that they indicate the internal unfolding of the discourse. Each of them coheres with the item "several features" in the sentence "Several features of the outbreak are of importance". They therefore contribute to the organization of the passage into a coherent text. No substitution and ellipsis cohesion ties were identified in the passage.

In summary, three cohesion types — Reference, Conjunction and Lexical — are identified in the analysis of RAP.7 reported here. There is unsurprisingly

overwhelming evidence to suggest that the text coheres primarily by means of lexical cohesion, and in particular lexical repetition. No substitution and ellipsis ties were identified in the text.

6.5.2 Cohesion in the Abstract

An insight into the pattern of cohesion in the ABS is hereby provided by the analysis of cohesive elements realised in ABS.15. Below is the full text and a table of the various cohesion ties identified in it.

Abstract

Passage:

¹To evaluate the influence of the angiotensin-converting enzyme inhibitor enalapril (2.5 to 40 mg per day) on the prognosis of severe congestive heart failure (New York Heart Association [NYHA] Functional Class IV), we randomly assigned 253 patients in a double-blind study to receive either placebo (n = 126) or enalapril (n = 127). ²Conventional treatment for heart failure, including the use of other vasodilators was continued in both groups. ³Follow-up averaged 188 days (range, 1 day to 20 months). ⁴The crude mortality at the end of six months (primary end point) was 26 per cent in the enalapril group and 44 per cent in the placebo group — a reduction of 40 per cent (P = 0.001). ⁵By the end of the study, there had been 68 deaths in the placebo group and 50 in the enalapril group — a reduction of 27 per cent (P = 0.003). ⁶The entire reduction in total mortality was found to be among patients with progressive heart failure (a reduction of 50 per cent), whereas no difference was seen in the incidence of sudden cardiac death.

⁷A significant improvement in NYHA classification was observed in the enalapril group, together with a reduction in heart size and a reduced requirement for other medication for heart failure. ⁸The overall withdrawal rate was similar in both groups, but hypotension requiring withdrawal occurred in seven patients in the enalapril group and in no patients in the placebo group. ⁹After the initial dose of enalapril was reduced to 2.5 mg daily in high-risk patients, this side effect was less frequent.

¹⁰We conclude that the addition of enalapril to conventional therapy in patients with severe congestive heart failure can reduce mortality and improve symptoms. ¹¹The beneficial effect on mortality is due to a reduction in death from the progression of heart failure. (N. Engl. J Med 1987; 316:1429-35.)

Table 28: Sample Analysis of Cohesion in the ABS

Sentence No.	Presupposing Item	Cohesion Type	Presupposed Item
1	we	R	the Consensus Trial Study Group (see Title)
2	heart failure	L	congestive heart failure
	other vasodilators	L	enalapril
	both groups	R	either placebo (n = 126) or enalapril (n = 127)
3	follow-up	L	patients (S.1)
4	<u>the</u> (enalapril group)	R	enalapril group
	enalapril group	L	enalapril
5	<u>the</u> (placebo group)	R	placebo group
	placebo group	L	placebo
6	<u>the</u> study	R	double-blind study
	study	L	study
	deaths	L	mortality
	<u>the</u> placebo group	R	placebo group
	placebo group	L	placebo group
	<u>the</u> enalapril group	R	enalapril group
	enalapril group	L	enalapril group
	reduction	L	reduction
	mortality	L	mortality
	patients	L	patients
progressive heart failure	L	severe congestive heart failure (S.1)	
7	cardiac <u>death</u>	L	68 <u>deaths</u>
	NYHA classification	L	New York Heart Association (NYHA)
8	<u>the</u> enalapril group	R	enalapril group
	enalapril group	L	enalapril group
	reduction	L	reduction
	heart failure	L	congestive heart failure (S.1)
9	both groups	R	placebo group and enalapril group
	similar	R	(sentence 7)
	placebo group	L	placebo group
	enalapril group	L	enalapril group
	enalapril	L	enalapril
10	<u>this</u> side-effect	R	hypotension requiring withdrawal
	side-effect	L	hypotension requiring withdrawal
11	we	R	the Consensus Trial Study Group (see Title)
	enalapril	L	enalapril
	severe congestive heart failure	L	severe congestive heart failure
	mortality	L	mortality
11	mortality	L	mortality
	reduction	L	reduction
	death	L	death (S.5)
	progression of heart failure	L	progressive heart failure

The title of the text under which the Abstract analysed above was written is "Effects of Enalapril on Mortality in Severe Congestive Heart Failure". An interesting observation is that the text coheres primarily by the repetition of most of the lexical items in its title. First, the lexical item "enalapril" is introduced into the discourse in the text in Sentence 2. Then it is repeated without an accompanying modifying element in Sentences 9 and 10. In addition to this, it is also repeated as the NP "enalapril group" in Sentences 4, 5, 7 and 8. The second lexical item found in the title which is also repeated in identical wordings in the text is "Mortality". It is introduced into the discourse in Sentence 4, and thereafter is repeated in Sentences 6, 10 and 11. The third lexical item from the title which gets repeated in identical form is "congestive heart failure", first used in Sentence 1 and subsequently repeated in Sentences 2 and 10. Other lexical items which, though not in the title get repeated using the same words are: "placebo group" (S. 1, 4, 5, 8), "Patients" (S. 1, 6), "death" (S. 5, 6, 11), "Progressive heart failure" (S. 6, 11), "reduction" (S. 5, 6, 7, 11). Apart from these cases involving lexical repetition, there are also instances of lexical cohesion in which items cohere because they appear to be equivalent in the way they are used in the text. For example, "vasodilators" and "enalapril" where enalapril is itself a type of vasodilator; "side-effect" and "hypotension requiring withdrawal" where the latter presupposes the former. An instance of lexical cohesion by means of a synonym or near-synonym also exists in the text — "death" and "mortality".

Another type of cohesion identified in the text is Reference Cohesion. The predominant form of reference cohesion identified in the text is "demonstrative reference". Thus, the item "both" (S.2) serves to specify the referent of the presupposing NP "both groups" referring to "placebo (n = 126)" and "enalapril (n = 127)" in Sentence 1. The same form of reference is repeated in Sentence 8. Another demonstrative reference cohesion item identified in the text is the determiner "the" modifying the noun, as in the NP "the study", where the definite article as determiner serves to specify the presupposed referent of the NP "double blind study" (S.1). The

third demonstrative reference item in the text is "this" as in the NP "this side-effect" (S.9). The item refers to "hypotension requiring withdrawal" (S.8). There is also an instance of personal reference cohesion in the text. This is realized by the personal pronoun "we" which in the context of discourse in the text presupposes "the Consensus Trial Study Group" presented after the title to the text. Since the title and its accompanying elements can be possibly regarded as part of the text, a cohesive relation can be established between it and an element in the main text or part of it.

The result of this analysis shows that the two main cohesion types are employed in the text — Lexical cohesion and Reference cohesion. Of the two types, lexical cohesion is not surprisingly the predominant means of cohesion employed in the text. The analysis also shows that neither conjunction, substitution nor ellipsis are used as cohesive devices in the text. Thus, the observations made in this analysis are similar in some respects to those made in the analysis of cohesion in RAP.7 (see Tables 24, 25, 26, 27). The analysis of both sample texts shows that lexical cohesion is the predominant form of cohesion employed in them. Both sample texts also seem to favour the repetition of lexical items, especially those related to the topic of discourse. They also make use of other forms of lexical cohesion such as synonyms, equivalence and antonyms. A look at the analysis of the sample texts from both genres (ABS and RAP) shows that they make use of reference cohesion. There is also evidence that they tend to make more use of demonstrative reference cohesion items than either the personal or comparative types. Finally, neither sample texts coheres by means of ellipsis or substitution cohesion. On the minus side, some differences may be noticed in the realization of cohesion types and items in the sample texts analysed from the RAP and the ABS. First, there is the absence of conjunction cohesion items in the ABS text. There is also the absence of "naming" lexical items in the ABS.

The similarities observed in the analysis of the sample texts from both genres in the preceding discussions would seem to suggest that the analysis of cohesion does not reveal anything striking or interesting in the way texts in both genres use the resources

of cohesion. However, an analysis based only one text from a whole corpus may not provide sufficient evidence to enable any meaningful generalization. On the other hand, it could provide an insight into what to expect in the rest of the texts in the corpus. It is with regards to this latter purpose that the observations made in the sample texts are both striking and interesting. Therefore, armed with the insight provided by those observations, I proceeded to analyse texts in the rest of the RAP and ABS texts in order to determine how typical the observations revealed in the sample texts are of the nature of cohesion in the genres.

...ing breast cancer before age 45 does not alter significantly for women who started using oral contraceptives (OCs) before they were 25 and went on doing so for more than four years. ²That is the main conclusion of an American study — the largest of its kind to date. ³In contrast to an earlier investigation, which raised considerable concern (New Scientist, 27 October 1983, p.264).

...⁴While David and his colleagues in the Cancer and Steroid Hormone (CASH) Study Group of the Atlanta Center for Disease Control look only at OC use by 2000 women, diagnosed with breast cancer in the two years up to December 1982, with data by 2005 women without the disease (The Lancet, Vol II, p.970).

...⁵The researchers also found no change in breast cancer risk in women who used high-pregnancy "postmenopausal" OCs before the age of 25 but stopped within the next six years. ⁶The earlier study had also made this suggestion, which led the Committee on Safety of Medicines to advise age to issue a list of pills and their potency, with advice to doctors to prescribe only low-pregnancy OCs.

...⁷Examining all these contradictory findings which must confuse "scientists, family planners and general practitioners", The Lancet suggests that they may be reconcilable, on the assumption that the earlier study in fact reflects the fact that young women in California may have used the pill more than elsewhere in the United States.

...⁸If long-term OC use in early life does affect the risk, the leader says, "it might show itself in terms of excess cases 10 or 20 years after exposure". ⁹In support of this possibility, the journal cites other surveys — for example, young Japanese women exposed to the atomic bombs in Hiroshima and Nagasaki, who only showed an increase in oestrogen-related breast cancer 15 years after their exposure.

...¹⁰The article goes on to calculate the probable time course of the effect of a threefold increase in risk, but only after 10 years — following four years or more of OC use before the age of 25. ¹¹The assumption is that there is a 10-year latent period before the risk becomes detectable.

...¹²In Britain, far less than five per cent of breast cancer cases would have been attributable to OC use in 1980, compared with perhaps 10 per cent in 1985 and 20 per cent in 1990. ¹³In fact, the Royal College of General Practitioners in 1981 found such an increased risk, though it was of marginal statistical significance.

6.5.3 Cohesion in the Journalistic Reported Version (JRV)

An indication of the pattern of cohesion in the JRV is provided in the analysis of text JRV.2 presented below. As usual, the text is first written out. This is followed by a detailed analysis of cohesion and a discussion of the various cohesion types realized in the text.

Passage:

Pill use and risk of breast cancer

¹The risk of getting breast cancer before age 45 does not alter significantly for women who started using oral contraceptives (OCs) before they were 20 and went on doing so for more than four years.

²That is the main conclusion of an American study — the largest of its kind to date. ³It contradicts an earlier investigation, which caused considerable concern (*New Scientist*, 27 October 1983, p.264).

⁴Bruce Stadel and his colleagues in the Cancer and Steroid Hormone (CASH) Study Group of the Atlanta Center for Disease Control compared OC use by 2088 women, diagnosed with breast cancer in the two years up to December 1982, with that by 2065 controls selected at the same time (*The Lancet*, Vol II, p.970).

⁵The researchers also found no change in breast cancer risk in women who used high progestogen "potency" OCs before the age of 25 and went on with them for more than six years. ⁶The earlier study had also made this suggestion, which led our Committee on Safety of Medicines two years ago to issue a list of pills and their potency, with advice to doctors to prescribe only low progestogen OCs.

⁷Commenting on these contradictory findings which must confuse "scientists, family planners and general practitioners", *The Lancet* suggests that they may be reconcilable, on the assumption that the earlier study in part reflects the fact that young women in California may have used the pill sooner than elsewhere in the United States.

⁸"If long-term OC use in early life does affect the risk", the leader says, "it might show itself in terms of excess cases 10 or 20 years after exposure". ⁹In support of this possibility, the journal cites other surveys — for example, young Japanese women exposed to the atomic bombs in Hiroshima and Nagasaki, who only showed an increase in radiation-related breast cancer 15 years after their exposure.

¹⁰The article goes on to calculate the probable time course of the effect of a threefold increase in risk, but only after 10 years — following four years or more of OC use before the age of 25. ¹¹The assumption is that there is a 10-year latent period before the risk becomes detectable.

¹²In Britain, far less than five per cent of breast cancer cases would have been attributable to OC use in 1980, compared with perhaps 10 per cent in 1985 and 20 per cent in 1990. ¹³In fact, the Royal College of General Practitioners in 1981 found such an increased risk, though it was of marginal statistical significance.

Table 29: Sample Analysis of Cohesion in the JRV

Sentence No.	Presupposing Item	Cohesion Type	Presupposed Item
1	breast cancer	L	breast cancer (Title)
	oral contraceptives	L	Pill (Title)
2	that	R	the risk of getting breast cancer before age
3	it	R	an American study
	investigation	L	study
4	OC use	L	oral contraceptives (S.1)
	breast cancer	L	breast cancer (S.1)
5	<u>the</u> researchers	R	Bruce Stadel and Colleagues
	researchers	L	Bruce Stadel and his colleagues in the Cancer and
	also	C	(sentence 4)
	breast cancer risk	L	the risk of getting breast cancer
6	<u>the</u> early study	R	an earlier investigation
	early study	L	earlier investigation
	also	C	(sentence 5)
	<u>this</u> suggestion	R	no change in breast cancer risk in women who used high progestogen
	suggestion	L	no change in breast
	pills	L	Pill (Title)
	low progestogen	L	high progestogen "potency" OCs
7	<u>these</u> contradictory findings	R	(1) the researchers also found.. (2) the earlier study had also made this suggestion ...
		L found that
	<u>the</u> earlier study	R	an earlier investigation
	earlier	R	early study
	earlier study	L	early investigation
	young women	L	women who used high progestogen "potency" OCs before the age of 25
	<u>the</u> pill	R	high progestogen "potency" OCs
	pill	L	pill
	United States	L	American

/continued

(Table 29 continued)

8	OC use early life <u>the leader</u>	L L R	OC use before they were 20 (S.1) The Lancet
9	<u>this possibility</u> possibility <u>the journal</u> journal	R L R L	it might show itself it might show itself The Lancet The Lancet
10	the article threefold increase in <u>risk</u> OC use	L L L	an American study the <u>risk</u> of OC use
11	<u>10 years</u> latent period the risk	L L	after 10 years increase in <u>risk</u>
12	in <u>Britain</u> breast cancer cases OC use	L L L	<u>American</u> study breast cancer OC use
13	such an increase	R	far less than 5 per cent in 1988, 10 per cent in 1985, 20 per cent in 1990

As with other sample texts so far analysed in this Section, the sample text analysed above seems to cohere mainly by means of lexical cohesion ties. Four forms of lexical cohesion may be identified in the text. They are lexical repetition, the use of equivalent lexical items, synonyms and antonyms. The first form is exemplified by the repetition of the lexical items "breast cancer" (Sentences 1, 4, 5 and 12); "oral contraceptives" (Sentences 1, 4, 8, 10 and 12) and "pill/pills" (Sentences 6 and 7). The second type, which is the most widespread in the text is represented by the cohesive relation which holds between the items:

- a) oral contraceptives/pills
- b) researchers/Bruce Stadel and his colleagues
- c) suggestion/... no change in the breast cancer risk in women who used high progestogen "potency" OCs before age 25 and went on with them for more than six years
- d) contradictory findings/the researchers also found
the early study also made this suggestion

- e) (young) women/women (who used) before the age of 25
- f) early life/before they were 20
- g) possibility/.... it might show itself in

What these items have in common is that they refer to the same entity not that they are lexically related. Thus, some of them would seem to belong to the category of Reference cohesion in Halliday and Hasan's model, e.g. "researchers/Bruce Stadel and his colleagues", but even then they are not normal examples of Reference in Halliday and Hasan's cohesion model. Others, e.g. "suggestion/...no change in the breast cancer ...", and "contradictory findings/the researchers also found" and "possibility/....it might show itself in" would seem to belong to the category of "general nouns or words". However, Hasan's (1984) modification of the category of lexical cohesion in Halliday and Hasan (1976) would classify these items as "Instantial" lexical cohesion items whose relation with each other is that of "equivalence". It is, therefore, in accordance with this latter classification that the items are grouped as lexical cohesion items in this study.

The third form of lexical cohesion represented in the text is achieved by means of synonyms as in the relations which exist between the lexical items "investigation" (Sentence 3) and "study" in the NP "American study" (Sentence 2). A similar relation exists between the "naming" lexical item "United States" (Sentence 7) and the item "American" in the phrase "American Study" (Sentence 2). The final form of lexical cohesion identified in the text — antonyms — is characterized by the relation which exists between the lexical items "low " (Sentence 6) and "high " (Sentence 5) in the NP's "low progestogen" and "high progestogen" resectively.

Another category of cohesion identified in the passage above is Reference cohesion. Most of the reference cohesion items identified in the text are those which are used to indicate demonstrative reference. They include the items "that" (Sentence 2) which refers anaphorically to the whole of the propositions in Sentence 1; the item

"it" (Sentence 3) also an anaphoric reference to the lexical item "American Study" (Sentence 2); the definite article "the" serving to specify the reference of the presupposing items "researchers" (Sentence 5) "study" (Sentences 6 and 7) and "journal" (Sentence 9). Other items in this category are the demonstrative pronouns "this" (Sentences 6 and 9) in which it is used to refer to the propositions in Sentences 5 and 8 respectively, and the demonstrative pronoun "these" (Sentence 7) referring anaphorically to propositions in Sentences 5 and 6. Also represented in the text is an instance of comparative reference. It is represented in the cohesive relation which is established between Sentence 13 and Sentence 12 by means of the item "such" in the expression "such an increase" which indicates a comparison between the observation made by the "Royal College of General Practitioners in 1981" with the increased risk discussed in Sentence 12.

