TOWARDS A VALIDATED DISCOURSE
ANALYSIS OF SCIENTIFIC TEXT

Graham Crookes
In recent years, a number of graduate students in the Department of English as a Second Language have selected the thesis option as part of their Master of Arts degree program. Their research has covered a wide range of areas in second language learning and teaching. Many of these studies have attracted interest from others in the field, and in order to make these theses more widely available, selected titles are now published in the Occasional Paper Series. This series, a supplement to the departmental publication Working Papers, may also include reports of research by members of the ESL faculty. Publication of the Occasional Paper Series is underwritten by a grant from the Ruth Crymes Scholarship Fund. A list of available titles and prices may be obtained from the department and is also included in each issue of Working Papers.

The reports published in the Occasional Paper Series have the status of "progress reports", and may be published elsewhere in revised form.

Occasional Paper #4 is an MA thesis by Graham Crookes. His thesis committee members were Jack Richards (chair), Craig Chaudron, Michael H. Long. This work should be cited as follows:

ABSTRACT

This study reviewed analyses of the discourse structure of scientific experimental-article introductions, and subjected a well-known one (Swales 1981) to empirical test.

A corpus of 96 articles was selected by stratified random sampling from 12 journals identified according to post-1980 citation frequency. Raters coded one quarter of the corpus according to the original Swales system, and found the system inadequate. Following revision of the system, 75% of a second quarter of the corpus was able to be rated at a satisfactorily high level of interrater agreement. Results were consistent with the existence of four basic units of discourse. Some subject-specific variation in structure was found. Social science article introductions were more complex than others.

Implications for EST materials writing are drawn. The importance of empirical validation of discourse analyses is stressed.

# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>iii</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>vii</td>
</tr>
<tr>
<td>CHAPTER I</td>
<td>1</td>
</tr>
<tr>
<td>CHAPTER II</td>
<td>23</td>
</tr>
<tr>
<td>CHAPTER III</td>
<td>29</td>
</tr>
<tr>
<td>CHAPTER IV</td>
<td>41</td>
</tr>
<tr>
<td>APPENDIX</td>
<td>50</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>58</td>
</tr>
</tbody>
</table>
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Article introductions from the first quarter of the corpus conforming to the original Swales model</td>
<td>34</td>
</tr>
<tr>
<td>2</td>
<td>Results</td>
<td>40</td>
</tr>
<tr>
<td>Figure</td>
<td>Page</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>38</td>
<td></td>
</tr>
</tbody>
</table>

Unit Definitions
Chapter I

Overview

The present thesis investigates the discourse structure of the scientific experimental-research paper, in particular its introductory section. It also presents a methodology for validating such analyses of discourse structure.

In Chapter I, the genre of the scientific paper is discussed in general terms* and the experimental-research paper is defined. Its importance for students of English for Science and Technology (EST) is indicated. The need for an analysis of its structure is justified through reference to schema theory, in terms of the effect that knowledge of the structure of a text has on its comprehension and production. A number of analyses of the structure of the scientific paper are reviewed.

In Chapter II, the main deficiencies these analyses have in common are reviewed. A comparison is made between them and the work of Swales (1981), which is identified as the most important recent research in this area. A more detailed critique is made of Swales' work. Its main problem is identified as a lack of empirical validation.

Chapter III outlines the way in which standardized procedures for determining the consistency with which raters apply analyses of behavior can be applied to the validation of an analysis of discourse. If a group of trained raters can apply a particular model of analysis to the subject of
investigation and obtain a satisfactory level of agreement, it may be taken that the system has psychological reality. The remainder of the chapter describes the method used in the present study. First, Swales' original system, and then a revised version of it, were applied to a carefully selected corpus of texts. The results of this investigation are presented.

Chapter IV reviews the results and draws conclusions. Aspects of the methodology and implications for EST materials writing are discussed, and ideas for further research and other work already in progress in this area are described.

Introduction

The growth of science and the use of English as its principal language of information dissemination have vastly increased the number of non-native science students in institutions of higher learning around the world. In response to their needs (Julian, Lowenstein and Slattery 1978, Robertson 1983), we have seen the development of programs and research in English for Science and Technology (EST).

EST is one type of English for Specific Purposes (ESP), which was a response to the demand for service courses in ESL which were directed to the specific needs of students.
Such service courses (often short-term and subject-specific) were generally produced for the private sector, and reflected a concern for accountability in their focus on precisely the skills and content which participants would have to be capable of using on completion of the course. EST courses were initially developed as service courses for university level students of science and technology. These students' need for English reflected the international nature of the language as the primary mode of communication in science. It also reflected the fact that even in many parts of the non-English speaking world, higher education science courses are often conducted in English (Allen and Widdowson 1978).

The rapid growth of EST in the 1970s resulted in a sudden demand for materials, and a variety of texts were published in a short space of time. These texts were designed to aid the development of skills necessary to handle types of scientific discourse. Their quality was variable. A characteristic deficiency was the absence of appropriate research or analysis of the nature of the tasks that students of science and technology had to carry out in English (Swales 1980, Mackay and Mountford 1978:vii), or of the materials, register and types of discourse they needed training in. A noted authority in the field has observed

ESP textbooks have been in many respects an educational failure. (Swales 1980:11)
There was also an absence of texts designed to handle the needs of more advanced students (see, for example, Bates 1978:94). One genre of EST which constitutes an essential part of the reading of advanced science students is the scientific paper. As Hill, Soppelsa and West (1982) observe, the ability to read and write experimental-research papers is important for success in the sciences. However, few ESL teachers are comfortable teaching students how to read and produce experimental-research papers (Schachter 1981). There are few ESL texts which deal with how to write such papers, despite science students' established preference for technical, subject-related materials to be used in ESL classes (Hill et al. 1982).

The quality and quantity of ESP/EST research has improved in recent years, as witnessed by the appearance of a new professional journal devoted to the field (The ESP Journal, begun in 1980), and by the formation of an ESP Special Interest Group in the TESOL organization. However, such work often displays an uncoordinated and unsystematic approach to the issues investigated, and is often characterized by an inadequate data base or by a lack of rigor in the way data is gathered and analyzed. This undoubtedly reflects the fact that many practitioners work under less than favorable conditions, such as on attachment to third world universities, some of which have limited
facilities for research (see e.g., Wingard 1981). For example, a recent article was published in TESOL Quarterly in 1983 on teaching students to read and write scientific articles (Hill et al.). It was concerned with their structure as well as with pedagogical techniques. Nevertheless, it made no reference to the most important other piece of work in this immediate area (Swales 1981). The latter work, although a thorough and insightful analysis in many respects, was an interim report which presented one researcher's analysis of article structure. It made no attempt to submit the results to empirical test.

The present study is based on that of Swales (1981) and addresses the same pedagogical needs attended to by Hill et al. (1982). It attempts to provide an empirical test of the Swales analysis, to develop a revised analysis and to present a methodology for future work. In doing so, it is motivated by an awareness of the need for replication and for researchers to build upon each others' work.

The scientific paper

The scientific paper can be broadly defined as a type of scientific writing, based on a single investigation, whose purpose is to contribute to the progress of science or technology (Peterson 1961:6). The present work is concerned only with papers published in professional, refereed, scientific journals. In addition, abstracts and "short
papers", or communications are excluded. A short paper is a report of work of immediate interest, intended principally to effect the quickest possible communication with others in the field, but does not constitute a full and official communication of the work, despite publication in an official journal. Such papers are becoming increasingly common.

Morris (1966:204) identifies two basic types of scientific paper published in journals: the theoretical type, and the experimental-research paper. Peterson (1961:133, 169-170) refers to the review article as an additional important and separate type. The experimental-research paper, to quote Hill & & (1982)

reports experimental or *ex post facto* research designed to test a hypothesis or theory. (334)

According to the conventional wisdom of rhetoricians and technical writers, such papers can be assumed to incorporate different rhetorical structures. It is these structures which make them distinctive and which ESL/EST students have to be able to recognize if they are reading such papers, (Carrell 1984) and produce, if they are writing them.

Scientific papers are subject to some constraints concerning form and style. The requirements of the journal editor and referees who represent the scientific community
apply to all papers. The experimental-research paper must also reflect hypothetico-deductive scientific method. This leads to a degree of standardization which suggests that such papers may share a common basic structure or schema, or employ common units of discourse.

The need for EST materials concerned with discourse structure

The pedagogical desirability of a knowledge of discourse structure has been stated by Pitkin (1969) in relation to the teaching of English composition to native speakers:

We need presently and will continue to need more efficient models for teaching our students to read connected discourse with understanding. (138)

Selinker, Todd Trimble and Trimble (1976) suggest that a similar need is felt in EST. Advanced ESL/EST students, they wrote

often seem unable to comprehend the total meaning of EST discourse even when they understand all of the words in each sentence and all of the sentences that make up the discourse. (282)

For Selinker et al., it is the students' lack of knowledge of the rhetorical structure of the discourse which hinders them.

Both Hill et al., (1982), and Swales (1981) are concerned with similar problems in their work. These and
other researchers assume a need for direct teaching of the rhetorical structure and organization of scientific papers. It might be inferred from these concerns that knowledge of a text's structure aids comprehension and production of such texts. A relationship exists between an individual's psychological conception of a form and his ability to comprehend and utilize it. When such a form is discoursal, this conception is referred to as a 'schema'.

Schema theory

Schema theory is a developing area of investigation which lies at the intersection of cognitive psychology and linguistics. It is principally concerned with the way in which various types of background knowledge affect understanding and recall. Carrell (1983), in a recent review, distinguishes between formal schemata, which deal with the rhetorical structure of discourse, and content schemata, which deal with general world knowledge.

Schemata have been shown to guide the comprehension not only of events and actions...but also to guide the interpretation of the linguistic representation of these events, scenes, activities -- i.e., oral and written texts.

Early work of relevance is Mandler and Johnson's (1977) work on story grammars. This refers to the underlying cognitive structure of a narrative, in terms of setting, event structure and episode. The proficient reader, Mandler
and Johnson argue, has internalized such a story grammar, and uses it to process stories.

Further support for such a notion comes from the work of Kintsch and van Dijk (1975, 1978), which suggests that knowledge of the schema of a piece of discourse is critical for its effective recall and understanding. Kintsch and van Dijk (1978) further develop a model of text processing, and partially summarize their earlier work:

There are a number of highly conventionalized text types. If a reader processes such texts in accordance with their conventional nature, specified well-defined schemata are obtained. These are shared by the members of a given cultural group and, hence, are highly suitable for research purposes. Familiar examples of such texts are stories (Kintsch and van Dijk 1975) and psychology research reports (Kintsch 1974). These schemata specify both the schematic categories of the texts (e.g., a research report is supposed to contain introduction, method, results and discussion sections) as well as what information in each section is relevant to the macrostructure (e.g., the introduction of a research report must specify the purpose of the study).

For Kintsch and van Dijk, it is the schema which determines which of the many propositions in a text are relevant or irrelevant to the reader, and thus directly affects how and whether they are processed or recalled. If the reader has a schema...that is not well defined, the outcome of the processing of the text will be "haphazard".

with obvious problems for comprehension, whereas if s/he is
familiar with the conventional nature of the text, well-defined schemata will be produced which will aid comprehension and recall. The reference to the culture-specific nature of schema is borne out, at least for the narrative genre, by the work of Colby 1970). The ESL/EST student who does not share such schemata may experience difficulties in reading comprehension.

Early EST rhetorical approaches to discourse structure

The work of Kintsch and van Dijk demonstrates how knowledge of schema or text structure is reflected in comprehension. It provides a justification for further analyses of text structure, such as the present one. It also provides needed support for earlier investigations into text structure which used different and possibly weaker methodologies. In the field of EST rhetoric, reference is made by Lackstrom, Selinker and Trimble (1973) to the rhetorical choices the writer of scientific English makes. A hierarchical relationship is posited between the choices to be made. A choice at the highest level (Level A, for example, 'Detailing an Experiment'), "to a large extent determines rhetorical choice at a lower level".

If the author's purpose is to present the results of an experiment, that is a choice at Level A, then he has no choice but to inform his reader of the apparatus that he has used in that experiment, a predetermined choice at Level B. (133)
In a subsequent revision (Selinker et al. 1976), Level B items are referred to (283) as 'general rhetorical functions', as opposed to, for example, 'specific rhetorical functions' such as definition or classification.

The principle focus of Lackstrom, Selinker and Trimble's investigations was the relationship between grammar and rhetorical choice, and later the role of implicit presuppositional information in scientific text. However, their investigations led other researchers to work which was more closely related to schemata.