A third category of cohesion identified in the text is that of "conjunction". It is represented by the lexical item "also" (Sentences 5 and 6). As a conjunction cohesion item, it links Sentence 5 to Sentence 4. Its function is to indicate that the observations made by the researchers is another point to be considered in conjunction with information provided in the previous sentence.

Finally, no substitution and ellipsis cohesion items were identified in the text.

In terms of types of cohesion realized in the text, the JRV sample text analysed above shares a number of features in common with the sample texts analysed from the ABS and the RAP. As in both sample texts, and as in nearly all texts, this sample text coheres primarily by means of lexical cohesion. However, unlike in the other genres, lexical repetition is not the main method of lexical cohesion employed in the text. Again, as in both sample texts, instances of reference and conjunction cohesion may be found in this text, while ellipsis and substitution cohesion types are conspicuously absent. But it must be pointed out that the NP "...the largest of its kind", in which "study" is ellipled, can arguably be seen as a case of cohesive ellipsis, though it is intrasentential (sort of) - if a dash counts as keeping the item within a sentence. Given

the lack of conventional punctuation it could be seen as either an NP in apposition or a highly elliptical sentence. However, the latter is somewhat implausibly so. Therefore, it would appear that all three sample texts demonstrate fairly close patterns in the realization of cohesion types. But is this same pattern sustained in the analysis of all texts in the corpus? Assuming they are, does the role of each cohesion type and the items which realize them differ in each genre in line with differences in the purpose of discourse in each genre? Answers to these and other related questions will be provided in the next section on "Cross-genre differences".

6.6 CROSS-GENRE COMPARISON

This discussion of the differences in cohesion between texts in the three genres is based on evidence from an analysis of all texts in the corpus along the same lines as those reported for the sample texts in Section 6.5 above.

The analysis of sample texts in Section 6.5 shows that the three texts from the RAP, ABS and JRV respectively do have much in common with regard to the cohesion types they tend to use. Following this observation, all texts in the corpus were analysed using the same procedure as that employed in the analysis of the sample texts. The results obtained from that analysis would seem to support the observations made in the sample texts. That is, that the genres are related in terms of the cohesion types they characteristically tend to use. Therefore, in the discussion that follows, we shall begin by examining these general similarities before going on to discuss the differences. Since the discussion is qualitative, we shall be concerned with similarities and differences in the way cohesion types and their sub-categories are realized and in the function which cohesion types perform across the genres. These similarities and differences will be illustrated with exemplifications from texts in the corpus.

Just as was observed in the analysis of the sample texts (Section 6.5), the analysis of texts in the entire corpus shows that three main cohesion types tend to

feature as cohesive devices in all three genres. The cohesion types are: Reference, Conjunction and Lexical. In other words, Substitution and Ellipsis cohesion types do not feature or at least feature only minimally as cohesive devices in texts in all three genres. Of the three main cohesion types identified in the analysis, lexical cohesion tends to predominate over the others in all three genres. Again, this is consistent with the observation made in the analysis of the sample texts referred to above. This observation is understandable because it is reasonable to expect that in technical texts, such as those which constitute the corpus of this study, information will almost invariably be imparted through the consistent use and repetition of technical and specialist vocabularies which would have the effect of binding sentences together. It is also reasonable to expect that in all kinds of texts, the choice of lexical items will be constrained by purpose, topic and audience. The repetition of lexical items will, therefore, form a major means by which sentences will be linked together.

Just as the observation regarding the high occurrence of lexical cohesion in the genres is understandable, so also is that regarding the non-occurrence of Substitution and Ellipsis cohesion types. Both cohesion types are known to be more common in spoken than in written English, and even where they do occur in writing, Ellipsis is often preferred to Substitution (Halliday and Hasan, 1976:199, Witte and Faigley, 1981: 190). Therefore, the non-occurrence of Ellipsis and Substitution cohesion types in the three genres appears to be consistent with these views.

The contentions made in the preceding paragraphs raise one fundamental question — "If the genres are similar with respect to the cohesion types they characteristically employ, in what respects, then, are they different?" This question can only be answered by a detailed examination of the sub-categories of each cohesion type that is realized in each genre for evidence of how and why they are realized. The order of discussion is Reference, Conjunction and lastly Lexical cohesion.

6.6.1 Differences in Reference Cohesion

In functional terms, the aspect of Reference cohesion which has direct influence on the overall purpose of discourse in each of the genres is "Personal Reference". Personal Reference items identified in the analysis contribute to cohesion by:

1. referring to authors only
2. referring to authors and other persons
3. referring to other researchers (citations)
4. referring to aspects of research subject-matter
5. referring to research participants

By far the commonest personal reference item observed in the analysis are those which are used to refer to authors, especially of the RAP. This observation sounds somehow unusual considering the fact that it is often thought that the requirement of impersonality constrains writers of scientific research papers from making reference to themselves in the text. As Myers (1989:4) points out:

Of course in reality scientists have a network of informal contacts, collaborations, and long-standing personal commitments that do not require great social distance. But none of this is to emerge in print; there are no first names, and one can refer to oneself in the third person.

This may probably be due, as Myers (op cit: 14) points out to the fact that scientific knowledge is supposed to be taken as universal; therefore any implication that a belief is a personal belief weakens it. But, as the analysis here shows, there are situations when such requirements do not always apply. In those situations, the personal reference items contribute to cohesion by linking the sentences in which they occur with a reality in the text or in the reader's world knowledge. In the RAP, the items by which personal reference cohesion is realised are the first person plural pronoun "we" and the possessive pronoun "our". Observations from the RAP texts

analysed in this study suggest that as far as making reference to writers is concerned, the pronoun "we" which was identified in 50 instances is used more often than the pronoun "our" with 25 occurrences. Instances of the use of the items in the RAP are shown below:

As part of a large study of human papilloma-virus (HPV) and herpes simplex virus (HSV) genomes in persons with genital cancers in Western Scotland and northwest England, we conducted a small investigation to determine whether HPV genomes were present in clinically and histologically normal paired, matched control tissues from patients with cancer.

We found that the cloned HPV-16 DNA probe hybridized to sequences in 8 of 11 clinically and histologically normal tissue samples from sites adjacent to those at which tissue was excised by radical surgical treatment.

(RAP.10: *The NEJM*, October 23, 1986)

In a study on the predictive value of the AST in a group of European primigravidae, 45% of women with a positive test at 28-32 weeks went on to have PIH and pre-eclampsia later in pregnancy. This percentage accords with the 52% of angiotensin-sensitive women in our placebo group in whom PIH and pre-eclampsia eventually developed.

(RAP.3: *The Lancet*, January 4, 1986)

In the examples above, the pronouns "we" and "our" cohere with the names of the writers of the research article listed after the title at the top of the text. In addition to this, the possessive pronoun "our" coheres with any prior use of the equivalent item "we" in the preceding text. In Halliday and Hasan (1976: 50) these items which refer to the speech roles of writers are said to be typically exophoric. However, as Halliday and Hasan (op cit) point out, they can also be anaphoric in many varieties of written English, particularly in those texts in which the "context of situation" includes "a context of reference". While the RAP cannot be said to belong squarely to this group of written texts, its use of the personal pronoun "we" and the possessive "our" is certainly endophoric and cohesive because the title/name of writers usually at the top of the text is arguably part of the context of discourse in RAP texts. Therefore, in

modelling cohesion from the point of view of the reader, items in the title, including the names of writers contribute to textual cohesion.

As in the RAP, the personal reference items which contribute to cohesion in the ABS are the personal pronoun "we" and the possessive pronoun "our". And as in the RAP, the pronoun "we", with 3 occurrences in the data, is used more often than "our" with only one identifiable instance of use in the data. Below are examples of instances of the use of both items in the ABS.

To evaluate the influence on the angiotensin-converting enzyme inhibitor enalapril (2.5 to 40 mg per day) on the prognosis of severe congestive heart failure (New York Heart Association [NYHA] Functional Class IV), we randomly assigned 253 patients in a double-blind study to receive either placebo (n = 126) or enalapril (n = 127).
(ABS.15: *The NEJM*, June 4, 1987)

Our results show a strong association between the presence of HPV-16 genomes and genital tumors and between HPV-16 genomes and histologically normal tissue within 2 to 5 cm of the tumors.
(ABS.10: *The NEJM*, October 23, 1986)

The two reference cohesion items in the texts above cohere with the names of the writers of the articles, usually at the top of the text. The reasons given for why these items are considered to cohere with authors' names in the case of the RAP also applies here. As has already been pointed out elsewhere in this Thesis, texts in both genres share a common title and are written by the same authors.

The discussion so far shows that there is a correspondence between the RAP and the ABS in terms of the personal reference cohesion items used to refer to writers in their texts. This relationship is made more obvious by the fact that texts in both genres share the same title and are written by the same authors. However, as we have seen, the use of these items tends to be more prevalent in the RAP than in the ABS.

JRV writers also tend to make reference to writers of the research articles they are popularizing. However, the items by which the references are made differ from those already identified in the RAP and the ABS. Thus, the analysis shows that two types of reference cohesion items are used to make reference to writers of RAP texts in

their popularized JRV equivalents. The items are: the third person pronoun "they" and the possessive pronoun "their". Instances of the use of these two personal reference items in the JRV are shown below:

Young's team is now growing squamous epithelial cells in tissue cultures and infecting them with EBV virus. They describe their results as encouraging. (JRV,4: *New Scientist*, 6 March 1986)

The research by a group of workers with Dr Gerald Berenson of the Louisiana Medical Center, New Orleans is described in the current "New England Journal of Medicine". Their results identified the type of fat and oily molecules to be found most frequently in the offspring of parents with heart disease. (JRV.8: *The Times* 1 October 1986)

The significance of the observations made regarding the differences between the RAP/ABS and the JRV does not lie so much in the fact that both genres make use of different personal reference cohesion items. Rather, it lies in how and why the use of these personal reference cohesion items reflect constraints imposed on the writer by the purpose, audience and medium of discourse in these genres. The item which best illustrates these constraints is the personal pronoun "we" as used in the RAP and the ABS.

As has been pointed out, in the RAP and the ABS, the first person plural pronoun "we" coheres with the names of the authors of texts in both genres, which are invariably listed at the top of texts. It is a truism that scientific (including medical) research papers display impersonality and objectivity in the presentation of research events. In other words, in reporting the results of a research investigation, the actions, judgments and choices which the writers make as researchers are not expected to form part of the final report on the investigation. Similarly, emotive statements which may be construed as furthering the image of the writer at the expense of the empirical aspect of the work being reported are completely avoided by the writer. As has been pointed out in Chapter 4, these characteristic features of scientific research reporting reflects

what Gilbert and Mulkay refer to as the "empiricist repertoire" in which scientists attempt to depict their actions and beliefs as a natural medium through which empirical phenomena make themselves evident. Despite this fact, writers of research articles and their accompanying abstracts still make reference to themselves by means of what Fowler and Kress (1979) refer to as the "exclusive we" or Halliday and Hasan (1976: 53) as the "impersonal we".

A number of reasons may account for the tendency for authors of RAP and ABS texts to refer to themselves collectively by means of the impersonal or exclusive "we" in violation, as it were, of the "requirement" of impersonality often associated with the reporting of scientific research investigations written for a professional audience. One of the reasons for this may relate to the observation by Halliday and Hasan (op cit) to the effect that the "impersonal we" is often used to overcome the tiresome effect of the use of passives in discourse. Following from this, it can be said that the use of the pronoun "we" in the RAP and the ABS is constrained by the need for authors of texts in these genres to occasionally vary their tense forms from the characteristic passive form of scientific discourse (Cooray, 1965; Swales, 1976) to an active form. Such variation, it can be argued, makes the text less boring and tiresome for the reader. Thus, it may be contended that in addition to providing a means by which the text hangs together as text, the first person plural active "we" is used in the RAP and BAS to mitigate the probable negative effect which the persistent use of the passive can have on readers of texts in these genres.

Another factor which constrains the use of the pronoun "we" in the RAP and the ABS is the need to indicate to the reader when a shift has occurred from one discourse segment to another. This function is best understood if viewed against the background of the concept of "Moves" proposed in Chapter 4. Thus, the pronoun "we" in association with other elements in the sentence serves to indicate the commencement of a new discourse segment or a shift in argumentation, as may be seen from the example below:

1. Recurrent spontaneous abortion has been treated by means of immunisation with paternal cells and cells from multiple donors. Organ allograft rejection is diminished by previous immunisation with blood transfusion, and the majority of women in the spontaneous abortion group unlike normal multiparous women did not have detectable cytotoxic antibodies against paternal lymphocytes.
2. Having achieved good results with our preliminary attempts at immunisation, we decided to carry out a controlled trial of treatment with paternal cells. We wanted to find out whether the success was truly

(RAP.1: *The Lancet* 27 April 1985)

Two segments of information or "moves" may be identified in the text above. The first segment is an attempt to provide some background information on the subject-matter of discourse in the entire text — i.e. "Recurrent spontaneous abortion". The second segment beginning with the sentence "Having achieved good results with our preliminary attempts" is concerned with presenting the new research by making reference to the writers and the purpose of the study. The pronoun "we" and the verb phrase "decided to carry out" are items which contribute to indicate a shift in orientation from presenting background information to presenting the new research. Such indicators or signals are known to aid discourse processing by readers. Following from all the above, it may be contended that some instances of the use of the pronoun "we" in the ABS and the RAP is constrained by the need for writers to signal to readers shifts in the structure of discourse or the commencement of new segments of discourse.

A third factor which constrains the use of the pronoun "we" in the RAP and the ABS is the need to indicate to the reader points in the logical development of the argument where the writer makes a unique choice (Tarone et al, 1985: 195). This function is related to the one discussed in the preceding paragraph, except that in this case the pronoun "we" indicates a shift in argumentation within a "Move" or discourse segment rather than between segments, as can be seen in the text below:

¹There is increasing evidence that the psychological vasodilation and low vascular resistance to flow in pregnancy depend on production of biologically equivalent amounts of vasodilator prostacyclin (PGI₂) and vasoconstrictor TXA₂ in the foetal and maternal circulation. ²In PIH and pre-eclampsia, production of these two eicosanoids is tilted heavily in favour of TXA₂; enhanced sensitivity to angiotensin II may be the first sign of such an imbalance. ³Low-dose aspirin as used in our study suppresses platelet TXA₂ synthesis, as indicated by the reduction of thrombin-induced formation of malondialdehyde to approximately 100% of placebo values, whereas the effect on vascular PGI₂ production is said to be negligible. ⁴However, we did not measure plasma concentrations or urinary excretion of prostacyclin metabolites. ⁵Thus we can only speculate that the low-dose aspirin used in our trial prevented PIH and pre-eclampsia by

(RAP.3: *The Lancet*, 4 January, 1986)

The text above is a discussion of an observation made during the study. First, the writer makes reference to what is already known about the two eicosanoids, prostacyclin (PGI₂) and TXA₂ in pregnancy. This functions as a form of background information which lends weight to the observation made in the study — i.e. that low-dose aspirin suppresses platelet TXA₂ synthesis. The use of the possessive pronoun "our" to modify the noun "study" in the sentence, "Low-dose aspirin as used in our study suppresses" (Sentence 3) marks out the result as one which the writers consider unique to their study, and therefore, distinct from previous results on the subject. In Sentence 4, the writers point out a unique choice they made in the investigation — they did not measure plasma concentrations or urinary excretion of prostacyclin concentrates. Following this, they arrive at a unique conclusion — "that low-dose aspirin used in our trial prevented PIH and pre-eclampsia by". In each of these situations, the personal pronoun "we" is used as subject to refer to the agents who carried out the action. By making overt reference to themselves in these contexts, the authors would seem to be suggesting that they are committed to these actions and choices and, therefore, are willing to take responsibility for their accuracy or otherwise. Thus, in the RAP in particular, the personal pronoun "we" is both a reference cohesion item as well as an item which serves to indicate points in the text at which the writer makes a unique choice.

In the JRV, as we have seen, two items — the personal pronoun "they" and the possessive pronoun "their" are used to refer to authors of RAP texts being popularized. Again, this raises the question of the extent to which the use of these reference cohesion items in the JRV reflects the influence of the purpose, audience and medium of discourse on writers. Answers to this question cannot be provided in isolation from the use of personal names as a cohesive property of discourse in the JRV. The discussion of the role of personal names as cohesive features of text in the JRV cannot be made in this Section, since personal names belong to the category of Lexical cohesion (Halliday and Hasan, 1976: 341). Suffice it to say here that the use of the personal pronoun item "they" and the possessive pronoun item "their" in the JRV is a logical consequence of the use of personal names in texts in the genre to refer to authors of RAP texts very often in their capacity as researchers rather than as authors.

The analysis also shows that personal reference cohesion items may be used to refer to writers of the texts and other persons, presumably the reader. This form of reference is unique only to the JRV and even then its occurrence in the genre is highly limited. Two items — the personal pronoun "we" and the possessive pronoun "our" are used to realise this function in the JRV. It has been pointed out (Myers, 1989: 7) that one of the functions of the personal pronoun "we" in research papers is as a positive politeness device to mitigate claims and denials of claims. Citing an example from a research paper in molecular biology, Myers (op.cit) contends that "one way of making a criticism while minimizing the FTA (Face Threatening Act) is for writers to use pronouns that include themselves in the criticism." While no such pronouns were identified in my RAP texts, probably due to the data size, a few of them were identified in the JRV. Therefore, following from Myers (op.cit), it can be said that the pronoun "we" as used in the JRV to include writers of the texts and others is a positive politeness device which functions to minimize the FTA where a criticism is being made. In line with the above contention it can be seen that the role of the pronouns

"we" and "our" in the JRV differs from those identified in the RAP and ABS in this study. The sample texts below serve to illustrate the function of these pronouns as discussed above.