Hepworth (1979) refers to Lackstrom, Selinker and Trimble as having listed the rhetorical functions for an experimental report as the Statement of the Problem, the Hypothesis, the Deduction of the Consequences, Description of Procedures, and so on. (149)

These are again general rhetorical functions although Hepworth does not make this distinction. He states that they are "very seldom explicitly identified in the text of any report" (149). Consequently

while the rhetorically-competent discourse analyst and the rhetorically competent scientist may be able to recognise these features...it is not at all clear that the same will be true for many ESL/EST students. It is, in fact, quite likely that the ESL/EST student will need considerable practice in learning to use these rhetorical features in writing. (153)

This pedagogical justification for a deeper analysis is
supplemented for Hepworth by the fact that merely identifying rhetorical functions as constituting the structure of the introduction to an experimental report is a post hoc procedure (149). Rather, he proposed to identify them by way of the "explicit discourse functions [they] employed". These are causality, contrast, time or space order and so on (149), such as are "common to any discourse" (158). These are what Selinker et al. (1976:283) call "rhetorical techniques".

Hepworth makes use of a discourse analysis system called Discourse Blocs (Pitkin 1969, 1973, 1977a,b). This system divides a discourse into blocks according to the explicit discourse function each one manifests. Types of functions are again the rhetorical techniques, stated in a binary form such as contrast/contrast, cause/effect, focus/assertion, or includer/included.

Hepworth's hypothesis (1979:153) was that there would be a correlation between such units and the rhetorical functions suggested by Lackstrom, Selinker and Trimble for the introduction to an experimental report: Statement of Problem (PROB), Hypothesized Solution (HYP), and Declaration of Consequences (D of C).

The description of method employed is brief:

In order to test our hypothesis, we carefully analysed the introductions to several experimental reports. (149)
After this test, for which no details of procedure or corpus are given, Hepworth found that a typical report introduction was a cause/effect bloc, (in Pitkin's terms), with the cause part containing a concession/assertion bloc of the form 'true X/yet Y'. Finding that this did not fit with the analysis of Lackstrom et al. Hepworth proposed to add a new unit to the Pitkin system: problem/solution:

We intuitively feel that cause/effect fits nicely with what we might call problem/solution.

(155)

The main result of his discussion is to suggest that discourse functions and the "rhetorical functions of EST" be equated thus -

<table>
<thead>
<tr>
<th>discourse functions</th>
<th>PROBLEM</th>
<th>SOLUTION</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>previous data</td>
<td>hypothesis</td>
</tr>
<tr>
<td>rhetoric functions</td>
<td>research lacking</td>
<td>deduction of consequences</td>
</tr>
<tr>
<td></td>
<td>current conflicting</td>
<td></td>
</tr>
<tr>
<td></td>
<td>theory data</td>
<td></td>
</tr>
</tbody>
</table>

(after Hepworth 1979:157, fig. 9)

Hepworth seems to be saying that the introduction to an experimental report exemplifies a special type of discourse structure - the problem/solution type. It should be noted that he has moved away from Pitkin's model somewhat in terms of the direction of analysis. It is not so much that the rhetorical units employ the discourse function 'problem/solution', as the other way around: the introduction to a report may be seen as one particular
example of a more general structure.

For the more specific purpose of the present investigation, it emerges that one possible set of rhetorical functions to be utilized in a report introduction is as outlined above, and, insofar as schema may be considered an organizing device for rhetorical forms, this set may be indicative of schema structure as well.

Prescriptive approaches to EST discourse structure

In his major study of discourse structures, 'The Thread of Discourse' (1975), Grimes directs our attention to writing teachers:

I find insights that contribute to discourse linguistics coming from people who are primarily skilful practitioners of the art of teaching writing, yet whose knowledge of linguistics itself precludes their making the kind of systematizations a linguist would make.

(14)

a very good start on the study of discourse patterns...could probably be made simply by bringing together systematically all the things that rhetoricians have said that speakers of...[a] language should or should not do.

(12)

It is from such "skilful practitioners of the art of teaching writing" that two of the most recent descriptions of the schema and archetypal content of the scientific paper come (Hill, Soppelsa and West 1982, Day 1979).

Hill et al. approach the problem from the point of view of developing materials and methods for teaching ESL
students to read and write scientific papers of the experimental research type. A description of their structure is a preliminary to this. They apply the Aristotelian 'beginning/middle/end' framework and show how such papers typically proceed from the general, in the Introduction, to the particular, in the middle section which they call 'Procedure', and return to the general in the final section, 'Discussion'.

'Procedure' is frequently divided into 'Methods' and 'Results' sections, they say. 'Methods' is a 'chronologically-ordered' descriptive section. 'Results' describes the manipulation of the data obtained from the process described in the methods section and gives the empirical results from such manipulation

They characterize the discussion section as essentially a mirror image of the introduction. It moves from the solution of the problem which motivated the study to the implications of the solution for the larger field.

It also "presents limitations of the study" which provide "areas for future research".

Day's work (1979) is in the tradition of technical writing, and is addressed to an audience of research scientists regardless of first language. It covers the area in great detail, providing prescription on matters of style and form, as well as steps to be taken to ensure publication. The extensive series of quotes which follows
provides an expansion of the comments of Hill et al.

The introduction...
(i) It should present first, the nature and scope of the problem investigated
(ii)...the pertinent literature should be reviewed
(iii) The method of investigation should be stated
(iv) The principal results of the investigation should be stated

For methods, the normal order of presentation is chronological [but]...related methods should be described together and straight chronological order cannot always be followed. The materials and methods section is the first section in which sub-headings should be used. When possible, subheading should match those to be used in Results.

There are usually two ingredients of the Results section. First, you should give some sort of overall description of the experiments, providing the 'big picture', without, however, repeating the experimental details previously mentioned...Second, you should present the data.

The Discussion is harder to define than the other sections.

1 Try to present the principles, relationships and generalizations shown by the results
2 point out any exceptions or any lack of correlation, and define unsettled points
3 Show how your results and interpretations agree (or contrast) with previously published work
4 Discuss the theoretical implications of your work, as well as any practical applications
$State your conclusions$
6 Summarize your evidence for each conclusion

The Discussion should end with a short summary or conclusion regarding the significance of the work

This is not so much Day's personal opinion, as a
codification of the collective wisdom. For example, O'Connor (1978:132) refers to the "conventional IMRAD structure (Introduction, Methods, Results and Discussion)" [my emphasis]. Trelease and Yule (1925) make much the same statements that Day puts forth fifty years later.