¹Unfortunately, we still await the development of an effective drug to block the actions of bradykinin and laslyl-bradykinin. (JRV.3)

²Americans have higher average levels of linoleic acid in their body fat: 16.3 per cent compared with 9.8 per cent in our population. (JRV.12)

In the first example, the writer is criticizing the fact that no effective drug has been developed to block the the actions of bradykinin and lasyl-bradykinin. In the second example the writer is criticising the low level of linoleic acid in the British population. In both examples, the writers minimize the FTA which the criticisms may have on the organizations directly involved in the activities by including themselves in the criticisms. Apart from minimizing the FTA in both examples, the pronouns "we" and "our" are also intended to have a persuasive impact on the JRV reader. The use of the pronoun "we" in particular relates to what Fowler and Kress (1979) refer to as the "inclusive we" which they contend implicates the addressee in the content of the discourse and "is, therefore, ostensibly more intimate and solidary." In the context of the examples above, and from the point of view of the decoder, the personal pronoun "we" as used in the JRV can be said to function to bring referents together in highlighting the negative implications which the results obtained in the research being popularized might have for them. Therefore, it can be argued that the use of the "inclusive we" in the JRV is an attempt by the writer to persuade the reader to share his point of view about the results of the research being popularized.

The third role which personal reference items tend to enter into in texts examined in the corpus is making reference to other researchers, otherwise referred to in the literature as citations. The analysis shows that the genres differ in the extent to which they make use of citations. The most obvious difference in this regard is the complete absence of citation in the ABS text. This means that citations can be eliminated from summaries of research articles without doing damage to the main issues discussed in the texts. In contrast, the RAP and the JRV tend to cite the work of other researchers in their texts. However, the genres differ in the extent to which they make reference to other researchers and in the purpose which such references serve. With regard to the former, the analysis shows that the practice of referring to other researchers is not as common in the JRV as it is in the RAP. An examination of JRV texts shows that very few instances of explicit reference to other researchers are made, and even then, only one of such cases appears to be cohesive. That instance is shown in the text below:

In the same issue of B.M.J., I.T. Mangonda and Colleagues at Leicester General Hospital report the results of a study of patients with kidney cancer. They suggest that although blood transfusions appeared to improve their chances of survival, this was really because their cancer was treated at an earlier stage.

(JRV.6: *New Scientist*, September, 1986)

The observation made in the JRV contrasts with that made in the RAP where every text examined in the genre had more than one instance of reference to other researchers. The JRV and the RAP also differ with regard to the purpose which reference to other researchers serve in their texts. Myers (1989: 6) points out that citations are used in scientific research papers to either support, deny or supersede claims made by writers about their research. He also points out that they are used to hedge claims, make criticisms and avoid "Face Threatening Acts" (FTAs) (Brown and Levinson (1987). Most of these functions were identified in the RAP texts examined in this study. Unlike in the RAP, "reference to other researchers" serves only one purpose in the JRV. It functions to reinforce rather than dismiss or deny claims made by RAP writers about their research. Thus "reference to other researchers" can be said to have a persuasive impact in the JRV.

Other roles of personal reference cohesion items identified in the corpus include reference to aspects of subject-matter and reference to participants. There appears to be no distinctive difference either in the realization of these items or in the functions which the items which realize these purposes perform across the genres.

Finally, as indicated at the beginning of this Section, there appears to be no difference in the realization and function of Demonstrative and Comparative reference cohesion items across the three genres. With regards to Demonstrative reference, it was observed that all three genres tend to make use of the demonstratives "this", "that", "those" and the non-selective demonstrative "the" for reference purposes. Demonstratives, particularly "this" and "that" are known to be used in discourse to refer to events, actions and propositions rather than to physical objects (Huddleston 1984: 296). All three genres examined in this study are primarily concerned with the description of events, actions and propositions. In view of this it is understandable that they all tend to make use of these demonstrative items for reference purposes.

Again, demonstratives are known to have descriptive and contrastive functions (Kamp, 1981). The communication of the results of a research investigation in the form of either a research paper, an abstract or science popularization involves the description and contrasting of events, actions and results. The texts examined in this study exhibit all of these characteristics. Thirdly, demonstratives function to direct the hearer's focus of attention to a new reference object that was not previously in focus (Pinkal, 1986). Related to this is the observation made in this study to the effect that demonstrative reference items tend to be used to sum up propositions in immediate referential distance. For example:

¹This parameter estimate indicates that a higher serum cholesterol level of 20 mg/dl (0.52 mmol/L) is associated with relative risk of 1.17 (95% confidence interval, 1.15 to 1.19). ²Since the mean serum cholesterol level for the entire cohort was 214.6 mg/dl (5.55 mmol/L), this translates into an estimate that a 9% higher serum cholesterol level was associated with a 7% greater CHD death rate — that is, a 1% higher serum cholesterol level was associated with an almost 2% higher CHD risk.

(RAP.11: *JAMA*, November 28, 1986)

In the example above, the demonstrative item "this" (Sentence 2) refers anaphorically to the whole of the propositions contained in the first sentence. Therefore, it can be said to function to sum up propositions contained in sentences which are in immediate referential distance. The same function is performed by the demonstrative "that" in the text below from the JRV. Examples also exist in the ABS to illustrate this.

The risk of getting breast cancer before age 45 does not alter significantly for women who started using oral contraceptives (OCs) before they were 20 and went on doing so for more than four years. That is the main conclusion of an American study — the largest of its kind to date.

(JRV.2: *New Scientist*, 7 November 1985)

An examination of the texts for comparative reference items shows that cohesion in the genres does not depend very much on comparative reference despite the fact that comparison is a major feature of discourse in the genres. This contention is based on the observation that most of the instances of comparison identified in the texts were intra-sentential, and therefore not cohesive. The result is that while the use of comparison may be a very important feature of discourse in these texts and therefore distinctive among the genres, it is not, in most cases, cohesive, at least in Halliday and Hasan's terms. Therefore, very few instances of cohesion by comparative reference were identified in the genres. The instances involved the use of such items as "similar", "such" and "other" . As these cases show, the items in this category are so few that no meaningful comparison can be based on them. A look at a larger sample may reveal more items, but as far as the sample in this study is concerned, these are the items which were found to occur.

6.6.2 Differences in Conjunction Cohesion

The analysis of sample texts undertaken in Section 6.5 shows that texts in all three genres make use of conjunction cohesion items . However, it does not indicate

the extent to which conjunction cohesion is distinctive among the genres. This can only be determined by further examination of texts in the genres for conjunction cohesion items whose use and function are related to the purpose of discourse in each genre. As usual, conjunction cohesion items in the RAP will first be discussed. This will be followed by a discussion of items identified in the ABS and the JRV.

The analysis of RAP texts for conjunction cohesion items does not reveal anything unique or distinctive in the way they are used. In other words, despite the fact that in the RAP, a wide variety of conjunction items tend to be used to achieve cohesion between sentences, nothing in their use suggests that they are constrained by either the purpose or audience of discourse in the texts in which they occur. However, the analysis shows that in addition to Reference cohesion items discussed in subsection 6.7.1 above, sentences in RAP texts also tend to cohere by means of Enhancement, Extension and Elaboration Conjunction items. The Enhancement Conjunction items are either those which express spatio-temporal or cause-conditional relations. The Extension items belong to the groups: Positive addition, Adversatives and Variation while the Elaboration items are either Particularizers, Summatives or Exemplifiers. Below is a comprehensive list of categories and items of conjunction cohesion identified in the RAP.

Category	Item
Enhancement	spatio-temporal
Enhancement	cause-conditional
Extension	Positive addition
Extension	Adversatives
Extension	Variation
Elaboration	Particularizers
Elaboration	Summatives
Elaboration	Exemplifiers

Table 30: Conjunction Cohesion Types in the RAP

A. Spatio-Temporal		
i)	Then + After an interval	e.g. <u>After</u> (RAP.5)
ii)	Then + specific Time Interval	e.g. <u>since</u> 1970 (RAP.7)
iii)	Then + repetition	e.g. <u>first, second</u> , etc.(RAP4)
iv)	Before + previous occurrence	e.g. <u>formerly</u> (RAP.5)
B. Extension		
i)	Positive Addition	e.g. <u>in addition</u> (RAP. 11) <u>furthermore</u> RAP.6)
ii)	Adversative	e.g. <u>however</u> (RAP.10)
iii)	Variation	e.g. <u>in contrast</u> (RAP.9)
C. Elaboration		
i)	Particularizers	e.g. <u>most importantly</u> (RAP2) <u>specifically</u> (RAP.11) <u>in particular</u> (RAP) <u>particularly</u> (RAP.)
ii)	Summatives	e.g. <u>in brief</u> RAP.6) <u>overall</u> (RAP.)
iii)	Exemplifiers	e.g. <u>that is</u> (RAP.11)

As has already been pointed out, the purpose, and audience of discourse in the RAP do not appear to have a distinctive impact on the conjunction cohesion items identified in texts in the genre. Perhaps, the only exception to this contention is the

group of items under the category "Particularizers" which, as can be seen in the example below, may function to indicate author commitment, thus having a persuasive impact.

¹The results of this study suggest that OC use has no significant effect on the risk of breast cancer in women under 45 years of age. ²Most importantly, we found no significant increase or decrease in the risk of breast cancer before age 45 years according to age at first use of OCs or duration of use.

(RAP.2: *The Lancet*, November 2, 1985)

In the text above, the particularizing item "most importantly" performs two functions. First, it enables the second sentence to cohere with the first by indicating that the proposition which follows it is a discussion of one of "the results of this study" which "suggests that OC use has no significant effect on the risk of breast cancer in women under 45 years of age". Secondly, it enables the reader to predict the degree of certainty, and/or confidence with which the author holds the propositions which the item precedes. Thus, in the text above, the item "most importantly" would seem to suggest that the author is certain about, and committed to the fact that "no significant increase or decrease" was found "in the risk of breast cancer before age 45 years according to age at first use of OCs or duration of use". To this end, its function relates to what is known as the modulation of propositions (e.g. Pindi, 1988) or epistemic modality (e.g. Coates, 1983).

Just as in the RAP, the Conjunction cohesion items used in the ABS are not quite revealing about the nature and purpose of discourse in texts in the genre. In spite of this apparent similarity, the ABS differs from the RAP in the types and amount of Conjunction items it characteristically employs for cohesive purposes. First, only two out of the three main categories of conjunction cohesion identified in the RAP tend to be realised in the ABS. The categories are "Enhancement" and "Extension". In other words, Elaboration items may not be part of the cohesive devices used in the ABS.

The tendency not to use Elaboration items in the ABS is understandable in view of the explanation by Halliday (1985b: 207) that:

In ELABORATION, one clause elaborates on the meaning of another by further specifying or describing it. The secondary clause does not introduce a new element into the picture but rather provides a further characterization of one that is already there, restating it, refining it, or adding a descriptive attribute or comment.

In view of the above, it is not surprising that elaboration conjunction items do not seem to be used for cohesive purposes in the ABS which essentially is a summary of the main results obtained in the RAP text which it accompanies. Since the ABS is only a summary of its accompanying RAP text, it does not have scope for further characterization, restating or refining of a proposition already made in it. Therefore, it is not likely to connect sentences by means of elaboration conjunction items.

Apart from the above difference, only one type of spatio-temporal item was identified in the ABS — the temporal adverbial "after" indicating a "then + after an interval" relation. Similarly only one cause-conditional cohesive item was identified in the genre — the item "therefore". In terms of Extension Conjunction, only one "adversative" item was also used for cohesive purposes in the ABS — the item "however". In the same vein, the only "variation" conjunction item identified in the genre is the contrastive conjunct "in contrast". In short, conjunction cohesion is not an important means of achieving textuality in the ABS. This points to the fact that although sentences in texts may be conceptually related, they are not necessarily explicitly sequentially related or linked. That is, each sentence is more or less an independent entity which shares a common cognitive content with other sentences in the text. It is by means of this common cognitive content, rather than by explicit signals of Enhancement, Extension or Elaboration that sentences in ABS texts are linked to one another.

In the JRV, as with the RAP, much of the use of conjunctive cohesion items is ordinary, not suggesting any distinctive discourse effect relative to the purpose and

audience of discourse for which texts in the genre are written. Table 31 below shows the conjunction cohesion categories and items realised in the JRV.

Table 31: Conjunction Cohesion Types in the JRV

1. ENHANCEMENT		
A. Spatio-Temporal		
i)	Then + after an interval	e.g. after (JRV.11)
ii)	Then + specific time interval	e.g. last week (JRV.14)
iii)	Then + repetition	e.g. One, two ... (JRV.1)
B. Cause-Conditional		
i)	Simple conditional	e.g. therefore (JRV.1)
2. EXTENSION		
i)	Positive Additional	e.g. and (JRV.7)
ii)	Adversative	e.g. but (JRV.3)
		yet (JRV.5)
		unfortunately (JRV.13)
3. ELABORATION		
i)	Exemplifiers	e.g. in Europe and ..(JRV4)
ii)	Clarification	e.g. in fact (JRV.2)
		indeed (JRV.11)

As the result in the table above shows, the JRV realises the same categories of conjunction cohesion items as the RAP. However, the JRV differs from the RAP in the type of Extension and Elaboration items it tends to use for cohesive purposes. This observation which constitutes a major source of difference between the JRV and the RAP also indicates that JRV texts may be distinguished from texts in the other genres by the use which it makes of conjunction items. Thus, unlike in the RAP, the "positive addition" item "and" is used in sentence initial position to link sentences together as the example below shows.

¹Proud and his colleagues found high levels of tissue kallikrein and kinins in the fluid from the lungs of these patients. ²And returning to the problems of the recent cold weather, the researchers also found the kinin and kinin-producing enzyme in the inflamed lungs of people suffering from bronchitis and pneumonitis.
(JRV.13: *New Scientist*, 26 February 1987)

The item and as used in the text above has one primary function. It marks the sentence in which it occurs as a writer-defined continuation of discourse, thus linking the sentence with the one preceding it. This contention is related to the observation made by Schiffrin (1987: 152) that "and" is a "structural coordinator of ideas which has pragmatic effect as a marker of speaker continuation".

The "adversative" conjunction item "but" is also used in sentence initial position to achieve cohesion between sentences in JRV texts as the example below demonstrates:

¹A high level of LDL in adults is regarded as a high-risk factor.
²But the Bogalusa comparison of children and parents uncovered other tell-tale signs in the youngsters, before the cholesterol balance registers an abnormal condition. (JRV.8: *The Times* 1 October 1986)

The adversative item "but" in the text above links Sentence 2 to Sentence 1 by its contrastive function. What is being contrasted is the idea of a high LDL risk in adults (Sentence 1) and "other tell-tale signs" uncovered about the risk in "the youngsters" (Sentence 2). This use of "but" relates to the observation by Schiffrin

(1987: 152) that "but" marks an upcoming unit as a contrasting action, adding that "but" "does not coordinate functional units unless there is some contrasting relationship in either their ideational or interactional content".

The use of "and" and "but" in the ways described above is unique to the JRV because they are features of language use which are more characteristic of spoken than written discourse and are more likely to be associated with informal than formal written texts. To this end, they represent one of the means by which JRV writers attempt to break barriers between themselves and their audience. Therefore, in its simulation of spoken language style the JRV attempts to break down barriers which exist between writers and their audience in its texts.

Another set of conjunction items which were found to be unique to the JRV texts in this study are those which tend to have clarification functions. They are the items "in fact" and "indeed". Their use is shown in the examples below:

In Britain, far less than five per cent of breast cancer cases would have been attributable to OC use in 1980, compared with perhaps 10 per cent in 1985 and 20 per cent in 1990. In fact, the Royal College of General Practitioners found such an increased risk, though it was of marginal significance.

(JRV.2: *New Scientist*, 7 November, 1985)

Reducing fat and cholesterol in the diet can lower cholesterol levels, as previous studies have shown, and a drop in blood cholesterol does cut the heart attack risk. Indeed, in the decade since the study was carried out, many Americans have given up smoking and consumed less fat — and the number of deaths from heart disease has already started to drop.

(JRV.11: *Newsweek*, December 8, 1986)

In the texts above, the items "in fact" and "indeed" function to link the propositions contained in the sentences where they occur with those in the sentences preceding them. In both cases the items indicate that the writers wish to reinforce and clarify propositions made in the preceding sentences. The need to reinforce and clarify propositions is one which is central to discourse in the JRV considering the audience-

level and the level of shared knowledge which exists between the writer and the reader of the JRV text.

In contrast to the use of clarification items discussed above, Table 31 shows that the JRV does not make use of "particularizing" elaboration items to achieve cohesion in texts. This may be due to the nature of the popularized texts examined in this study — those written by persons other than the original authors of the research being popularized. As we saw in the RAP, particularizing conjunction items function to indicate level of author commitment to research results. Since authors of JRV texts merely report on the research of others, there appears to be no reason why they should show strong commitment to the results in the original texts. Another related observation is the absence of "summative" elaboration conjunction items in the JRV. The analysis of RAP texts reveals that cohesive summative items tend to occur in Introduction and Methods sections, both of which are not favoured areas for the extraction of information for popularization in the JRV (see the analysis of Moves, Chapter 4). This may account in part for the absence of the items in the JRV.

6.6.3 Differences in Lexical Cohesion

As with the other types of cohesion already discussed in this Section, differences in lexical cohesion are not accounted for in quantitative terms, but rather by reference to tendencies in the use and function of items which enable sentences to cohere as related units of text. As has been pointed out elsewhere in this study (Section 6.5), the model analysis of the three sample texts taken from each of the three genres suggests that the most common form of cohesion employed in the genres is lexical cohesion. The analysis also shows that the most common form of lexical cohesion employed in the texts is lexical repetition. Other forms of lexical cohesion also identified in the analysis are the use of equivalent items, synonyms, meronyms and antonyms.