Literary and stylistic approaches to EST discourse structure

A different tradition is built upon by Hutchins (1977)--that of the literary critic and the analyst of narrative and folk tale. Hutchins says that in the past, the study of the "macro-structure" of texts (which he glosses as the "overall organisation" of texts (18)), "has been concentrated almost wholly on narrative texts". He addresses himself to the expository text in the form of what he refers to as the "scientific paper", which he defines as an article of a learned journal which argues for the revision of some accepted opinion in some area of academic study (18-19)

Hutchins first reviews the work of Propp (1968), Longacre (1974), Greimas (1966) and Brémond (1970) on the macro-structures of narrative texts. Propp's work on Russian folk-tales deals with the arrangement of episodes which Propp calls 'functions' (sections of the tale which can be summarized in a phrase, such as 'the treachery of the villain'). Longacre's work posits more abstract units, such
as 'climax', or 'denouement'. Greimas argued that Propp's functions could be reduced to a sequence: 'orientation, complication, evaluation, resolution and coda', which matches the level of abstractness of Longacre and also of Brbmond, who postulated a cycle: 'equilibrium, degradation, disequilibrium, amelioration, equilibrium'. The cycle consists of two parts - equilibrium to disequilibrium, and from there back to equilibrium again. Longacre's goes from *setting* to *climax* and from there to *closure*. Hutchins argues that such a two-stage pattern is evident in both Greimas' and Longacre's analyses, and then seeks similar forms in the macro-structure of the scientific paper. He takes as a starting point Gopnik's Linguistic Structures in Scientific Texts (1972). According to Hutchins, Gopnik identified 'three basic types of scientific paper: the *controlled experiment* type, the *hypothesis testing* type, and the *techniques description* type.' He shows that all three of these types can be recast in terms either of semantic progression or in terms of Brbmond's cycles. On this basis he is prepared to argue for a generalized structure, whose applicability is such that he claims:

the following structure seems to underlie many scientific papers:

| *current* hypothesis/paradigm |
| demonstration of inadequacies |
| statement of 'problem' |

18
the 'solution'

statement of 'new' hypothesis or of alternative hypothesis
testing of hypothesis or hypotheses
'proof' of hypothesis or of one of the alternative hypotheses

implications of 'solution'

(30-31)

From this position Hutchins then goes on to discuss the place of the scientific paper in both science as a whole and in the historical tradition of rhetoric. His basic position remains that the generalization made with regard to narrative macro-structure can be brought over to expository discourse, specifically the scientific paper, and he uses arguments based on Gopnik's analysis to justify this. The structure outlined above may be intuitively appealing, but the strength of Hutchin's argument is limited by the fact that Gopnik's data base was not actually scientific papers, but a corpus of 28 short summaries of the intended presentations to be given at a scientific meeting. The texts analysed were in fact conference abstracts.

Hutchins' fairly 'delicate' analysis, as Hoey (1979) describes it, has since been supplemented by Swales' (1981) work on article introductions. This is more in the tradition of EST rhetoric referred to earlier, but its conclusions bear comparison to those of Hutchins. With a more satisfactory data base (16 articles from each of physics, biology/medicine, and the social sciences), Swales
identified four 'moves' in the introductory sections of the papers, which in turn could be sub-categorized.

Move 1 Establishing the field
   [by] showing centrality
   stating current knowledge
   ascribing key characteristics
Move 2 Summarizing previous research
Move 3 Preparing for present research
Move 4 Introducing present research
   [by] giving the purpose
   describing present research

Swales' moves are obviously more closely tied to the nature of his corpus than Hutchins' units, which echo more abstractly the scientific process, and which are intended to reflect underlying trends for expository prose in general. There are some interesting similarities, however. It might be suggested that "establishing the field" and "summarizing previous research" (Moves 1 & 2) were close to a statement of the 'current' hypothesis/paradigm; that Move 3, "preparing for present research" is often effected by a demonstration of inadequacies; and that Move 4, "introducing present research" might contain or constitute a statement of 'problem'.

Markers and signalling

Reference has so far been confined to existing work concerning the possible units of the experimental-research paper. It is also necessary to consider briefly how those units may be signalled. One of the limitations of existing
work is that mentioned by Hoey (1979), referring to the work of Hutchins:

Hutchins' discussion of the structure [of the paper] is the most considered but it still leaves some important questions unanswered. Perhaps the most crucial of these is...how are the structure and its component parts identified by the reader/listener? 

One obvious possibility is by way of lexical items. Work done on discourse structure signals in English compares unfavorably with work done on some other languages, and Hoey says "What work there has been in English on this aspect of discourse has been sketchy and off the cuff". Early work concentrated on the way conjuncts indicate the relationship between clauses, but the work of Farnes (quoted in Hoey), and that of Winter (1977), establishes that the signalling of such relationships is not restricted to conjuncts.

The function of a sentence or group of sentences in a discourse may be signalled by a conjunct, a clause or even a sentence...the function of a sentence need not be signalled from within the sentence...it is quite normal for the function of a sentence or group of sentences to be signalled in advance. (Hoey, 1979:29)

Murphy and Candlin (1979) analysed a stretch of scientific discourse (not an experimental-research paper). They used the model for analysis of oral classroom discourse developed by Sinclair and Coulthard (1975). In attempting to apply this model, they focused attention on the way that the passage provides signals equivalent to those used in oral discourse to indicate the boundaries of exchanges. Where
words like 'Right.' and 'Now..' are used in speech, they point out that in written text, typographical conventions may be used instead (in addition to lexical items, clauses etc.). Given that the Sinclair-Coulthard system was not designed for written scientific text at all, Murphy and Candlin found deficiencies, but noted that what the system did do successfully was gloss metalinguistic devices of the utterer, the devices used to guide the reader/hearer through the text. However, the materials designer needs to interpret and characterize the network of meaning relationships that constitute the message the hearer/reader is expected and wishes to interpret. (48)

This "network of meaning-relationships" is elsewhere referred to as "plot".

With the exception of Swales, most analysts of EST discourse give very little attention to this problem. Swales' position will be discussed in the following chapter.
Inadequacy of existing analyses

The main problem with existing analyses that have been referred to as 'rhetorical' arises out of the characteristics of the rhetoric tradition. This tradition prescribes, but does not attempt to verify that what it prescribes is what happens. Alternatively, the analysis is deliberately abstract, intended to provide a way of perceiving a system at some underlying level.

Day's work, for example, although thorough and useful, cannot be taken as establishing the accuracy of the system advocated by technical writers. It is prescriptive, and therefore not necessarily descriptively adequate. The fact that writers advocate the use of certain rhetorical structures in technical writing does not necessarily mean that such structures are in fact employed in technical writing. A parallel may be drawn between prescriptive approaches to the scientific article and to the paragraph.