Following the results obtained in that initial inquiry, all texts in the corpus were analysed for lexical cohesion in order to determine whether, in addition to enabling sentences to hang together, the choice and function of lexical cohesion items is constrained by the purpose and audience of discourse in each genre. In other words, whether the choice and function of lexical cohesion items is distinctive in the genres. In order to do this effectively, all the lexical cohesion types identified in Hasan (1984) (see Section 6.5) were analysed for in the texts. However, it was found out during the analysis that not all of the categories of lexical cohesion being searched for had direct impact on the purpose of discourse in texts in the genres. Therefore, the discussion presented here is based on those categories of lexical cohesion in Hasan (1984) which tended to have distinctive impact on the purpose of discourse in the genres, and which therefore relate to the kind of problems we are interested in in this study.

As with the analysis of the sample texts referred to above, an examination of texts in the corpus reveals that "repetition", "equivalence" and "naming" are the three main types of lexical cohesion realised in the three genres. Other types of lexical cohesion identified in texts across the three genres, albeit in low proportions, are "synonymy", "antonymy" and "meronymy". Of the three main types only "equivalence" and "naming" appear to have been used distinctively in the genres; the type of items which cohere by Repetition being the same in the three genres. Thus, in functional terms, there appears to be no significant difference between the genres in the realization of lexical cohesion items in text.

Lexical items which contribute to textual cohesion in the RAP, ABS, and JRV as a result of their being repeated after initial use were observed to belong to the following categories:

1. Technical lexical items
2. Technical abbreviations
3. Symbols and formulae

An overwhelming majority of the items which tend to be repeated and by means of which sentences in texts in the three genres cohere are Technical lexical items. Some of these items are those which encode diseases, e.g., cancer, cervical and prostate cancer, lung cancer, metastatic cancer; meningitis, meningococcal disease, etc. Others are items that encode "basic medical terms". That is, items which are fundamental to medical discourse in the sense that they are common to all kinds of medical texts (Salager, 1983: 56), e.g. lymphocytes, cells, cases, controls, infant mortality, blood transfusion, patient, immunization, general practitioners, etc. The third type of items in this category are those which encode "specialist medical terms". These are items which, though used in the description of particular medical problems or ailments, are heterogeneously distributed in Medical English literature (Salager, 1985: 6). They are rarely used outside the field of medicine and related fields. They include words like placebo treatment, double-blind trial, interplacental arterial thrombosis, primigravidae, apolioproteins, pharmacogrologic effects, leukapherisis, optic neuropath, hyperapobetalipoproteinemia, etc. Items under these three categories were found to be frequently repeated, often in identical wordings, where they did occur in texts in the three genres. Perhaps the only difference between the genres in this respect is in the quantity of the items that really do occur in text.

The next group of lexical items which tend to be repeated in texts in the three genres are technical abbreviations. Quite often abbreviations of words are used only after their full forms have been presented and probably understood by the reader. The abbreviations used cohesively in texts across the three genres include: OC's, PIH, EBV, HSV, HSV DNA, CHD, HMWK, AMI, LAK, LDL, etc. It must, however, be pointed out that relatively fewer abbreviations were identified in the JRV compared to either the RAP or ABS. Thus, only three technical abbreviations — OC, EB virus and LDL — were identified in all the JRV texts examined. The tendency in the JRV is for

the abbreviations to be introduced alongside their full versions, as in the example below:

The findings provide an even better indicator of risk than the previous discoveries by other doctors of the importance of the ratio in blood serum of the two types of cholesterol: low-density lipoprotein (LDL) and high-density lipoprotein (HDL).

A high level LDL in adults is regarded as a high risk factor.

(JRV.8: *The Times* 1 October 1986)

The third group of lexical repetition items identified in the analysis of texts in the three genres is "Symbols and formulae". Not many symbols and formulae were found to be used cohesively in texts across the three genres. The number identified in the ABS was much fewer than that in either the RAP or the JRV. All the same, the following symbols and formulae were found to have contributed to cohesion in all three genres: HSV-2, HPV-16, TSA₂, PGI₂, C3d, 1-HMWK, HB5 (1gG2a), etc.

Lexical repetition tends to perform two functions in the texts examined. First, it tends to indicate the continuity of subject-matter of discourse in the text. In principle, every item which is repeated more than once contributes to the continuity of the subject-matter of discourse in text. However, the level of continuity which lexical items can provide varies in accordance with how close they are to the topic of discourse. In other words some lexical items tend to perform the function of indicating continuity of subject-matter more than others. In this study, it was observed that "technical lexical items", particularly those used to "encode diseases" tend to perform the function of indicating continuity of subject-matter more than either "technical abbreviations" or "symbols and formulae". The reason for this is probably because as has been shown in this chapter, technical lexical items are very closely associated with the topic of discourse in texts in the genres. The term "topic of discourse" is used here in its non-technical sense to mean that which the entire text is about. It corresponds largely with the title of the text, but extends beyond it to include all forms of exposition

on the subject-matter being discussed within the text. To this end, the term subject-matter can be said to be interchangeable with the concept of topic of discourse as it is used in this study. Therefore, this intuitive characterization of the term "topic" differs from the much narrower conceptualization of the same notion by Keenan and Schiefflin (1976) Li and Thompson (1976), Chafe (1972), Halliday (1967) and Givon (1983), which are restricted to the characterization "sentence topic".

Based on the conception of topic or subject-matter stated above, the repetition of technical lexical items and technical abbreviations as well as symbols and formulae function to indicate the continuity of subject-matter, as the example below demonstrates:

¹Pregnancy-induced hypertension (PIH) and pre-eclampsia appear to be associated with increased production of Thromboxane A₂ (TXA₂), a potent vasoconstrictor and stimulator of platelet aggregation, by the placenta and by platelets. ²TXA₂ has been put forward as an aetiological factor for the vasoconstriction, platelet-hyperactivity, and uteroplacental arterial thrombosis, which characterize PIH and pre-eclampsia. ³A daily dose of aspirin as low as 1-2 mg/kg effectively inhibits platelet cyclo-oxygenase and synthesis of TXA₂ by platelets and therefore may have a favourable effect. ⁴Thus, Crandon and Isherwood obtained a history of aspirin intake during pregnancy from primigravidae and showed that in women who had taken aspirin more than once every 2 weeks throughout pregnancy, PIH and pre-eclampsia were significantly less common than in women who had no history of aspirin ingestion. ⁵Similarly, results of a more recent randomised but not double-blind study suggested that aspirin might prevent recurrent pre-eclampsia in multigravidae. ⁶However, the possible prevention of PIH and pre-eclampsia in primigravidae by suppression of platelet TXA₂ production with low-dose aspirin has not been investigated.

(RAP.3 *The Lancet*, January 4, 1986)

The title of the text from which this example was extracted is "Low-dose Aspirin Prevents Pregnancy-Induced Hypertension and Pre-eclampsia in Angiotensin-Sensitive Primigravidae". An examination of the lexical items underlined in the text reveals that virtually all of them belong to an item in the title of the text and that one or the other of the items is repeated in every sentence in that text. Thus the items form

what Halliday and Hasan (1976) refer to as "cohesive chains" or sequences which tend to indicate the continuity of subject-matter. A typical example of how the chain works may be seen in the continuity provided by four lexical items, below:

Table 32: Lexical Chains in a Sample RAP Text

Sentence No.	(I) Pregnancy-Induced Hypertension (PIH) and Pre-eclampsia	(II) Thromboxane A ₂ (TXA ₂)	(III) Platelet aggregation	(IV) Aspirin
1	Pregnancy-Induced hypertension (PIH) Pre-eclampsia	Thromboxane A ₂ (TXA ₂)	Platelet aggregation	
2	PIH and Pre-eclampsia	TXA ₂	Platelet- hyperactivity	
3		TXA ₂	Platelet Cyclo- oxygenase	Aspirin
4	PIH and Pre-eclampsia			Aspirin
5	Recurrent Pre-eclampsia			Aspirin
6	PIH and Pre-eclampsia	TXA ₂		Low-dose Aspirin

The second function which lexical repetition tends to perform in texts in all three genres is "the identification of key concepts". The term "key concepts" is being used here to refer to key phrases or expressions or to what is often referred to as "key words" in a research paper. The function of indicating key concepts is most commonly realized by "technical abbreviations". But it may also be realised by "technical lexical items". The distinction between technical abbreviations and technical lexical items in

context is thin since one is an abbreviation of the other which is usually given in the text. However, the tendency to constantly repeat abbreviations instead of the full version cannot be fortuitous. Such repetitions, apart from conforming to conventions of writing in science, tend to highlight the importance of the item to the universe of discourse in text. Therefore, it can be said that the purpose of employing technical abbreviations and repeating them over stretches of text is, amongst others, to identify the concepts they represent as key concepts in the universe of discourse in the text. A useful example might be the constant repetition of "PIH" and "TXA₂" in the example text analysed for continuity of subject matter above. But more importantly, the use of technical abbreviations is an indication of the effect of audience on discourse in the genres. This is so because the decision to use technical abbreviations often involves making assumptions about the expertise of the reader, his familiarity with the terms and or sufficient concern for the topic to process initial mention when it is given in full.

As has been stated earlier, there is hardly any difference in the way lexical repetition operates and in the realization and function of repeated lexical items across the three genres. Therefore, in this respect the three genres can be said to be functionally related. However, they may differ in the density of lexically repeated items, but this is outside the scope of this study. It can then be said that, while lexical repetition is an important feature of cohesion in the genres, it does not seem to constitute a distinctive characteristic for differentiating the three genres.

One category of lexical cohesion which is certainly distinctive among the genres is "Naming". In this study, "naming" is taken to mean the use of proper nouns (personal names) to achieve textual cohesion. It involves the repetition of names of participants across sentence boundaries.

An examination of texts in the corpus shows that personal names which tend to contribute to textual cohesion across the genres belong to two main categories:

1. Those that refer to RAP writers

2. Those that refer to other researchers (citations).

These categories have also been discussed under Reference Cohesion. But that discussion was limited to the use of personal and possessive pronouns as reference cohesion items. In this Section, however, attention will be focused on the use of personal names as items of textual cohesion, for as has been stated, Halliday and Hasan (1976: 34) point out that the repetition of personal names in text is, strictly speaking, lexical cohesion.

In line with the above, the analysis of texts in the corpus reveals that the most common cohesive role which personal names play in the genres is to refer to RAP writers. The tendency to refer to RAP writers by their personal names appears to be unique to the JRV as no instances of such reference was observed in either the RAP or the ABS, except perhaps for authors' names which usually follow the title of the text. However, these names do not usually get repeated in the main text; hence they cannot be said to be cohesive in lexical terms.

The result obtained with regard to the use of personal names to refer to RAP writers in the JRV raises one fundamental question — how is the tendency referred to above a reflection of the constraint imposed on JRV writers by the purpose, audience and medium of discourse in the genre? The tendency to refer to RAP writers by their personal names in the JRV texts is due in part to the fact that the JRV is an attempt by a different writer to report what someone else has written. It may also be due to the fact that, though the JRV is a form of scientific writing (Widdowson, 1979), it nevertheless is essentially journalism. As journalism, the JRV tends to manifest some of the characteristics associated with journalistic prose. One of those characteristics is the tendency for it to have a participant line. This is supported by the contention of Smith and Frawley (1983) that journalism is "something of a narrative genre" because, like fiction, it involves a description of facts, events and participants in these events. Related to this is also the fact that the tendency to refer to writers of RAP texts in the

JRV by their real names is constrained by the need for the JRV writers to present what Myers (1986) refers to as "the narrative of nature" in which "the plant or animal, not the scientific activity is the subject".

In addition to the observations made above, an adequate explanation of the tendency to refer to RAP writers by their names in JRV texts may lie in examining how these names are realized in context and the functions they perform in those contexts. This is necessary because, although proper names, as used in the JRV, have a referring function, the naming conventions adopted in the genre can result in the names having a persuasive impact. Three naming conventions were identified in the JRV texts analysed. The first naming convention is reflected in the tendency not to accompany the names of RAP writers with the academic titles of their owners. For example:

1. Neil Blumberg and his colleagues (JRV.6)
2. Blumberg and his colleagues (JRV.6)

This naming convention is one instance of the informal writing style which is characteristic of JRV texts. Casual or informal reference to RAP writers can be seen as a strategy adopted by JRV writers to appeal to their audience, for as Adams-Smith (1987: 637) points out, "a journal or newspaper must appeal to its readers in order to sell, so the tone adopted is less formal, even flippant at times". The use of an informal naming convention is, perhaps, one of those less formal features of writing adopted in the JRV to appeal to the interest of the readership.

A second naming convention identified in the JRV is the tendency to first refer to RAP writers by their names and titles and to drop the titles in subsequent references in the text. For example:

1. Professor D J P Baker and Dr C Osmond (JRV.5)
2. Baker and Osmond (JRV.5)

The tendency to first of all refer to RAP writers by their titles and names and to drop the titles in subsequent references can be said to have a persuasive impact. For instance, the fact that the results of the research being popularized in JRV was reported by "Professor D J P Baker" in conjunction with "Dr C Osmond" might persuade the reader about the implications of the research which JRV writers are keen to put across. Having done this, the writer reverts to the less formal and casual convention of naming which is characteristic of JRV texts and which is more appealing to its audience.

The third naming convention identified in the JRV is that in which the researcher is named within the context of the establishment in which the research was conducted. It is more or less an integral part of the first two types already discussed. For example:

1. Neil Blumberg and his colleagues at the University of Rochester, New York....
2. Professor D J P Baker and Dr C Osmond of the Medical Research Council's Environmental Epidemiology Unit, London

In this type of naming convention, the modification to the head noun — that is, reference to the context in which the research was conducted — would seem to function to underline the nature of the value which the JRV writer attaches to the research being popularized. The tendency to identify RAP writers with establishments where they work is also likely to have an impact on the reader's assessment of the quality of results obtained and the implications drawn from them. As Myers (1989) points out, "there are great differences in the relative power between a Nobel Prize winner and an assistant professor at the University of Idaho, as observed at a conference, or in letters, or referees' reports". Therefore, as pointed out earlier in this section, although proper names do have a referring function in the JRV, the naming conventions adopted in the genre can lead to the names having a persuasive impact.

In summary, the analysis of textual cohesion reported in this chapter has revealed that there is an apparent similarity between the three genres in terms of the cohesion types they characteristically employ. Thus, three cohesion types — Reference, Conjunction and Lexical — are the main textual relations which hold between sentences in texts across all three genres. Despite this apparent similarity, there emerged interesting pragmatic differences in the way items belonging to these cohesion types are realised and the functions they perform across the three genres.

Before commencing on this analysis, a number of working hypotheses were posited and questions to steer the study towards its set purpose were asked. The present chapter attempts to assess the extent to which issues raised in these hypotheses and questions have been met by the results obtained in the analysis of data. This chapter is also an attempt to indicate conclusions which may be drawn from the findings made in this study and the implications which the study has for genre analysis in general and English for Science and Technology (EST) in particular.

CHAPTER SEVEN

SUMMARY, CONCLUSIONS AND IMPLICATIONS

7.0 INTRODUCTION

This study is an attempt to characterize differences in the patterns of discourse organization between three parallel genres of written medical discourse - the Abstract (ABS) accompanying a research article, the Research Article Proper (RAP) and the Journalistic Reported Version (JRV) of the same research article published in popular science magazines and newspapers. It adopts three approaches to the characterization of discourse organization in texts. The approaches adopted are: schematic structure, thematic progression and cohesion. These three approaches are, in fact, related aspects of discourse. In the relationship which exists between them, the aspect of schematic structure is superordinate to those of thematic progression (TP) and cohesion. The analysis of schematic structure accounts for the organization of texts in the corpus into segments which are made up of characteristic linguistic features. The segments also include within them utterances which are ordered in accordance with their FSP into identifiable patterns of TP. They also embody sentences which are bound together by means of the relations of textual cohesion. The results obtained with each of these related approaches are discussed in Chapters Four, Five and Six respectively, alongside discussions on variations which exist between the genres with respect to the realization of each approach.

Before embarking on the analysis, a number of working hypotheses were posited and questions to steer the study towards its set purpose were asked. The present chapter attempts to assess the extent to which issues raised in these hypotheses and questions have been met by the results obtained in the analysis of data. This chapter is also an attempt to indicate conclusions which may be drawn from the findings made in this study and the implications which the study has for genre analysis in general and English for Science and Technology (EST) in particular.

7.1 SUMMARY OF RESULTS

7.1.1 Summary of Results on Schematic Structure

The discussions on differences in schematic structure (Chapter Four) began with the characterization of the surface feature characteristics of each genre (Section 4.2) after which attention was focussed on schematic structures proper. In terms of the iconic or the surface feature organizational format of texts in the three genres, it was observed that the RAP differs distinctly from the ABS and the JRV in having a standard frame-like format for the presentation of information in its texts. This format is the IMRD (Introduction, Materials and Methods, Results, Discussion) organizational plan of most research articles in science and technology journals. Thus, in contrast to the RAP, the ABS and the JRV texts examined in this study do not have explicit or formal frame-like formats against which information may be mapped. However, they do have identifiable paragraph boundaries which indicate discourse organization. The genres were also observed to differ distinctly in the form in which titles are written, in the amount of information contained in the propositions which make up titles and in the level of abstraction and complexity of titles. Since the ABS and the RAP are usually written under the same title, the distinction made with respect to titles is between the ABS and the RAP on the one hand and the JRV on the other. Therefore, with regards to titles, it was observed that the JRV tends to make use of short, simple titles in contrast to the ABS and RAP which favour long, complex titles. Titles in JRV texts also tend to focus only on one medical problem and are often couched in such a way as to attract attention, sometimes by making an emotive appeal. In contrast, titles in RAP texts were observed to focus on more than one medical problem at a time and couched in a manner intended to reflect the complexity of the problem addressed in the text. Related to the issue of simplicity and brevity of titles in the JRV is the fact that they are clearer and easier for the audience to process than titles in ABS and RAP texts. This is

due to the fact that titles in RAP texts contain more abstract and complex propositions in their titles than JRV's. Thus titles in RAP's contrast with those in JRV's most of which were observed to be made up of simple expressions, which like newspaper headlines, tend to embody few co-ordinations and subordinations. Titles in both genres were also observed to differ in the verb forms they characteristically contain. Thus, titles in RAP and ABS texts tend to be made up of verbless propositions while propositions in titles in JRV texts tend to have verbs, most of which are finite. It was also observed that titles in the two groups of genres tend to perform different functions. With regard to this, titles in JRV texts function to indicate "the consumer approach", the "look-after-yourself approach", "the environmental approach" and "the medical approach" to medical discourse (Karpf, 1988). Titles in the ABS and the RAP, on the other hand tend to function only to indicate the "medical approach". Finally, in terms of the surface feature characteristics of discourse in the genres, it was observed that ABS and JRV texts in this study do not contain accompanying visuals such as graphs, tables and figures which appear to be normal features of RAP texts.