For more than a hundred years, and even now, a particular type of analysis concerning the paragraph has prevailed, with prescriptions concerning the positioning of topic sentences and subsequent types of development. However, the work of Rodgers (1966) establishes the unfounded nature of the analysis through an investigation of its historical development. That of Braddock (1974) and
Stern (1976) calls into question statements concerning the existence and positioning of the topic sentence, and suggests that even experts cannot identify paragraphs from other than their typographical markers.

Other problems attend the work of Hepworth (1979). He criticises Lackstrom, Selinker and Trimble for a post hoc approach to their units of analysis, and bases his on 'explicit discourse functions' such as 'causality, contrast, time and space order, and so on'. Insofar as Hepworth was writing in the rhetorical tradition, it is not surprising that he did not attempt to define these units explicitly, nor give details of his analytical procedures or corpus.

In arriving at 'Problem/Solution' as the discourse function underlying article introductions, Hepworth came to the same conclusion as Hutchins, whose argument is based on showing that other linguists and literary analysts have found closely paralleling structures in a variety of types of texts, and that a similar one appears in the scientific article.

It may be argued that we cannot assign anything more to these analyses than a verdict of 'not proven', since their units (as Boey 1979 points out) are not clearly defined, their analytical techniques, when used, are not indicated, nor are their corpuses referenced (in the case of Hepworth) or appropriate (in the case of Hutchins). Above all, there is no attempt made to subject them to empirical test.
The work of Swales, by contrast, is far more rigorous. (See above, for details of the four-move system arrived at by Swales.) Here, for the first time, we are presented with a referenced and appropriate corpus, units identified, a carefully detailed explanation of the analysis of the corpus and of remaining problems. There is only one major problem - the research lacks empirical validation. Swales himself indicates his awareness of this. He recognizes the danger that

the discourse analyst labels something as \( x \)
and then begins to see \( x \) occurring all over the place

(Swales 1981:13)

"One way out" that he suggests, which various EST analysts have taken (Tarone, Dwyer, Gillette and Icke 1981, Pettinari 1981 and others) is to consult a specialist as to what is 'really' going on. Swales did not do this. (This may have been for logistical reasons, since he refers to limitations of time in his preface.) He recognizes that

I am open to the charge that

my unsubstantiated and ill-defined terminological labels...are little more than a reflection of my own perceptual predispositions.

(14)

Part of the object of the present work is to indicate how an analysis of this sort may avoid laying itself open to such a charge. See below, chapter 111.

Other difficulties which exist apply to the sampling.
There are two valid approaches: either one should follow Roe (1977), and select according to the texts used by a particular group of EST students or courses, or a random sample could be taken and the procedure indicated. Swales' work is limited by the fact that it grew out of work intended to investigate the structure of the section of article introductions which reports previous research. That he did not confine his work to this area is to our benefit. However, the original orientation of the work resulted in problems concerning the corpus selected. Full details of the random selection method used are not given, but it is clear that Swales deliberately selected only those articles which contain reports of previous research within them. Having widened the investigation to the general structure of article introductions, his conclusions are limited in their generalizability by this somewhat inappropriate corpus. He does mention this in terms of the four-part analysis he proposes, of which part 2 would occur a priori, but not in terms of the effect this would appear to have on the generalizability of the conclusions.

A further problem lies in the way Swales states his interim conclusion about the structure of the article introduction:

I would therefore like to claim, on the basis of the 48 introductions in the corpus, that in something over 70% of the cases the author or authors make four sequenced moves.
The 70% is footnoted as \[43/48 \times 48/48 \times 40/48 \times 46/48 = 71.5\%\]. If in greater than 70% of the cases, a sequence 1234 was observed, this statement would be justified. But the calculation appears to refer to figures for suppliance for each move. It is not clear whether, of the 43 introductions which had a move 1 (and which also had move 2s, since that was a basis for selection), 40 had move 3s as well, or whether alternatively, some of the 5 introductions which did not have move 1s had move 3s, and so on.

Then there is the question of product versus process. Swales dismisses his earlier attempt to apply the Sinclair-Coulthard model to these texts, because

Such an approach led to an analysis of product, whereas a process analysis that attempted to identify the obligatory and optional moves a conventional journal author might make in the construction of a text could well be more directly beneficial for the kinds of application that I had in mind.

It seems here that Swales is suggesting that the analysis referred to indicates the author's actions:

I would like to claim that the author or authors ... make four sequenced moves.

It would thus be an analysis of the process of writing, rather than of the product. But later he says

I have not so far tried to assign a place for article introductions within some texttypology, nor have I made any serious effort to explain what I think is 'really happening' when a writer drafts or redrafts an introduction.
It might be agreed that a process-based analysis is more useful than a product-based one, but both Swales' and the present study are analyses of product.

Remaining problems apply to the definitions and signals of the moves, and they will be considered as part of the description of the revision of the Swales model, in Chapter 111.
Chapter III

Validating a discourse analysis

In developing an analysis of behavior which proposes that a particular type of behavior is made up of sequences of units, it is standard practice for the validity of such an analysis to be established by showing that the units can be defined in such a way that a group of trained raters can record the incidence of units of behavior at a sufficiently high level of agreement. If a system represents a realistic depiction of the activity, it will be possible to obtain such agreement. Having a group of trained observers carry out an analysis of such activity prevents the analysis from being purely subjective. The Swales model is open to the criticism that it is a purely subjective analysis, as Swales himself recognises. But if it can be shown that it can be applied by a group of trained raters to an appropriate corpus at an adequate level of agreement, it could justifiably be regarded as reflecting reality.

Goals of the present study

In the above survey, we have seen that a knowledge of the discourse structure of text is critical for its understanding and production. In addition, in the area of EST, the analyses reviewed have been shown to be lacking on methodological grounds. In particular, they are open to the charge of subjectivity. The goals of the present study
were to validate a discourse analysis of the structure of one scientific text type.

Method

In outline, the procedure was as follows:

1) a corpus was selected
2) raters were selected
3) the overall design was explained to the raters
4) definitions of the units of analysis were presented and discussed
5) unit boundary markers were presented and discussed
6) worked examples were presented
7) raters practiced analysing simple texts, whose structure had already been established
8) more complex texts were analysed, interrater reliability scores calculated, and disagreements discussed by the raters, both between themselves and with the trainer
9) step 8 was repeated until a satisfactory level of interrater agreement was attained.
10) the corpus was rated
11) analysis

The corpus

Initially, a corpus of 96 scientific articles was selected. The same basic categories as used by Swales were
adopted, for reasons of comparability: "the hard sciences", "the Biology/Medical field" and "the Social Sciences". Within each of these three sections, four journals were chosen, according to Garfield's criterion of popularity. Garfield (1981, 1982) ranks journals according to number of times cited in a particular period. Thus it is possible to determine the importance of a particular journal in terms of its likelihood of being encountered by anyone reading or doing research in a given area. By working down the unified rank list for times cited in the post-1980 period, it was possible to build up a group of the four most "popular" journals for each of the three areas.

Having settled upon the titles, and having chosen 1980 and after as the period that selection would be restricted to, a table of random numbers was used to provide a point of entry to a particular month in the period from 1/80. From the issue thus selected (if the journal appeared monthly) and the first of the month otherwise, four articles were then selected. Articles were examined, beginning with the first in the relevant issue, and in sequence thereafter, and the first four suitable were entered into the corpus. Articles were rejected if they appeared to be written exclusively by non-native authors, if they appeared to be review or theoretical articles, and if selection would mean having more than one article by the same author in the corpus. If four suitable articles were not to be found in
the same issue of the journal, the subsequent issue was moved to, and so on.

Training of raters; testing the Swales model

It was decided that raters should be individuals with some linguistic sophistication. Specialists (i.e. scientists or technical writers) were not selected, principally due to the logistical problems envisaged.

A pair of MA(ESL) students was trained (myself and one other) using the extensive sequence of fully analysed article introductions and representative examples given by Swales, as well as a number of article introductions not included in the corpus, but taken from the same journals.

Regretably, Swales does not define the moves clearly, but provides a "term" - thus in reference to Move 1:

The term I have eventually settled on to characterize this opening move is Establishing the Field. (1981:23)

The term is then explained in considerable detail, and examples given.

One quarter of the corpus was selected (by stratified random sampling from the original corpus), and rated. Results are shown in Table 1. Despite ten hours of joint effort, plus individual study of the materials, satisfactory interrater agreement could not be arrived at, unless articles which appeared not to conform to the Swales model
were not included. Excluding them, interrater reliability was high: \( \text{kappa} = 0.96 \) (Cohen 1960).
Table 1 -- Article introductions from the corpus conforming to the original journal text no. tone

<table>
<thead>
<tr>
<th>journal</th>
<th>text no.</th>
<th>tone</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>hard sciences</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>J. Astrophysics</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>J. Chemical Physics</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Physics Review B</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td></td>
</tr>
<tr>
<td><strong>biology/medicine</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New Eng. J. of Medicine</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Lancet</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>J. Cell Biology</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

The remainder of the first quarter articles, including all 8 from the journals), did not conform to the original. There were two texts from each journal. Interrater reliability for the above 8
Difficulty with the system did not merely reflect the raters' lack of training, intelligence or understanding of English—it rapidly became clear that some introductions deviated so far from a strict four-move schema as to call into question such a system's descriptive adequacy. As mentioned, the terms used in the Swales model to characterize moves are brief. Their explanations are very detailed, but despite this some appeared to be too vague to apply without unacceptable levels of disagreement.

More significant was the question of boundary markers. Brown and Yule (1983:69) refer to formulaic expressions which are used as markers to indicate the boundaries of blocks of discourse. When there are no boundary markers, reliance must be placed on the notion of topic, which, they say, is very difficult to define. Swales, referring to Move 2, says "in contrast to the other three moves, there tends to be no signal of onset." This was generally found to be the case in the present work. Swales also gives ten examples of the openings of Move 2, which, unfortunately, as he says, "exhibit an uncanny resemblance to Move 1 signals" (33-34). An explanation of the difference between Move 1 and Move 2 signals is given (34), but the problem caused by this similarity in signals was too severe to be overcome. Authors exhibited a disturbing tendency to summarize previous work (the definition of Move 2) at any place in the Introduction. It became clear that revision of
both Move 1 and Move 2 was necessary to clarify the point at which the former changed into the latter.

The other major problem concerned the applicability of the conception that Introductions have four moves, one of each type. Swales' corpus does not appear to contain any introductions with more than four moves, except in one exceptional case. There, "the first of two Move 3 elements is embedded inside Move 2." If there is not a hierarchical relationship implied by the word `embedding' (impermissible given the definition of the moves), then the alternative is a Move 2, followed by a Move 3, followed by a Move 2. It seemed that, given our corpus and the definitions derived from Swales' terms and explanations, there were very many examples of such 'embedding', involving all four moves.

It appeared to be the case that though the units of analysis were well motivated (despite a possible Move 5 -- see Chapter IV), their definitions were insufficiently clear in some cases. The limitations on the sequences in which they might appear and their optionality implied by the Swales model did not permit its easy application to the corpus at hand.

Consequently, a slightly revised model was developed. The original system embodied a lengthy description of some ten pages of text for each unit. For the revised version, short definitions were prepared (see Figure 1). Units
of discourse were referred to as 'types' rather than 'moves'. The emphasis on conciseness, and on changes in level of generality was introduced with regard to Type Is. The function of reporting past research for the purpose of summarizing it was introduced for Type 2s. The definitions of the other two units reflect more closely the original Swales units. The requirements that there be only four units, and that they be in the order 1-2-3-4, were both eliminated. All introductions were to be coded - there was no category for 'deviants'. Finally, three new raters, again all MA(ESL) students were obtained. It was felt that the author (one of the first pair of raters) no longer represented a 'normal' subject with regard to article introductions, due to overexposure.
Type One (1)

This unit establishes the topic. It often appears at the start of article introductions, and manifests itself in a number of ways:

- It asserts the centrality of the topic of the paper to current research; it implies that the research is mainstream; it indicates common knowledge regarding the topic, or states the current state of knowledge; it may describe or make an assertion about a key aspect of the topic; it establishes that the topic is prominent; it indicates the interest or importance of the topic.

It is concise. It makes statements of a high level of generality.

Type Two (2)

Type Two refers to past research for the purpose of summarizing either all relevant research, or merely that section of past research referred to. In simple introductions, it often follows a 1. When it does so, it is marked by a drop in the level of generality of reference.

Type Three (3)

Type Three justifies the research. It has a number of manifestations. It indicates a gap or a defect in the research, questions its validity or otherwise evaluates it; it airs a problem, or raises a question or hypothesis; it indicates that a finding should be extended, explored or taken further.

Type Four (4)

Type Fours are essentially metacommments on the following research. They give the purpose of research, introduce it briefly, or follow-up a statement that a finding should be extended.

Figure 1 -- Unit Definitions
Following training, raters attained a high interrater agreement figure of 0.9 (Fliess 1971) on a test run of a small section of the first quarter of the corpus. They then rated completely the second quarter of the corpus. Financial and temporal limitations precluded rating the whole of the corpus: a further one quarter of the original corpus was coded (different to that used to test the Swales model). The results are summarized in Table 2.