In terms of the organization of discourse in texts in the genres into identifiable schematic units or MOVES, it was observed that the genres differ in the following respects:

- 1 The distribution of Moves which texts in each genre are capable of realizing.
- 2 The distribution and type of constituent elements which make up the Moves which texts in each genre are capable of realizing.
- 3 The type of linguistic features by which Moves are characterized in texts across the genres.
- 4 The order in which Moves realized in texts in each of the genres are presented.

With regard to Item 1 above, the result shows that RAP texts tend to embody more Moves than ABS and JRV texts. Thus, while RAP texts tend to be made up of eleven Moves, ABS and JRV texts were found to be capable of realizing only nine Moves. The Moves which each genre realizes are characterized either as "Normally Required" or "Optional" depending on their regularity of occurrence in texts in the corpus. By means of this classification it was possible to observe that the JRV differs distinctly from the ABS despite the apparent similarity between both genres in the number of Moves each of them is capable of realizing in text. In terms of this, out of the nine Moves of the ABS, only two are "normally required moves" while seven are "optional moves". In the JRV, on the other hand, seven out of the nine moves which texts in the genre are capable of realizing are "normally required moves" while only two are "optional moves". The RAP differs from either of these two genres in having nine "normally required moves" and only two "optional moves". Three information reduction strategies were found to be in use and these constrain the number of moves which the ABS and the JRV select from the RAP. They are: Omission, Combination and Construction. Of these strategies, Omission appears to be the predominant strategy of information reduction adopted in the ABS and the JRV.

With regard to item 2 above, it was observed that the RAP with twenty-eight "constituent elements" differs very strongly from the ABS with fourteen constituent elements and the JRV with nineteen constituent elements. Constituent elements from the RAP which are deleted in ABS and JRV texts would seem to represent information which are either unimportant or irrelevant to the needs of the audience and the purpose of discourse in the genres. Apart from differences in the quantity of constituent elements which texts in each genre are capable of realizing, there is also the difference in the quality of constituent elements realized in texts across the genres. This difference is manifested very strongly by the constituent elements identified in JRV texts. The analysis of JRV texts shows that they embody constituent elements which are unique to

the JRV and which are not found in either the RAP or the ABS. The constituent elements which are unique to the JRV are:

- (a) Stressing the local angle
- (b) Explaining principles and concepts
- (c) Indicating comments and views.

These constituent elements would seem to serve the purpose of improving readability by improving on the level of shared knowledge which exists between the writer and his audience.

With regard to item 3 above, the analysis shows that Moves in the RAP tend to have more sentences than those in the ABS and the JRV (Table 21). In contrast, it was observed that the ABS has the highest average number of words per sentence relative to the total number of words per text. However, the difference between it and the RAP in this respect is almost imperceptible. The JRV was found to have the lowest average sentence length. This result would seem to be a reflection of the level of complexity of sentences and of discourse in the genres. Other linguistic differences identified in the analysis include the tendency for transformations to occur in the forms of the verb when identical information is being reported across genres. For instance, a transformation from the present perfect active in Texts 'A' (RAP) and 'C' (JRV) (Section 4.6.3) to the past tense passive in text 'B' (ABS). Finally, the genres were found to differ in the extent to which they make use of interrogative forms. In this regard, it was observed that only the JRV incorporates interrogative forms; no interrogative forms were identified in either the RAP or the ABS.

With regard to differences in the order in which Moves are presented across the genres (item 4 above), it was observed that there is a much closer relationship between the RAP and the ABS than there is between both genres and the JRV. That is, the RAP and the ABS seem to adopt identical strategies in the organization of the moves

they contain which is quite different from that adopted by the JRV. In the RAP and the ABS, Moves can be said to follow each other in a logical and sequential manner following the conventional order of presentation of argumentation in learned scientific journals. In other words, though the ABS tends not to select all the Moves found in the RAP, the omission of RAP Moves does not necessarily lead to the re-organization of the pattern in which information is normally presented in scientific journal articles. In contrast to the RAP and the ABS, there appears to be considerable flexibility in the way Moves are ordered and sequenced in the JRV to the extent that an impression is created that there is no uniformity in the way various JRV writers organize information in texts in the genre. However, a close look at the regularity of Moves in the JRV shows that they can be classified into three main divisions: Initial, Medial and Final Moves. Using the same models as those found in the RAP, included in the initial moves are moves 1, 9, 2 and 3; in the medial moves are 7, 4 and 5 and in the final moves are 10 and 11. The Moves in the JRV were characterized as not following the same sequence as in the RAP (see Section 4.6.4, Chapter 4).

7.1.2 Summary of Results on Thematic Progression

Differences in the Thematic Progression patterns employed in texts across the three genres were discussed under the following headings:

- The realization of TP patterns across genres
- The distribution of TP patterns
- The characteristic elements which realize theme across genres.

In terms of the TP patterns characteristically realized by texts in each of the genres, it was observed that the simple linear and the constant TP patterns are features

of text in all three genres. That is, every text in the corpus was found to make use of the simple linear and the constant TP patterns to organize and develop discourse. The analysis also showed that the TP with a split rheme is a feature of discourse organization in only the RAP and the JRV. Finally, the TP with a derived theme was identified in the RAP only.

The occurrence of the simple linear and the constant TP patterns in all texts in the corpus indicates that they are fundamental patterns for the organization of information in texts in the genres studied. The differences in the number or types of TP patterns realized by texts across the three genres is accounted for in terms of differences which exist between the genres in the length of their texts, and, therefore, the scope which they have for the development of information.

In terms of the distribution of TP patterns, it was observed that the genres vary in the extent to which they make use of the simple linear and constant TP patterns to organize and develop the information they contain. The Constant TP pattern was found to predominate in the RAP, while in the JRV, the predominant pattern was found to be the simple linear TP pattern. The ABS, however was observed to make use of both the simple linear and the constant TP patterns fairly evenly. The high occurrence of the constant TP pattern in the RAP is accounted for by reference to the tendency for authors of research reports to thematize specimens and subjects in the research, particularly in the Methods and Results sections of the research paper. Since these are things which the researcher holds in common with the reader as far as the experiment is concerned, the tendency to thematize them can be said to reflect an attempt by authors of research papers to take into consideration the knowledge which they share with the reader in the organization of information in utterances in RAP texts. On the other hand, the high occurrence of the simple linear TP in the JRV is accounted for by reference to the low level of shared knowledge which exists between the writer of a JRV text and its reader. This situation, it is argued, constrains the writer to constantly select from information

presented as rheme in a preceding utterance in constructing the theme of the utterance following it as a means of mitigating the effect which a lack of shared knowledge can have on the quality of the writing and the reader's ability to process information.

No one-to-one correspondence was found between the distribution of TP patterns and the organization of discourse in texts in the genres into identifiable schematic structures. However, some schematic units were found to show preference for one TP pattern over others. Thus, the simple linear TP pattern was observed to predominate in schematic units of text which are made up mainly of expository or explanatory information. In contrast, the constant TP pattern was observed to occur with a high frequency in schematic units which involve descriptive and procedural information. The former are represented by schematic units identified in the Introduction and Discussion sections in RAP texts. The latter consist mainly of units in the Methods and Results section in RAP texts. Related to the findings reported above is the observation that the Derived TP pattern tends to occur in paragraph structures in RAP texts which involve itemization or serialization of events.

In terms of the characteristic elements which realize theme across the genres, it was observed that in the RAP, schematic units of text which are organized in accordance with the constant TP pattern tend to make use of complex nominal subjects as theme or the combination of a prepositional phrase adjunct and a nominal group subject as theme. In either case, the items which function as theme are most often non-agent subjects. Another item which regularly functions as theme is a constant TP pattern of organization in the RAP is the "there" particle in a there-construction. The tendency for utterances in RAP texts to make use of non-agent subjects is seen as an attempt by authors of research papers to reconcile the FSP of utterances with the demands of the conceptual presentation of scientific and medical knowledge which places emphasis on impersonality and the scientific activity in the presentation of facts. It is also considered as being constrained by the need for detachment from the subject-matter of discourse by

authors of research papers. In the JRV, on the other hand, it was observed that schematic units which are organized following the constant TP pattern tend to make great use of personal pronouns and proper or personal names as theme. The use of personal names as theme in the JRV is seen as a reflection of the relationship which the genre shares with narrative genres in having a participant line. It is also seen as a strategy aimed at involving the reader in the subject-matter of discourse in JRV texts.

7.1.3 Summary of Results on Cohesion

Overall, the analysis of cohesion reveals that the ABS, RAP and JRV tend to make use of the same types of cohesive devices in their texts. That is, all three genres make use of Reference, Conjunction and Lexical Cohesion items. It also means, albeit in relative terms, that they all tend to make much use of lexical cohesion items as a means of textual cohesion. Furthermore, it means that all three genres appear not to make cohesive use of Substitution and Ellipsis.

The results obtained with Reference Cohesion shows that Personal Reference is the predominant form of reference cohesion employed in texts in the three genres. Personal reference cohesion items identified in texts in the corpus were classified according to participant roles in the discourse into:

- 1 Those that make reference to authors of the research article.
- 2 Those that make reference to authors and other persons.
- 3 Those that refer to other researchers (situations).
- 4 Those that refer to aspects of research subject-matter.
- 5 Those that refer to research participants.

With respect to Item 1 above, the ABS and the RAP were found to differ distinctly from the JRV in the personal reference items they characteristically use. While the ABS and the RAP tend to make use of the first person plural pronoun 'we' and the possessive pronoun our to refer to authors, the JRV makes use of the pronouns 'they' and the possessive their for the same purpose. The difference in the choice of items in both groups of genres lies in the fact that in one situation, the authors and the researchers are the same persons while in the other, the author of the text is not the same as the authors of the research article.

In terms of cohesive force, reference to the authors of the research in question and others was found to be unique to the JRV. The personal pronoun we and the possessive our are the means by which this function is realized in the JRV.

In terms of reference to other researchers, especially in the form of citations, it was observed that the ABS differs markedly from either the RAP or the JRV in not having citations in its texts. It was equally observed that citation, or reference to other researchers is not a function which personal reference cohesion items can be strongly identified with in the JRV. Reference to other researchers by means of personal reference cohesion items is much stronger in the RAP where the quality of the research reported in the text is almost invariably measured by the number of citations it contains. Although those citations which involve the use of personal names are not instances of reference cohesion, subsequent reference to them by means of personal pronoun definitely are. In addition to the fact that the JRV and the RAP differ in the extent to which personal reference items are used to refer to other researchers, both genres were also found to differ in the purpose which such a reference serves. While in the RAP the writer presents his own claims to deny, support or supersede claims made in the cited works, citations function only to reinforce claims made by researchers in the JRV.

No distinctive differences were found either in the realization of items or in the function performed by the personal reference cohesion items which make reference to

aspects of subject matter (item 4) and reference to participants (item 5) across the genres.

The other Reference cohesion types examined in the corpus are Demonstrative and Comparative reference cohesion. No distinctive difference was found across the genres in the way demonstrative items are used for cohesive purposes. The demonstrative items identified in texts across the genres are this, that, and those as well as the non-selective demonstrative the. The selective demonstratives, especially this and that are used to contrast events, actions and propositions as well as refer to events, actions and propositions. They are also used to direct the reader's attention to a new referent object and to sum up propositions in immediate referential distance.

It was found that the genres do not rely very much on comparative reference for cohesion. This is because a lot of the instances of comparison found in the texts are intersentential, rather than intra-sentential. Therefore, while comparison may be a strong linguistic feature of discourse in the genres, its use is not, in most cases, intrasententially cohesive, which is the kind of cohesion under consideration (cf. Halliday and Hasan). Thus, very few instances of comparative reference were observed in texts in the corpus. Those few instances involved the use of the items similar, such and other.

The genres were found to differ distinctly in terms of Conjunction cohesion. The main distinction, though, is between the popularized genre (JRV) on the one hand, and the professional genres (ABS and RAP) on the other. In line with this, it was observed that the JRV differs from the other two genres both in the conjunction items it uses and in the way they are used. While not much can be said about the extent to which the use of conjunction cohesion items in the ABS and RAP is constrained by the purpose, audience and medium of discourse, the use of cohesive conjunction items in the JRV is very revealing of the effects which these variables have on discourse in the genre. Thus, as the result reported in Chapter Six shows, the JRV is the only genre

which makes use of the additive item AND and the adversative item, BUT in sentence initial position in text. The way the items are used in the JRV is characteristic rather of spoken or informal than written or formal discourse. In addition the JRV was also found to make use of conjunction cohesion items which have clarification functions. The items are in fact and indeed. It cannot be denied that the need to constantly clarify propositions is central to effective communication of information to a less-specialist or even lay audience. As has already been pointed out, the use of conjunction cohesion items in the ABS and the RAP have nothing unique about them. Perhaps this is due to the highly formal nature of discourse in these genres. Despite this apparent similarity, some form of difference was found between the ABS and the RAP in the types of conjunction items they tend to use for cohesive purposes. Worthy of note in this regard is the tendency for sentences in ABS texts not to cohere by means of elaboration conjunction items. Elaboration conjunction items function to link sentences together in text by providing further characterization of a preceding sentence or a sentence that is already in the text. Since the ABS is a summary of the RAP, it is understandable that it does not make use of elaboration conjunction items for cohesive purposes. In addition to this, the ABS also tends to make limited use of spatio-temporal, cause-conditional, adversative and extension cohesive conjunction items compared to the RAP as only one instance of each of these conjunction cohesion types was identified in the ABS text analysed in this study.

With regard to Lexical Cohesion, three main types of lexical items tend to function to hang sentences in the genres together. They are items which are repeated (lexical repetition) in exactly the same wordings, items which have equivalent meanings (equivalence) with a previously used lexical item and items which are the names of persons or places (naming). Of these three, only equivalence items and naming items contribute to distinguishing the genres in terms of cohesion. In other words, hardly any difference was observed between the genres in the way lexically repeated items are realized and in the functions they seem to perform. The most distinctive form of lexical

cohesion identified in the genres is naming. Personal names which contribute to cohesion in the genres are of two types:

- those which refer to RAP writers
- those which refer to other researchers (citations).

The use of personal names of RAP writers as a cohesive device is unique to the JRV.

Names of RAP writers are not repeated in the body of the text in RAP and ABS texts.

Three naming conventions are adopted in the JRV. They are:

- Academic titles and names
- Academic titles and names at initial mentioning and names without titles in subsequent mentioning
- name only and institution or establishment.

The naming convention adopted was found to be part of the informality of writing in the JRV and a strategy for establishing solidarity with the audience.

7.2 CONCLUSIONS

The results obtained in this study lend weight to the following general conclusions:

- 1 Although the ABS, RAP and JRV texts examined in this study report on the same subject-matter and in the same field, one should not be misled into assuming that they belong to the same genre, or that one or the other is a sub-genre of another.

It has been pointed out elsewhere in this thesis (Chapter 2, section 2.2.3.1) that a genre is not only a recognized communicative event with a shared public purpose with aims mutually understood by participants within the event, but also a recognized communicative event which in terms of positioning, form and intent is structured and standardized (Swales, 1985). In line with this, it has been shown in this study that the three text-types differ distinctly in the iconic or surface organization of information they contain, in the organization of information in utterances in a manner that indicates author perspective, as well as in the cohesive relations which exist between sentences in text. In addition, the text-types have also been shown to be written for different purposes and with different audiences in mind. Therefore, based on the results obtained in this study, each of the text-types is a distinct genre.

- 2 A meaningful characterization of genre necessitates the isolation of patterns of discourse structure which in their totality reflect the communicative potential of a genre.

As has been shown in this study, it is not sufficient to characterize genres by isolating only one form of discourse structure and characteristic linguistic forms. An integrated approach of the kind adopted in this study is needed to fully account for the structure of discourse which genres manifest. A partial description of the discourse structure of genres might lead to confusions about genre differentiation and classification. Therefore, it is only by isolating patterns of discourse structure which in their totality reflect the communicative potentials of a genre that we can begin to understand in detail what constitutes the real message of a genre, how it is produced and how the message it contains is intended to be received by the audience. And, it is only by doing this that we can begin to understand what the term "genre" is meant to cover.

- 3 Despite the fact that genre-analysis is primarily concerned with the characterization of individual genres, it is also a viable model for the comparative study of texts.

This thesis has demonstrated the fact that even within what may be broadly termed "the language of medicine", texts which report on the same topic can differ in the patterns of discourse structure they characteristically employ regardless of the similarities which may exist between them in lexicogrammatical terms. Thus, it reinforces questions often asked regarding how meaningful it is to talk about things like "the language of medicine", "the language of engineering", "the language of agriculture", "the language of physics", etc. Perhaps, it might be better to talk about "the language of abstracts accompanying a medical research paper", the language of medical research articles", "the language of popularized medical research reports" etc instead.