Results

As may be seen from Table 2, in three-quarters of the sample, raters were able to reach agreement as measured by kappa (Fliess 1971) at above 0.6 (see also Gelfand and Hartman 1975, below). There were two texts for each journal. Sentences which were not coded unanimously were omitted in determining the sequence of units in each text. No structure is indicated for texts where kappa was below 0.6. The most common structures were 24 and 1234 (five occurrences each). The structure 1234 was not observed in any social science texts. The structure 24 was observed once in social science texts. There were on average 1.8 times more units in social science texts than in other texts.
## Table 2

### Results

<table>
<thead>
<tr>
<th>journal</th>
<th>text no.</th>
<th>structure (text units)</th>
<th>interrater agreement (kappa)</th>
<th>length (sentences)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>hard sciences</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Astrophys.J.</td>
<td>1</td>
<td>1 2 4 2 4</td>
<td>0.61</td>
<td>17</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>0.49</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Phys.Rev.B.</td>
<td>1</td>
<td>2 4</td>
<td>0.62</td>
<td>9</td>
</tr>
<tr>
<td>2</td>
<td>1 2 3 4</td>
<td>1.00</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>J.Chem.Phys.</td>
<td>1</td>
<td>1 2 3 4</td>
<td>0.66</td>
<td>23</td>
</tr>
<tr>
<td>2</td>
<td>1 2 3 4</td>
<td>0.93</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>J.Am.Chem.Soc.</td>
<td>1</td>
<td>1 2 4</td>
<td>0.59</td>
<td>14</td>
</tr>
<tr>
<td>2</td>
<td>1 2 4</td>
<td>0.78</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td><strong>biology/medical sciences</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New Eng.J.Med.</td>
<td>1</td>
<td>1 2 3 4</td>
<td>0.91</td>
<td>12</td>
</tr>
<tr>
<td>Z</td>
<td></td>
<td>0.50</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Lancet</td>
<td>1</td>
<td>2 4</td>
<td>0.77</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>0.55</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>J.Cell Biology</td>
<td>1</td>
<td>1 2 3 4</td>
<td>0.33</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>1 2 3 4</td>
<td>0.74</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>J.Physiology</td>
<td>1</td>
<td>2 3 4</td>
<td>0.70</td>
<td>11</td>
</tr>
<tr>
<td>2</td>
<td>2 4</td>
<td>0.70</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td><strong>social sciences</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Child Dev.</td>
<td>1</td>
<td>1 2 3 2 3 4</td>
<td>0.72</td>
<td>36</td>
</tr>
<tr>
<td>2</td>
<td>1 2 3 2 3 2 3</td>
<td>0.61</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td>Arch.Gen.Psych.</td>
<td>1</td>
<td>1 2 3 2 3 4</td>
<td>0.76</td>
<td>10</td>
</tr>
<tr>
<td>Z</td>
<td>1 2 4 2 4</td>
<td>0.65</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>Am.Soc.Rev.</td>
<td>1</td>
<td>2 4</td>
<td>1.00</td>
<td>17</td>
</tr>
<tr>
<td>Z</td>
<td>1 2 3 2 3 2 3 4</td>
<td>0.8</td>
<td>39</td>
<td></td>
</tr>
<tr>
<td>J.Pers.Soc.Psych.</td>
<td>1</td>
<td>2 4 2 3 2 3 4 2 4</td>
<td>0.24</td>
<td>74</td>
</tr>
<tr>
<td>Z</td>
<td></td>
<td>0.68</td>
<td>S7</td>
<td></td>
</tr>
</tbody>
</table>

40
Chapter IV

Discussion

It may be tentatively concluded that results are consistent with the idea that four basic units of discourse occur in article introductions. In some shorter article introductions, there is a tendency to find the simple four-move schema posited by Swales, but this is by no means the only possibility. In more complex passages, a variety of alternative arrangements are possible, involving repeated use of mainly the 'internal' units Two and Three. In the corpus under consideration, writers normally, though not universally, begin with a One and end with a Four (11/16 times).

Results must be interpreted with caution, since the figures for interrater agreement are low. Hartman (1977) comments:

No entirely agreed upon set of rules for deciding upon an acceptable value for trial (or session) reliability has yet been formulated. (113-4)

He quotes Gelfand and Hartman (1975) as "recommending that...kappa...should exceed 0.6". The statistic, used in the present study, is Fliess' (1971) generalisation of kappa to more than two raters. It is a non-parametric statistic derived from nominal categorical data. It is the most conservative of the measures available (others being pairwise kappa, phi, and percentage agreement). In the
present study, generalizations concerning introduction structure are based upon the 75% of cases where kappa is greater than 0.6.

The unit of coding was the sentence. The sentence was selected as the basic unit of analysis since it initially appeared that writers reflected the traditional conception of the sentence as constituting a complete unit of meaning (or 'thought'). Difficulties were envisaged in choosing a smaller but perhaps less well-defined unit, such as the clause or phrase. However, in some cases, writers utilized the flexibility of language, and complex sentences, to produce a sentence which contained elements of more than one unit of text, according to the system's definitions. Thus

Interest in these reactions has been sparked by the hope that clusters might possess unique catalytic activity, and by the proposal [reference] that clusters might be used as soluble models for metal surfaces.

Bavaro, Montangero & Keister (1983:4977)

In this case, following a previous Type One, the writer continues to refer to the interest of the topic (a major Type One lexical signal) but also references past work and raises two hypotheses. The raters in this case were obliged to decide which of these characteristics was most outstanding - was the writer primarily indicating the interest of the topic, or was he raising hypotheses which had already been mentioned in the literature? This case, and others like it, were not always coded unanimously. Another
particularly clear example of disagreement was the following:

Although fluorescent antibody studies have established the overall distribution of myosin within some vertebrate nonmuscle cells (1,2,8,11,19,20,21-24), the form and detailed arrangement of myosin has remained a mystery. (Herman & Pollard 1981:346)

This sentence might have been written

Fluorescent antibody studies have established the overall distribution of myosin within some vertebrate nonmuscle cells (1,2,8,11,19,20,21-24). However, the form and detailed arrangement of myosin has remained a mystery.

Had this been the case, unanimous agreement might have been expected, to the effect that the first sentence constituted a Type 2, and the second a Type 3. By taking the sentence as the basic unit of coding and forcing a decision, this fuzziness of boundaries necessarily manifested itself in rater disagreement at such points.