- 4 Though a genre is a recognized communicative event which is structured and standardized, it can also share features of discourse structure with other recognized communicative events with which it is clearly associated in terms of topic and field of discourse.

It can be seen that while the three genres in this study manifest distinct patterns of discourse organization, they also hold a number of features in common. It would seem that the extent to which one genre shares features of discourse structure with another depends on how closely related they are in terms of purpose and audience of discourse. Thus, the ABS and the RAP which, though different, share aspects of audience and purpose, have a lot more in common in terms of discourse structure than either of them has with the JRV.

- 5 The patterns of discourse organization employed in popularization by science journalists reflect the view held by Myers (1986) in relation to popularization by research scientists that popular science texts present science as "the narrative of nature".

This relates in some respects to the numerous strategies adopted in JRV texts to construct discourse as human interest stories. It includes the tendency to relate discourse in the JRV to the real world of the reader through the inclusion of constituent elements such as those which "stress the local angle"; the tendency to provide the reader with explanations of concepts and principles upon which researches are based; the tendency to present information relating to the main findings in the research as the initial segment of discourse in text; the tendency to refer to authors of research papers by their personal names; the tendency to adopt more of the simple linear TP pattern in the organization of utterances and to constantly thematize personal names and titles and names of places (towns, countries etc) as well as the tendency to adopt an informal, almost casual style in the use to which cohesive items such as conjunctions (the items and and but) and questions are put in text.

7.3 THEORETICAL IMPLICATIONS

7.3.1 Theoretical Implications for Discourse Organization and Varieties Differentiation

The implications of the results obtained in this study are both general and specific. In general terms, the results have implications for the description of discourse organization and the nature of variation which contextual factors have on the organization of information in text. Thus, in analysing the patterns of schematic structure, thematic progression and cohesion in texts in the corpus, an attempt is also made at ascertaining how the contextual factors of purpose, audience and context constrain the organization of discourse in the texts into chunks known as Moves; the configuration of utterances into theme and rheme and how their concatenation accounts

for the progression of information as well as ascertaining how the variables constrain the relations of cohesion which enable sentences to hang together.

With regards to schematic structure, it was observed that despite the fact that triads of texts in the corpus report on the same subject-matter in the same field of discourse, they differ in the way information they contain is organized into hierarchical schematic structures. Thus, whereas the ABS tends to have much fewer schematic units of information than the RAP or while the JRV organizes the schematic units it contains in a much flexible manner than the RAP, the discourse in texts in all of the three genres still remains coherent and well organized. Again, whereas the ABS and the JRV contain much fewer constituent elements (sub-moves) than the RAP, and whereas the JRV contains new constituent elements outside of those found in the RAP, the organization of discourse in texts in each genre does not distort the message which they contain. In other words, the fact that a Move or a constituent element in either the ABS or the JRV is not the same or is not preceded and followed by another Move or constituent element in strictly the same order of occurrence as in the RAP does not mean that the information they present is either incoherent or ill organized.

The results obtained with regard to thematic progression relate to those discussed above for schematic structures. For instance, it was observed that despite the fact that texts in the corpus report on the same topic and in the same field of discourse, they tend to differ in the extent to which they make use of one TP pattern over others to indicate the continuity of discourse in texts. Thus, whereas the ABS tends to strike a balance between the simple linear and the constant TP patterns and whereas texts in the RAP tend to be developed and organized mainly according to the constant TP pattern and whereas information in utterances in the JRV follow predominantly the simple linear TP pattern, texts in all three genres succeed in communicating the central message in the researches they report upon. This shows that discourse is not an arbitrary sequence of utterances. Writers provide their readers with a perspective from which they can decode

the message contained in an utterance, a perspective which, as Danes (1970, 1974) has recognized, involves the concatenation of Theme and Rheme into patterns of Thematic Progression. The predominant pattern of thematic progression adopted in a particular text would, therefore, seem to reflect the level of assumptions which writers have about their readers based on the level of shared knowledge which exists between them.

With regard to cohesion, it is significant that though texts in the three genres seem to make use of the same cohesion types, they differ in the categories and items which they use. The analysis of cohesion may be conducted for the purpose of pointing out the linguistic resources available in a particular text for making relationships between sentences. It may also be conducted to explain how texts are written to be understood. It is with this latter aspect in mind that the analysis of cohesion (Chapter Six) reported in this thesis incorporates notions such as homophora and paraphora which account for the relationship between the grammatical and lexical items of cohesion, and the context of situation of discourse in text analyses are incorporated into the description of the linguistic resources available for hanging sentences in texts in the genres together.

The implication for discourse organization of the three approaches to discourse structure adopted in this study is that they reinforce the fact that irrespective of similarities in subject-matter and field of discourse, a text that is appropriate in one context may be inappropriate in another (Blass, 1985). In other words, it points to the fact that different contexts constrain different patterns of discourse organization, be it at the level of the overall organization of discourse, or the level of organization of utterances or even the level of the relations of cohesion which hold between sentences. Therefore, a theory of discourse organization cannot consist simply of the quantification of discourse organizing elements. It must include principles which constrain writers organization of overall discourse structural units, organization of utterances and choice of elements by which sentences relate to each other in text.

Also in general terms, the results obtained in this study have implications for the concept of text variation in language use. In this regard, the study provides an alternative to the register analysis model of text variation which tends to group texts together on the basis of similarity of frequency of linguistic items. As Hopkins and Dudley-Evans (1988: 113) point out "one of the problems of many approaches to analysis has been that although they are capable of grouping texts together on the basis of similarity, they are rarely capable of differentiating on the basis of differences". This study provides an alternative to register analysis models by addressing two main issues which make classical register analysis inadequate for the solution of problems of text variation:

- 1 Problems with the delimitation of a descriptive corpus (e.g. Scientific texts).
- 2 Problems with the application of an analytical procedure on an inadequately differentiated corpus (e.g. Scientific English texts vs General English texts).

As an alternative, the approach adopted in this study is based on a delimitation of a textual corpus on the basis of contextual factors which constrain discourse production and reception defined in a way that reflects the realities of interaction. The textual corpus is then subjected to established qualitative procedures in discourse analysis and text linguistics which subsume within them the ways context determines various levels of text organization. That alternative procedure consists of the characterization of varieties of language in use as genre - a differentiation which Swales (1981) recognizes as a recognized communicative event with a shared public purpose and which, within varying degrees of freedom is structured and standardized. It also consists of the characterization of levels of discourse organization in terms of schematic structure, thematic progression and cohesion.

7.3.2 Some Wider Implications

The discussion presented in Section 7.3.1 above relates to the implications of this study to the general area of discourse analysis. In this section, we examine the specific implications which the study has for genre-analysis in English for Science and Technology (EST). First, the study is an addition to the studies which seek to characterize genres of EST discourse. As with most studies which adopt a genre-analysis approach, this study bases characterization of genre on the identification of schematic structures and the association of such schematic structures with specific underlying linguistic forms. Therefore, it conforms to the basic procedure in genre-analysis - to base the characterization of genre on the correlation between form and function. But beyond mere identification of schematic units, this study also characterizes the constituent elements or sub-moves which make up each move. It also characterizes moves into "optional" and "normally required moves". While these are fairly standard procedures in genre-analysis, to the best of my knowledge, they have never before been applied to the whole of the research article. Previous genre-based studies of research articles have tended to focus on specific sections, e.g. Introduction section only (Swales, 1981) or on the Discussion section only (McInlay, 1983; Dudley-Evans, 1986; Hopkins and Dudley-Evans, 1988).

The tendency to characterize genres only by correlating schematic units with their underlying linguistic forms does not reveal much about the interaction which goes on in genres. As Widdowson (1983: 102) points out, the danger in such an analysis is that it might lead us to suppose that form-function correlations are fixed and can be learned as formulae. However, unlike some other attempts at genre-analysis, this study includes an analysis of some interactive forms of discourse organization which are realized in genres such as thematic progression and cohesion. To this end, this study can be said to be an attempt at proposing a new direction to genre-analysis based on an integrated approach to the analysis of discourse structure in genres. It is also an attempt

at proposing a new approach to the identification of genres based on the concept of comparative/ contrastive analysis of related genres. The tendency to characterize aspects of texts, e.g. Article Introductions and Discussions, as genres leads to the comment by Widdowson (op cit: 101) that although Swales refers to his investigation as "genre specific", it is not entirely clear what the term genre is really meant to cover. Echoing the same views, Adams-Smith (1987: 12) contends that the definition of genre remains a problem, for while the concept is clear, it is not clear what its edges are. Illustrating, Adams-Smith cites a hypothetical example involving an AP article, appearing in the *New Scientist*, the *Journal of American Medical Association*, *Journal of Immunology* and the *Proceedings of the Immunopharmacology Society* respectively and wonders whether it is true to say that all five articles are instances of the same genre. Concluding, she contends that it is not easy to decide where to draw the boundaries between one genre and another.

The data used in this study certainly matches the example cited by Adams-Smith above. As evidence from the texts in the corpus has shown, each text in a triad of related texts bears the distinctive feature of a recognized communicative event with unique patterns of discourse organization. Therefore, this study is an answer to the question raised by Adams-Smith in that it indicates that it is not true to say that the five articles in her example are all instances of the same genre. Thus, the study is a contribution to the study of genre and the process of genre-analysis for it enables a much clearer understanding of what the term genre is meant to cover and what its edges are.

Finally, this study has practical relevance to the teaching of reading and writing. Pointing out the pedagogical rationale for investigations into the semantic organization of texts, Stubbs (1986: 50) states that:

We try at least to teach students explicitly about the organization of written language at the level of graphology, although our teaching may often be

inadequate. We do less well at the level of syntax, since there is as yet inadequate description of the differences in grammar between spoken and written English. But systematic teaching about semantic organization (including cohesion and paraphrase) and discourse (eg. narrative structure) is almost non-existent.

Therefore, this study contributes to the number of theoretical studies of the organization of discourse whose findings have obvious relevance to language teaching. In specific terms, the results obtained in this study have implications for the teaching of reading and writing in the sciences. Most importantly, the results are of great relevance to the teaching of English to non-native speakers who need a knowledge of how the language is used in Science and Technology, including medical science, in order to achieve the utmost in their fields of study and participate effectively as members of their respective discourse communities. This is not to suggest that the findings made in this study do not have implications for learners who are native speakers of English. In fact, Davies and Greene (1984) recommend aspects of the approach to text analysis adopted in this study to British Secondary School teachers in their book *Reading for Learning in the Sciences* based on *The Schools Council Project Reading for Learning in the Secondary Schools*. Therefore, this study has relevance for native and non-native speakers of English, though its implications appear to be greater for the latter than the former. However, since this study is a theoretical investigation into patterns of discourse organization, we can only speculate on the extent to which it can contribute to improving reading and writing in science and technology. Thus, results obtained in this study may contribute to learning in the sciences in the following ways:

- 1 Explicit knowledge of schematic units and their constituent elements might enable teachers to direct learners more accurately to regular patterns of information in text, thus enabling learners to learn to adopt the same procedure in their reading and writing tasks more consciously. And as Davies and

Greene (op cit) point out, it will also provide a framework for reading and learning and a framework for writing and notetaking.

2 Explicit knowledge of the concept of Theme/Rheme or "Given/New" and the notion of thematic progression will help teachers provide adequate explanations to learners when they produce sentences or paragraphs that are either not clear or that are discontinuous. The learners will be able to determine for themselves whether the sentences they produce are clear and whether there is continuity in their discourse. Bloor and Bloor (1987) have demonstrated, by means of writings by non-native speakers, practical ways in which the concept of Theme and Rheme can help improve writing quality.

3 A knowledge of the correct use of cohesive devices will enable learners to understand how texts are created through the relationships which exist between words.

The practical implications of this study for language teaching may be summed up with this observation by Holloway (1981: 215):

When they use the semantic theories described here, students can ask themselves, as they struggle to convey a particular idea in a sentence, "What is the old information I need to present so that I can tie this idea to what has gone before?" And when they look at a rough draft they can ask, "Have I reflected the connections of thought in my thesis statement implicitly or explicitly in the devices I have used for cohesion in the body of my paper?" or "Have the cohesive devices linked together not only the individual pieces of text but also my thoughts effectively?"

7.4 LIMITATIONS AND SUGGESTIONS FOR FURTHER RESEARCH

I set out in this study to throw some light on problems associated with genre variation and differentiation by showing how parallel genres differ in the way they organize discourse. It is for this reason that I adopted a comparative/contrastive approach as well as an eclectic approach to the analysis of research data. Both approaches have ensured that solutions have been provided to some of the persistent problems in genre-analysis. At the same time, they have exposed a number of weaknesses in a conceptual framework which attempts to characterize genres by reference to other genres to which they may relate. Some of the limitations which the adoption of the conceptual framework in this study has exposed may be rectified by future research which may decide to limit the number of genres and the variables in the analysis to two instead of the three as is the case in this study.

The second shortcoming of this study may be the limited size of the corpus. However, considering the fact that this study is proposing a new approach to the characterization of genre based on comparative and eclectic paradigms, it is only proper that the data be limited to a size at which meaningful comparison and application of variables of discourse structure can take place. Future studies, based on other forms of parallel texts and possibly statistical analysis on larger text samples, are clearly needed to validate the approach adopted in this study.

The third limitation of this study relates to the lack of absolute precision in the delimitation of Move and Constituent element boundaries often associated with schema-based analysis of discourse. This problem is not unique to this study. It can be associated with the nature of discourse analysis itself which tends to adopt approaches from various fields of learning (eg. cognitive psychology, artificial intelligence, etc), most of which were not originally designed for problems associated with text analysis. More research still needs to be done with a variety of text-types or genres if a solution is to be found to this problem.

The fourth limitation is related to the lack of sufficient and detailed empirical work of a "text as process" type to back up a number of assumptions made in this study. Thus, in most cases assumptions are based on intuitions gained from studies which have analysed "text as product". There is, therefore, a clear need for further research in the area of "text as process" to fill the existing gap in our knowledge of how contextual factors, including psychological ones, constrain the organization of written discourse.

In accordance with the broad limitations identified in the preceding discussion, I suggest the following as areas for further research:

- 1 The comparative/contrastive study of professional and popularized genres in other scientific and technological fields in order to determine whether the process of popularization is similar or different across scientific disciplines.
- 2 Cross-genre comparison of different types of popularized science texts, eg. popularizations in *Scientific American* versus popularizations in *New Scientist*; popularizations in *Newspapers* vs popularizations in quality magazines (eg. *The Economist* etc).
- 3 Cross-genre comparison of different types of professional science texts, eg. *News Articles* vs *Letters*; *Research Articles* vs *Review Articles* and vice versa.
- 4 Experimental studies with authors of professional and popularized tests as a means of determining the psycholinguistic dimensions of the organization of discourse and shedding more light on factors which constrain the choices which writers make in adapting some organizational strategies at certain points in text.

5 Experimental studies with EST learners to determine the extent to which an application of the approach adopted in this study and the findings therein can contribute to improving their ability to process texts in their subject areas.

6 Designing EST courses around the approaches and findings made in this study.

7 Studies are also needed to determine the extent to which the approach proposed in this study could be of use to error analysts and contrastive linguists, translators, interpreters, information scientists and computer specialists.

Al-Jabri, A.M. (1987) *Contrastive Text Analysis: A Study of English and Arabic*. Unpublished PhD thesis, University of Aston, Birmingham.

Anderson, R.V., Brockmann, H.J. & Wilson, C.H. (eds) (1978) *New Directions in Technical and Scientific Communication*. Amherst, Massachusetts: Fallout Press.

Anderson, R.V. & Pearson, D. (1974) 'A Information-theoretic View of Basic Processes in Reading Comprehension', in Pearson, P.D. (ed), *Handbook of Reading Research*. London: Longman.

Barber, C.S. (1932) *Arabic Grammar*. London: Cambridge University Press.

Barber, C.S. (1962) 'Some Morphological Characteristics of Modern Standard Arabic', in *Contributions to English Syntax and Phonology*, Stockholm: Almqvist and Wiksell.

REFERENCES

- Adams-Smith, D.** (1983), "Prescribing and General Practice: Style in Medical Journals", Paper presented at TESOL conference, February 1983.
- Adams-Smith, D.** (1984), "Medical Discourse: Aspects of Authors' Comment", *The ESP Journal*, Vol 3, no 1, pp25-36
- Adams-Smith, D.** (1986), "Aspects of Register Variation in Seven Popular Science Articles and Research Papers from which they were derived. Unpublished MA Dissertation, University of Birmingham
- Adams-Smith, D.** (1987a), "Variation in Field-Related Genres", in Dudley-Evans, T. (ed) *Genre Analysis and ESP, ELR Journal*, Vol I, pp10-32
- Adams-Smith, D.** (1987b) "The Process of Popularization-Rewriting Medical Research Papers for the Layman: Discussion Paper", *Journal of the Royal Society of Medicine*, Vol 80, pp634-636
- Al-Jabr, A.M.** (1987), Cohesion in Text Differentiation: A Study of English and Arabic. Unpublished PhD thesis, University of Aston in Birmingham
- Anderson, P.V., Brockmann, R.J. & Miller, C.R.** (eds), *New Essays in Technical and Scientific Communication: Research, Theory and Practice*, Farmingdale, N.Y: Baywood Publishing Company, Inc.
- Anderson, R.C. & Pearson, D.** (1984) "A Schema-theoretic View of Basic Processes in Reading Comprehension", In Pearson, P.D. (ed), *Handbook of Reading Research*, London: Longman
- Bartlett, F.C.** (1932) *Remembering*, London: Cambridge University Press
- Barber, C.L.** (1962), "Some Measurable Characteristics of Modern Scientific Prose, In *Contributions to English Syntax and Phonology*, Stockholm: Almqvist and Wiksell

Bazerman, C. (1983), "Scientific Writing as a Social Act: A Review of the Literature of the Sociology of Science", In Anderson, B.V; Brockmann, R.J. and Miller, C. R. (eds) *New Assays in Technical and Scientific Communication: Research, Theory and Practice*, Farmingdale, N.Y: Baywood Publishing Company, Inc

Bazerman, C. (1985) "Strategy-Laden Schema and Schema-Laden Strategies in the Reading of Physicists, *Written Communication*, Vol 2

Bazerman, C. (1988), *Shaping Written Knowledge: The Genre and Activity of the Experimental Article in Science*, Madison, Wisconsin: The University of Wisconsin Press.

Bellack, A.; Kliebard, H.M.; Hyman, R.T. & Smith, F.L. (1966) *The Language of the Classroom*, New York: Teachers College Press

Blakemore, D. (1988), "The Organisation of Discourse", In Newmeyer, F.J. (ed), *Linguistics: The Cambridge Survey IV, (Language: The Socio-Cultural Context)*, Cambridge, Cambridge University Press

Blass, R. (1985), "Cohesion, Coherence and Relevance, University College, London

Bley-Vroman, R. & Selinker, L. (1984), "Research Design in Rhetorical/Grammatical Studies: A Proposed Optimal Research Strategy". *English for Specific Purposes Newsletter*, issues 82/83/84, January, February, March

Bloor, M. & Bloor, T. (1987), "Application of the Theory of FSP to the Teaching of Writing", Paper presented at BAAL Seminar, Essex University

Brown, A.L.; Day, J.D. & Jones, R.S. (1983), "The Development of Plans for Summarizing Texts", *Child Development*, 54, pp968-979

Brown, G. & Yule G. (1983) *Discourse Analysis*, Cambridge: Cambridge University Press

- Brown, P. & Levinson, S. C.** (1987), *Politeness..* Cambridge, Cambridge University Press.
- Bruce, N.J.** (1983), "Rhetorical Constraints and Information Structure in Medical Research Report Writing.
- Burkett, D.W.** (1973), *Writing Science News for the Mass Media*, Houston, Texas: Gulf Publishing Company
- Candlin, C.N.; Brutton, D.J.; Leather, J.L. & Woods, E.** (1977), "Doctor-Patient Communication Skills," Graves Medical Audio-Visual Library
- Carrel, P.L.** (1983), Some Issues in Studying the Role of Schemata, or Background Knowledge in Second Language Comprehension. *Reading in a Foreign Language*, 1, 2, pp.81-92.
- Carter, R.** (1987), *Vocabulary: Applied Linguistic Perspectives*, London: Allen and Unwin
- Catford, J.C.** (1965), *A Linguistic Theory of Translation: An Essay in Applied Linguistics*, Oxford: The University Press
- Celce-Murcia, M. & Larsen-Freeman, D.** (1983), *The Grammar Book: An ESL/EFL Teacher's Course*, MA: Newbury House Publishers
- Chafe, W.L.** (1976), Givenness, Contrastiveness, Definiteness, Subjects, Topics and Points of View. In Li C.N. (ed) *Subject and Topic*, New York, Academic Press
- Chafe, W.L.** (1982), "Integration and Involvement in Speaking, Writing and Oral Literature", In Tannen, D. (ed) *Spoken and Written Language*, New Jersey: Ablex Publishing Company
- Coates, J.** (1983), *The Semantics Of the Modal Auxillaries*, London: Croom-Helm.
- Cooper, C.** (1985), Aspects of Article Introductions in IEEE Publications. Unpublished MSc Dissertation, The University of Aston in Birmingham

- Cooper, C.R. & Greenbaum, S.** (1986), *Studying Writing: Linguistic Approaches*, Beverley Hills: Sage Publications
- Cooper, M.M.** (1986), "The Ecology of Writing", *College English*, 48, pp364-375
- Cooray, M.** (1965), "The English Passive Voice", *English Language Teaching*, Vol 21, 3, pp203-210
- Comrie, B.** (1985), *Tense*, London: Cambridge University Press.
- Couture, B.** (1986), "Bridging Epistemologies and Methodologies in Written Language Function". In Couture, B. (ed), *Functional Approaches to Writing Research Perspectives*, London: Francis Pinter
- Cremmins, E.T.** (1982), *The Art of Abstracting*. Philadelphia: ISI Press
- Crooks, C.** (1984), "Towards a Validated Discourse Analysis of Scientific Texts". Unpublished MA Dissertation, University of Birmingham
- Curzon, D.** (1984), "Themes, Hyperthemes and the Discourse Structure of British Legal Texts. *TEXT*, Vol 4-1/3, pp31-56
- Dahl, O.** (1974), "Topic-Comment Structure in a Generative Grammar with a Semantic Base. In Danes, F. (ed), *Papers in Functional Sentence Perspective*, Prague: Academia
- Danes, F.** (1949), "The Intonation of the Interrogative Sentence", *Nase Rec*
- Danes, F.** (1964), "A Three-Level Approach to Syntax, *Travaux Linguistique de Prague*, 1, pp225-240
- Danes, F.** (1974), "Functional Sentence Perspective and the Organization of the Text". In Danes, F. (ed), *Papers in Functional Sentence Perspective*, Prague: Academia
- Danes, F.** (1970), "One Instance of Prague School Methodology: Functional Analysis of Utterance and Text". In Garvin, P.L. (ed), *Method and Theory in Linguistics*, The Hague: Mouton

- Davies, R. & Sklair, L. (1972), "Science and Technology Coverage in Four British Newspapers", 1949-69", Mimeo, London School of Economics
- Davies, F. & Greene, T. (1984), *Reading for Learning in the Sciences*, Edinburgh: Oliver and Boyd
- deBeaugrande, R. & Dressler, W.U. (1981), *Introduction to Textlinguistics*. London: Longman
- Dubois, B.L. (1981), "The Construction of Noun Phrases in Biomedical Journal Articles". In Lundquist, L.; Pincht, H., and Quistgaard, J. (eds), *Pragmatics and LSP*, Copenhagen
- Dubois, B.L. (1986), "From "New England Journal of Medicine" and "Journal of the American Medical Association" through the "Associated Press" to Local Newspaper: Scientific Translation for the Laity. *Wissenschaftssprache Und Gessellschaft*. Hamburg, Edition Akademie
- Dubois, B.L. (1987), "A Reformulation of Thematic Progression Typology. *TEXT*, 7(2), pp89-116
- Dudley-Evans, T. (1986), "Genre-Analysis: An Investigation of the Introduction and Discussion Sections of MSc Dissertations" in Coulthard, M. (ed.) *Talking About Text*. Discourse Analysis Monograph No.13, ELR. University of Birmingham
- Dudley-Evans, T. (1987) (ed), *Genre-Analysis and ESP*, ELR Journal, Vol I, pp1-9
- Edmonson, W. (1981), *Spoken Discourse*, London: Oxford University Press
- Eiler, M.A. (1986), Thematic Distribution as a Heuristic for Written Discourse Function. In Couture, B. (ed), *Functional Approaches to Writing Research Perspectives*: London: Francis Pinter
- Ellis, J. & Ure, E. (1969), "Language Variety: Register". In Meetham, A.R. (ed), *Encyclopedia of Linguistics, Information and Control*, London: Penguin press
- Farago, P. (1976), *Science in the Media*, London, Oxford University Press

- Ferguson, C.A.** (1983), "Sports Announcer Talk: Syntactic Aspects of Register Variation", *Language in Society*, 12, pp.153-172
- Firbas, J.** (1964), "On Defining Theme in Functional Sentence Analysis". *Travaux Linguistique de Prague*, Vol I, pp267-280
- Firbas, J.** (1974), "Some Aspects of the Czechoslovak Approach to the Problem of Functional Sentence Perspective". In Danes, F. (ed), *Papers in Functional Sentence Perspectives*, Prague: Academia
- Firbas, J.** (1975), "On "Existence/Appearance on the Scene" in Functional Sentence Perspective". *Prague Studies in English*, Vol XVI, pp 47-70
- Firbas, J.** (1976), "A Study in the Functional Perspective of the English and the Slavonic Interrogative Sentence. *BRNO Studies in English*, Part 12, pp9-57
- Firbas, J.** (1986), "Non-Thematic Subjects in Contemporary English. *Travaux Linguistiques de Prague*. Vol 12, pp239-256
- Fowler, R. & Kress, G.** (1979), "Critical Linguistics". In Fowler, R., Hodge, B., Kress, G. and Trew, T. (eds), *Language and Control*, London: Routledge and Kegan Paul
- Freed, R.C. & Broadhead, G.J.** (1987), "Discourse Communities, Sacred Texts, and Institutional Norms". *College Composition and Communication*, 38, pp155-165
- Freeman, S.H.** (1987), "Verbal Communication in Medical Encounters: An Overview of Recent Work". *TEXT*, 7(1), pp003-017
- Freeman, S.H. & Heller, M.S.** (1987), "Editor's Preface". *TEXT*, 7(1), pp001-002
- Fries, P.H.** (1981), "On the Status of Theme in English: Arguments from Discourse. *Forum Linguistium*, Vol VI, No I, pp1-38

- Fum, D., Guida,, G. & Tasso, C.** (1986), "Tailoring Importance Evaluation to Reader's Goals: A Contribution to Descriptive Text Summarization. In *Proceedings of Coling '86 - 11th International Conference on Computational Linguistics*. Bonn: University of Bonn, pp256-259
- Funkhouser, G.R. & Macoby, N.** (1970), *Communicating Science to Non-Scientists*. Institute of Communication Research Report, Stanford University, USA.
- Ghadessey, M.** (1983), "Information Structure in Letters to the Editor". *IRAL*, Vol 21, 1, pp46-56
- Gilbert, G.N. & Mulkay, M.** (1984), *Opening Pandora's Box: A Sociological Analysis of Scientists' Discourse*. Cambridge: Cambridge University Press
- Giora, R.** (1983), "Segmentation and Segment Cohesion: On the Thematic Organization of Text". *TEXT*, Vol 3(2), pp155-181
- Givon, T.** (1983), "Topic Continuity in Spoken English". In Given, T. (ed), *Topic Continuity in Discourse: A Quantitative Cross- Language Study*. Amsterdam: John Benjamin Publishing Co
- Glaser, R.** (1982), "The Stylistic Component of Language for Specific Purposes", *Fachsprache*
- Glaser, R.** (1975), "Emotive Features in Technical and Scientific English". In Ringbom, H. (ed), *Style and Text*, Stockholm: Skriptor
- Gleason, H.A.** (1965), *Linguistics and English Grammar*. New York: Holt, Rinehart and Winston
- Gleason, H.A.** (1968), "Contrastive Analysis in Discourse Structure". In Alatis, J. (ed), *Report on the 19th Annual Round Table Meeting on Linguistics and Language Studies: Contrastive Linguistics*, Washington DC: Georgetown University Press
- Gopnik, M.** (1972), *Linguistic Structures in Scientific Texts*. The Hague: Mouton

- Greatz, N.** (1982), "Teaching EFL Students to Extract Structural Information from Abstracts", Paper Presented at *The International Symposium on LSP*, Eindhoven, The Netherlands.
- Gregory, M.** (1967), "Aspects of Varieties Differentiation". *Journal of Linguistics*, Vol 3, pp177-198
- Gregory, M. & Carroll, S.** (1978) *Language and Situation*, London: Routledge and Kegan Paul
- Grimes, J.E.** (1975), *The Thread of Discourse*. The Hague: Mouton
- Grimes, J.E.** (1977), "Narrative Studies in Oral Texts". In Dressler, W.U. (ed), *Current Trends in Textlinguistics*, Berlin: Walter de Gruyter
- Gutwinski, W.** (1976), *Cohesion in Literary Texts*. The Hague: Mouton
- Halliday, M.A.K.** (1967), "Notes on Transitivity and Theme in English, Part II". *Journal of Linguistics*, Vol 3, pp67-68
- Halliday, M.A.K.** (1968), "The Users and Uses of Language". In Fishman, J. (ed), *Readings in the Sociology of Language*, The Hague: Mouton
- Halliday, M.A.K.** (1979), *Language as Social Semiotic*. London: Arnold
- Halliday, M.A.K.** (1985a), *An Introduction to Functional Grammar*. London: Robert Arnold
- Halliday, M.A.K.** (1985b), "Language Across the Culture", Paper presented at *The SAEMEO RELC Conference*, Singapore
- Halliday, M.A.K. & Hasan. R.** (1976), *Cohesion in English*, London: Longman
- Halliday, M.A.K., Stevens, P. & McIntosh, A.** (1964), *The Linguistic Sciences and Language Teaching*. London: Oxford University Press

- Halliday, M.A.K. & Hasan, R.** (1986), *Language, Context and Text: Aspects of Language in a Social-Semiotic Perspective*. Victoria: Deakin University Press
- Hanania, E.A.S. & Akhtar, K.** (1985), "Verb Form and Rhetorical Function in Science Writings: A Study of MS Thesis in Biology, Chemistry and Physics". *ESP Journal*, Vol pp48-88
- Hatim, B.** (1981), *Discourse Analysis Models: Towards an Applied Linguistic Definition of Text Variation*. Unpublished PhD Thesis, University of Exeter
- Hartman, R.R.K.** (1980), *Contrastive Textology: Comparative Discourse Analysis in Applied Linguistics*. Heidelberg: Julius Groos Verlag
- Hartman, R.R.K.** (1987), "Contrastive Textology and Bilingual Lexicography". In Monaham, J. (ed), *Grammar in the Construction of Texts*. London: Francis Pinter
- Hasan, R.** (1984), "Coherence and Cohesive Harmony". In Flood, J. (ed) *Understanding Reading Comprehension: Cognition, Language and the Structure of Prose*. USA International Reading Association
- Hasan, R.** (1985), "The Structure of a Text", in Halliday, M.A.K. & Hasan, R., *Language, Context, and Text: Aspects of Language in a Social-Semiotic Perspective*, Victoria: Deakin University Press
- Hein, N. & Wodak, R.** (1987), "Medical Interviews in Internal Medicine: Some Results of an Empirical Investigation", *TEXT* 7(1) pp37-66
- Hemminki, E.** (1982), "Quality of Clinical Trials -- A Concern of Three Decades." *Methods in Information in Medicine*, 21, pp81-85
- Heslot, J.** (1982), "Tense and Other Indexical Markers in the Typology of Scientific Texts in English". In Hoedt, J. et al (eds), *Pragmatics and LSP: Proceedings of the 3rd European Symposium on LSP*. Copenhagen: LSP Centre/UNESCO
- Hibberd, P. L. & Meadows, A. J.** (1980) "Information Contained in Clinical Trial Reports." *Journal Of Information Science*, 2, pp165-168
- Hill, T.** (1958), "Institutional Linguistics", *Orbis*, 7 No 441

- Hoey, M.** (1979), "Signalling in Discourse", *Discourse Analysis Monograph*, No 6, ELR, University of Birmingham
- Hoey, M.** (1980), "Viewing Discourse as Dialogue as an aid to English Language Teaching", *ASPECTS* (Journal of the Panhellenic Union of State Teachers of English), Vol 5, 3, pp21-24
- Hoey, M.** (1983), *On the Surface of Discourse*. London: George Allen and Unwin
- Hoey, M. & Winter, E.** (1986), "Clause Relations and the Writers Communicative Purpose". In Couture, B. (ed), *Functional Approaches to Writing*. London: Ablex Pulo and Francis Pinter
- Holloway, D.W.** (1981), "Semantic Grammars: How They Can Help Us Teach Writing. *College Composition and Communication*, No 32, pp205-217
- Hopkins, A.** (1985), "An Investigation into the Organizing and Organizational Features of Published Conference Papers. Unpublished MA Dissertation, University of Birmingham
- Hopkins, A. & Dudley-Evans, T.** (1988), "A Genre-based Investigation of the Discussion Sections in Articles and Dissertations". *English for Specific Purposes*, Vol 7, No 2, pp113-121
- Huckin, T. & Olsen, L.** (1984), "On the Use of Informants in LSP Discourse Analysis". In Pugh, A. and Uljin, J. (eds), *Reading for Professional Purposes*. London: Heinemann
- Huddleston, R.D.** (1971), *The Sentence in Written English: A Study Based on an Analysis of Scientific Texts*. Cambridge: Cambridge University Press
- Huddleston, R.D.** (1984), *Introduction to the Grammar of English*, Cambridge: Cambridge University Press
- Hudson, R.** (1984), *Word Grammar*. Oxford: Basil Blackwell

- Hutchins, J.** (1977), "On the Structure of Scientific Texts", *Journal of Linguistics*, No 5, pp18-39
- Jacoby, S.** (1987), Reference to Other Researchers in Library Research Articles. In Dudley-Evans, T. (ed), "Genre-Analysis and ESP", *ELR Journal*, Vol I, University of Birmingham
- Jarvis, J.** (1983), "An ESP Teachers Guide to Functional Analysis". *The ESP Journal*, Vol 2, 1, pp101-112
- Jones, G., Connell, I. & Meadows, J.** (1978), *The Presentation of Science by the Media*, Primary Communications Research Centre, University of Leicester
- Jones, L.K.** (1977), *Theme in English Expository Discourse*. Illinois: Jupiter Press
- Jones, L.B. & Jones, L.K.** (1985), "Discourse Functions of Five English Sentence Types. *WORD*, Vol 36, No I, pp1-19
- Jordan, M.P.** (1982), "Structure, Meaning and Information Signals of Some Very Short Texts: An Introductory Analysis of Everyday English Prose". *The Eighth LACUS FORUM*, pp410-417
- Jordan, M.P.** (1984), *Rhetoric of Everyday Prose*, London: George Allen unwin
- Jordan, M.P.** (1985), "Towards an Integrated Theory of Textual Cohesion", *MALS Journal*, Vol 10, pp9-40
- Jordan, M.P.** (1986), "Short Texts to Explain the Problem-Solution Structure and Vice Versa". *Instructional Science*, Vol 9, pp221-275
- Kahn, J.Y.** (1983), *Modes of Medical Instruction: A Semantic Comparison of Textbooks of Medicine and Popular Home Medical Books*. Berlin: Mouton Publishers
- Kamp, H.** (1981), "A Theory of Truth and Semantic Representation: In Groenendijk et al (eds), *Formal Methods in the Study of Language*. Amsterdam: John Benjamin