Other sources of disagreement are simply what appear to be mistakes, where raters disagreed over what appears to the writer to be a clearcut case. This must be taken as evidence of raters' lack of adequate training. The variability of texts suggest also, that a longer test run should have been undertaken. The effects of raters' understanding of topic are also a possible source of error.

Brown and Yule (1983) make instructive comments about the relationship between "formulaic expressions such as 'Once upon a time'" and topic change as boundary markers.
They point out that topic is something which is difficult to define and perceive clearly. It is perhaps created as the outcome of a negotiation between writer and reader. Brown and Yule refer to it as being located in the writer rather than in the text. Genres which make heavy use of formulaic expressions are better suited to the present type of analysis. Attempting the present form of analysis outside of such genres as the fairy story or the scientific experimental-research paper may be inadvisable.

At the outset of the inquiry, it was not clear what the balance was between topic change and formulaic phrase in determining the divisions between units of text in scientific experimental-research article introductions. Although the genre is formalised, variation exists in the degree to which writers adhere to such formalism. In particular, writers vary in the extent to which they use formulaic expressions to mark transition from one block to another, as opposed to relying merely on topic change. They may be following a simple, standard pattern, but if this is not signaled explicitly, it is left to the discourse analyst's or rater's understanding of the topic as a whole, and relations between different aspects of it, to decide where one block finishes and another starts. Thus, to give a concise example
These methods are useful for comparing intact filaments in different tissues. They do not indicate whether individual filament forming polypeptides (FFPs) have significantly large regions of related aminoacid sequences. (Milstein and McGuire 1981:312)

In this case, even a close reading by a non-specialist may miss the fact that the second sentence is raising a gap. It is necessary, having read the whole text, to come to an understanding of the relationship between aminoacid sequences and FFPs. Such an understanding is naturally taken for granted by the writers. There is also the negative in the second sentence, but this on its own is dangerously little to go on in coding the passage. There is no 'but', no 'a weakness is...', nor any of the large number of other possible lexical signals the writers could have used. A heavy reliance is placed on topic alone.

In the present work, a decision was made to use only non-specialists, principally for practical reasons. Also, even if it had been possible to gather a group of, say, astrophysics experts and have them rate astrophysics papers, conclusions about the structure of the general scientific article introduction would have been impossible. It is necessary to point out that the raters' lack of full comprehension of the topic of some of the articles may have limited the accuracy of their codings. However, it may be argued that although levels of interrater agreement are not uniformly high, this does not invalidate generalisations.
made concerning text structure drawn here. In addition, as de Beaugrande and Dressier (1981) point out

A science of texts demands its own terms and notions because of the nature of its object...we should work to discover regularities...rather than rules or laws....It is the task of science to systemize the fuzziness of its objects of inquiry, not to ignore it or argue it away.

(xiv-xv)

Finally, since it might be expected that any future analysis of this sort would be carried out by materials writers, the development of an analysis capable of being applied by non-specialists is desirable.

Subject-specific variation

There is some suggestion that the incidence of repeated cycles of internal moves increases according to subject: articles in social science journals tended to have much longer and more complex introductions. There were 1.8 times more units in social science introductions as in the other two groupings. Perhaps the newness of the field results in a lack of shared preconceptions and a greater need for both definition of terms and motivation of hypotheses. Considering also articles in the social sciences from outside the corpus itself (but from the same journals and type of articles), there was occasional use of topic-specific subheadings within the introduction, which did not occur in other journals. More significantly, stretches of unreferenced text presenting theoretical background sometimes occurred. Although there were none in the quarter
of the corpus on which rating was done, their existence suggests the possibility of a fifth unit, whose function would be "presenting general, non-referenced theoretical background".

Articles which were not primarily experimental or data-based were not included in the corpus. It should be obvious that articles on pure mathematics, for example, are a type which the current analysis would not apply to, since its units presuppose data-based work in the hypothetico-deductive mode of science. Articles of the social sciences which take mathematics as a model -- economics, for example, and some parts of linguistics -- would equally be unsuitable for the present analytical framework.

Pedagogical and materials-writing implications

Pedagogical implications which may be drawn from this study apply more to materials designed to aid EST reading than writing, because of the product nature of the analysis. Such materials would, as Swales has suggested, be aimed at making salient the possible structures of article introductions, devices used to signal them, and the way relations between the topics of different segments contribute to the text as a whole. However, the main point to be made is that the analysis such materials are to be based on must be a valid reflection of discourse structure in target texts, and proven so. If we inculcate an
inaccurate schema into our ESL/EST students concerning the expectations they should have of a text, we are making their task harder, not easier.

Further research

When the present work was initiated, it was envisaged as merely verifying existing analyses of the 'Introduction', and then going on to deal with the 'Discussion'. Belanger (1984) has begun work on the structure of the 'Discussion' section, which is of course, the natural progression from the present work. Indications are that a structure consisting of a number of basic elements which may be repeated in various subsequences is to be found in the 'Discussion' section, as in the 'Introduction'. It is anticipated that verification of such an analysis using the present technique may be more difficult, because the structure of the discussion section is believed to be less conventionalized and harder for even technical writers to make prescriptions about (Day 1979, see above). Because of the probable greater importance of topic in this section, it may be impractical to attempt an analysis across sciences. Instead, a subject-specific or journal-specific analysis may need to be undertaken, using a group of expert raters.

Finally, the work presented here has focused solely on product. Research in this area is also needed concerning
process, as Swales (personal communication) has suggested. Other techniques besides the use of appropriate schema are used by the good reader. In the present context, a knowledge of how the good ESL/EST reader handles scientific texts would provide information concerning the processes involved in reading. More broadly, research is also needed concerning how such texts are generated. Suitable techniques for research on composing and revising processes already exist (see Hearing 1984 for a recent review). Findings arrived at from this perspective would show the other side of the subject under investigation, and would have implications for the writing of EST materials which could complement those concerning EST reading materials.
APPENDIX: CORPUS


Abbreviations

J. Physiol. = Journal of Physiology
J. Am. Chem. Soc. = Journal of the American Chemical Society
Arch. Gen. Psych. = Archives of General Psychiatry
Am. Soc. Rev. = American Sociology Review
BIBLIOGRAPHY


Garfield, E. (ed.). 1982. *Social Science Citation Index.* Philadelphia: ISI.


Robertson, D. L. 1983. English use, needs and proficiency among foreign students at the University of Illinois at Urbana-Champaign. PhD (Dept. of Edn.) University of Illinois.


Stern, A. A. 1976. When is a paragraph? College Composition and Communication 27:3.