- Karpf, A.** (1988), *Doctoring the Media: The Reporting of Health and Medicine*. London: Routledge
- Keenan, E.O. & Schieffelin, B.B.** (1976), "Topic as a Discourse Notion. In Li, C.N. (ed), *Subject and Topic*. New York: Academic Press
- Kennedy, C. & Bolitho, R.** (1984), *English for Specific Purposes*, London: MacMillan
- Kintsch, W. & Van Dijk, T.A.** (1978), "Toward a Model of Discourse Comprehension and Production". *Psychological Review*, 85, pp363-394
- Kintsch, W. & Van Dijk, T.A.** (1975), Recalling and Summarizing Stories. *Languages*, 40, pp.98-116.
- Knorr-Cetina, K.D.** (1981), *The Manufacture of Knowledge: An Essay on the Constructivist and Contextual Nature of Science*. Oxford: Pergamon Press
- Knorr-Cetina, K.D. & Mulkay, M.** (1983) (eds), *Science Observed: Perspectives on the Social Study of Science*. London: Sage
- Kraigher-Vesel, N.** (1985), "Application of Discourse Analysis in Materials Design". *Paper presented at 5th European Symposium of LSP*, Leuven
- Kress, G.** (1983), "Linguistic and Ideological Transformation in News Reporting". In Davies, H. and Walton, P. (eds), *Language, Image, Media*. Oxford: Basil Blackwell
- Kreighbaum, H.** (1967), *Science and the Mass Media*. New York: New York University Press
- Lachowicz, D.** (1981) "On the Use of Passive Voice for Objectivity, Author Responsibility and Hedging in EST". *Science of Science*, No 2 (6), Vol 2
- Lackstrom, J.E.** (1977), "The Comprehension of EST: Arguments and Definitions". *Studies in Language Learning*, Special issue on LSP. Vol 2a

- Lackstrom, J.E., Selinker, L. & Trimble, L.P. (1972), "Grammar and Technical English". *English Teaching Forum*, Vol X, No 5
- Lamb, S. M. (1966), *Outline of Stratificational Grammar*. Washington D. C., Georgetown University Press.
- Leech, G. (1966), *English in Advertising*, London: Longman
- Li, C.N. & Thompson, S.A. (1976), "Subject and Topic: A New Typology of Language". In Li, C.N. (ed), *Subject and Topic*, New York: Academic Press
- Longacre, R. & Levinsohn, S. (1977), "Field Analysis of Discourse". In Dressler, W.U. (ed), *Current Trends in Textlinguistics*, Berlin: Walter de Gruyter
- Lopez, G.S. (1982), Article Introductions in Spanish: A Study in Comparative Rhetorics. Unpublished MSc Dissertation, University of Aston in Birmingham
- Lowe, I. (1987), "Sentence Initial Elements in English and their Discourse Functions". *Occasional Papers in Systemic Linguistics*, Vol 2, pp 5-33.
- Malcolm, L. (1987), What Rules Govern Tense Usage in Scientific Articles? *English for Specific Purposes*. Vol 6, No I, pp31-43
- Mandler, J.M. (1978), A Code in the Node: The Use of Story Schema in Retrieval. *Discourse Processes*, Vol I, pp14-35
- Maseide, P. (1987), "The Permanent Context Construction: A Neglected Dimension of Therapeutic Discourse. *TEXT*, Vol 7(1), pp67-87
- Mathesius, V. (1939), "On So-called Functional Sentence Perspective", *Slovo a slovensnort*. Vol 7, pp177-174
- Mathesius, V. (1941), "The Basic Function of Word Order in Czech". *Slovo a slovensnort*, Vol 7, pp169-180

- Mathesius, V.** (1942), "From Comparative Word Order Studies". *Casopis Pro Modernic Filologii*, Vol 28, pp181-190
- Maynard, S.K.** (1986), "Interactional Aspects of Thematic Progression in English Casual Conversation, *TEXT*, 6(1), pp73-105
- McInley, J.** (1983), "An Analysis of the Discussion Section of Medical Journal Articles. Unpublished MA Dissertation, University of Birmingham
- McPhill, A.** (1911), "Style and Medical Writing". *Canadian Medical Journal*, pp70-73
- Meadows, J.** (1987), "The Growth of Science Popularization: A Historical Sketch. *IMPACT of Science on Society*, No 144, pp341-346
- Meyer, P.G.** (1987), "Some Observations on the Signalling of Structure in Technical Discourse". In Monahan, J. (ed), *Grammar in the Construction of Texts*: London: Francis Pinter
- Meyer, B.J.F.** (1975), *The Organisation of Prose and its Effects on Memory*. Amsterdam: North Holland Publishing Company
- Minsky, M.** (1975), "A Framework for Representing Knowledge. In Winston, P. (ed), *The Psychology of Computer Vision*. New York: McGraw Hill
- Mosenthal, J. H. & Tierney, R. J.** (1984), "Cohesion: Problems with Talking about Text." *Reading Reaserch Quaterly*, 19, pp240-244
- Morrison, K.** (1981), "Some Properties of Telling-Order Designs in Didactic Inquiry". *Philosophy of the Social Sciences*. 11, pp245-262
- Myers, G.** (1985a), "The Social Construction of Two Biologists' Proposals". *Written communication*, Vol 2, No 3, pp219-45
- Myers, G.** (1985b), "Texts as Knowledge Claims: The Social Construction of Two Biology Articles". *Social Studies of Science*. Vol 15, pp593-630

- Myers, G.** (1986), "The Social Construction of Popular Science: The Narrative of Science and the Narrative of Nature". Paper read at an ELR Seminar, University of Birmingham. In Myers, G. (in press), *Writing Biology: Texts in the Social Construction of Science*. Madison: University of Wisconsin Press
- Myers, G.** (1988), "Lexical Cohesion and Readers' Knowledge in Science and Popular Science Texts. *Paper presented at an LSU Seminar*, University of Aston in Birmingham
- Myers, G.** (1989), "The Pragmatics of Politeness in Scientific Articles". *Applied Linguistics*, 10, 1, pp.1-35
- Myers, G.** (in press), *Writing Biology: Texts in the Social Construction of Science*. Madison: University of Wisconsin Press
- Nearle, S.** (1983), *Genre*, London: British Film Institute
- O'Connor, M. & Woodford, F.F.** (1976), *Writing Scientific Papers in English*. North-Holland
- Perfetti, C.A. & Goldman, S.R.** (1974), "Thematization and Sentence Retrieval". *Journal of Verbal Learning and Verbal Behaviour*, 13, pp70-79
- Perfetti, C. A. & Goldman, S. R.** (1975), "Discourse Functions of Thematization and Topicalization ." *Journal of Psycholinguistic Research*, Vol. 4, No. 3, pp 257-271.
- Pettinari, C.** (1983), "The Function of Grammatical Alternation in Fourteen Surgical Reports". *Applied Linguistics*, Vol 4, No I, pp
- Pettinari, C.** (1982), "A Comparison of the Production of Surgical Reports by Native and Non-native Speaking Surgeons. *Paper presented at Ninth International Systemic Workshop*, Toronto, Canada
- Pindi, M.M.K.** (1988), "Schematic Structure and the Modulation of Propositions in Economics Forecasting Text". Unpublished PhD thesis, The University of Aston in Birmingham

- Pinkal, M.** (1986), "Definite Noun Phrases and the Semantics of Discourse". In Proceedings of Coling 1986 (11th International Conference on Computational Linguistics) Bonn: University of Bonn
- Prince, E.F.** (1981), "Toward a Taxonomy of Given-New Information". In Cole, P. (ed), *Radical Pragmatics*, New York: Academic Press
- Proctor, M.** (1988), "Discourse Organization Patterns and their Signals". In Turney, A. (ed), *Applied Text Linguistics: Six contributions from Exeter*. Exeter Linguistics Studies, Vol 13, University of Exeter
- Quirk, R., Greenbaum, S., Leech, G. & Svartvik, J.** (1985) *A Grammar of Contemporary English*. London: Longman
- Reddick, R.J.** (1986), "Textlinguistics, Text Theory and Language Users". *WORD*, Vol 37, No 1-2, pp31-43
- Regent, O.** (1985), "A Comparative Approach to the Learning of Specialized Written Discourse". In Riley, P. (ed), *Discourse and Learning*, London: Longman
- Reid, T.B.W.** (1956), "Linguistics, Structuralism and Philology. *Archivum Linguisticum*, Vol 8 (2), No 28
- Rounds, P.L.** (1981), "On Hedging in Social Science Written Texts". *Linguistics* 564, University of Michigan, Ann Arbor
- Rumelhart, D.E.** (1975), "Notes on a Schema for Stories". In Bobrow, D.G. and Collins, A.M. (eds), *Representation and Understanding: Studies in Cognitive Science*. New York: Academic Press
- Rumelhart, D.E.** (1980), "On Evaluating Story Grammars". *Cognitive Science*, Vol 4, pp313-316
- Rumelhart, D.E.** (1981), *Understanding Understanding (Technical Report CHIP 100)*, La Jolla, CA: University of California, San Diego, Centre for Human Information Processing

- Rumelhart, D.E. & Ortony, A.** (1977), "The Representation of Knowledge in Memory. In Anderson, R.C., Spiro, R.J., and Montague, W.E. (eds), *Schooling and the Acquisition of Knowledge*, Hillsdale, N.J: Lawrence Erlbaum Associates
- Rumelhart, D.E. & Norman, D.A.** (1978), "Accretion, Tuning and Restructuring: Three Modes of Learning". In Cotton, J.W. and Klatzy, R.L. (eds), *Semantic Factors in Cognition*. Hillsdale, N.J: Lawrence Erlbaum Associates
- Rumelhart, D.E.** (1987), "Representation of Knowledge". In Aitkhead, A.M. & Slack, J.M. (eds.) *Issues in Cognitive Modelling*, London: Lawrence Erlbaum Associates/The Open University
- Salager, F.** (1983), "The Lexis of Fundamental Medical English: Classifactory Framework and Rhetorical Function (A Statistical Approach)". *Reading in a Foreign Language*, Vol I, No I, pp54-64
- Salager, F.** (1985), "Specialist Medical English Lexis: Classificatory Framework and Rheotrical Function - a Statistical Approach. *EMP Newsletter*, Vol 2, No 2, pp5-17
- Schank, R.C. & Abelson, R.** (1977), *Scripts, Plans, Goals and Understanding*. Hillsdale, N.J: Lawrence Erlbaum Associates
- Schiffrin, D.** (1987), *Discourse Markers*, Cambridge: Cambridge University press
- Selinker, L.** (1979), "On the Use of Informants in Discourse Analysis and Language for Specialized Purposes". *IRAL*, Vol XVII, No 3, pp189-215
- Selinker, L.** (1972), "Interlanguage". In *International Review of Applied Linguistics*, Vol X, No 3.
- Selinker, L., Trimble, L. & Vroman, R.** (1974), *Working Papers in English for Science and Technology*. College of Engineering, Seattle, University of Washington
- Selinker, L., Trimble, L. & Trimble, M.T..** (1978), "Rhetorical Function Shifts in EST Discourse". *TESOL Quarterly*, No 12

- Selinker, L., Tarone, E. & Hanzeli, V. (1981) (eds), *English For Academic and Technical Purposes: Studies in Honour of Louis Trimble*. Rowley, Mass: Newbury House
- Sgall, P., Hajikova, E. & Benesova, E. (1973), "Remarks on the Topic/Comment Articulation, II, *PBML*, 19, pp29-58
- Sharafuddin, A.M. (1987), "Science Popularization: A View from the Third World". *IMPACT of Science on Society*, No 144, pp347-353
- Sherrard, C.A. (1986), "Summary-Writing: A Topographical Study" *Written Communication*, 3, pp234-343
- Sinclair, J., McH. & Coulthard, R.M. (1975), *Towards an Analysis of Discourse: The English Used by Teachers and Pupils*. London: Oxford University Press
- Smith, R.N. & Frawley, W.J. (1983), "Conjunctive Cohesion in Four English Genres". *TEXT*, Vol 3-4, pp347-374
- Stevens, P. (1973), "Technical, Technological and Scientific English. *ELT Journal*, 27, 3, pp223-234
- Stubbs, M. (1986), *Educational Linguistics*. Oxford: Basil Blackwell
- Swales, J. (1976), "Verb Frequencies in English", *ESPMENA Bulletin*, 4, pp28-31
- Swales, J. (1981a), "Aspects of Article Introductions". *ESP Research Report*, No 1, The University of Aston in Birmingham
- Swales, J. (1981b), "The Function of One Type of Particle in a Chemistry Textbook". In Selinker, L. et al (eds), *English for Academic and Technical Purposes*. Rowley, Mass: Newbury House Publishers
- Swales, J. (1985), *Episodes in ESP*, London: Pergamon

- Swales, J.** (1986), *Language and the Professions: Current Developments and Future Prospects*. In Kaplan, R. (ed), *Annual Review of Applied Linguistics*, Cambridge: Cambridge University Press
- Swales, J.** (1987), "Communication Skills for the Public Sector". In *Proceedings of the Southern Africa Conference on Communication Skills for the Public Sector*, Lusaka: National Institute of Public Administration
- Tarone, E., Dwyer, S., Gillette, S. & Icke, V.** (1985), "On the Use of the Passive in Two Astrophysics Journal Papers". In Swales, J., *Episodes in ESP*. Oxford: Pergamon
- Tarone, E.** (1988), *Variation in Interlanguage*, London: Edward Arnold
- Taylor, D.E.** (1975), "The Transfer of Information from Journal to the Mass Media". Unpublished MSc Dissertation, City University, London
- Travenicek, F.** (1951), *Grammar of Standard Czech*. Prague
- Trimble, L.** (1985), *English for Science and Technology: A Discourse Approach* Cambridge University Press.
- Urquhart, A.H.** (1983), "The Effect of Rhetorical Ordering on Readability. In Alderson, C.J. and Urquhart, A.H. (eds), *Reading in a Foreign Language*, London: Longman
- Urquhart, A.H.** (1986), "Functional Analysis of Scientific Abstracts", College of St marks and St John, Mimeo
- Vachek, J.** (1975), (ed), *Valem Mathesius: A Functional Analysis of Present Day English on a General Linguistic Basis*. The Hague: Mouton
- Vande Kopple, W.J.** (1983), "Something Old, Something New". *Research in the Teaching of English*, 17, pp85-99
- van Dijk, T.A.** (1980), *Macrostructures*, Hillsdale, N.J: Lawrence Erlbaum

- van Dijk, T.A. (1985), "Structure of News in the Press". In VanDijk, T.A. (ed), *Discourse and Communication (New Approaches to the Analysis of Mass Media Discourse and Communication.)* New York: Walter de Gruyter
- van Dijk, T.A. (1986), "News Schemata". In Cooper, C.R. and Greenbaum, S. (eds), *Studying Writings Linguistic Approaches*. London: Sage Publication.
- Varantola, K. (1987), "Popularization Strategies and Tense Functional Shifts in Scientific / Technical Writing." in *UNESCO AISED-LSP Newsletter*, 10, 2, pp 33 - 52.
- Weill, H. (1884), *The Order of Words in the Ancient Languages Compared with that of the Modern Languages*. Boston; re-edited Amsterdam, 1978
- Weissberg, R.C. (1984), "Given and New: Paragraph Development Models from Scientific English". *TESOL Quarterly*, Vol 18, No 3, pp
- Weizmann, E. (1984), "Some Register Characteristics of Journalistic Language: Are they Universal? *Applied Linguistics*, Vol 5, No I, pp
- Wellek, R. & Warren, A. (1973), *Theory of Literature*, Middlesex: Penguin Books
- West, G.K. (1980), "That-Nominal Constructions in Traditional Rhetorical Divisions of Scientific Research Papers. *TESOL Quarterly*, Vol XIV, No 4, pp483-488
- Widdowson, H.G. (1979), *Explorations in Applied Linguistics*, London: Oxford University Press
- Widdowson, H.G. (1983), *Learning Purpose and Language Use*. London: Oxford University Press
- White, R.V. (1974), "The Concept of Register and TESOL". *TESOL Quarterly*, 8, 4, pp401-416

- Wingard, P.** (1981), "Some Verb Forms and Functions in Six medical Texts". In Selinker, L. et al (eds), *English for Academic and Technical Purposes*. Rowley, Mass: Newbury House Publishers
- Winter, E.O.** (1971), "Connection in Science Materials". *Papers and Reports No 7* London: Centre for Information on Language Teaching
- Winter, E.O.** (1976), "Fundamentals of Information Structure". Hartfield Polytechnic: Mimeo
- Winter, E.O.** (1977), "A Clause-Relational Approach to English Texts: A Study of Some Predictive Lexical Items in Written Discourse". *Instructional Science*, 6, 1, pp1-92
- Winter, E.O.** (1982), *Towards a Contextual Grammar of English*, London: George Allen
- Witte, S.P.** (1983), "Topical Structure and Revision". *College Composition and Communication*, Vol 34, 3, pp313-341
- Witte, S.P. & Faigley, L.** (1981), "Coherence, Cohesion and Writing Quarterly". *College Composition and Communication*, Vol 32, pp189-204
- Wulf, F.** (1938), [Uber die Veranderung von Vorstellungen (Gedachtnis und Gestalt)] In Ellis, W.D. (translated and condensed), *A Source Book of Gestalt Psychology*, London: Routledge and Kegan Paul.