

LEXICAL COHESION IN STUDENT ACADEMIC WRITING

by

SUSAN LOUISE VAN TONDER

submitted in fulfilment of the requirements for
the degree of

MASTER OF ARTS

in the subject

LINGUISTICS

at the

UNIVERSITY OF SOUTH AFRICA

SUPERVISOR: PROF EH HUBBARD

JANUARY 1999

Student number: 611-158-0

I declare that *Lexical Cohesion in Student Academic Writing* is my own work and that all the sources that I have used or quoted have been indicated and acknowledged by means of complete references.

SL van Tonder

SIGNATURE

(Ms SL van Tonder)

29 January 1999

DATE

Contents Page

SUMMARY

ACKNOWLEDGEMENTS

CHAPTER 1: INTRODUCTION	1
1.0 INTRODUCTION	1
1.1 STATEMENT OF PROBLEM	1
1.1.1 Research questions	1
1.1.2 Background considerations	3
1.1.3 Aims	4
1.2 PRELIMINARY PERSPECTIVE ON KEY CONCEPTS	5
1.2.1 Coherence	5
1.2.2 Cohesion	7
1.2.2.1 Cohesive relations between student writing and the question prompt	9
1.2.3 Lexical errors	10
1.3 RESEARCH FOCUS	11
1.3.1 The corpus	11
1.3.2 Hypotheses	12
1.4 STRUCTURE OF THE STUDY	14
CHAPTER 2: LEXICAL COHESION, LEXICAL ERRORS, COHERENCE AND ACADEMIC WRITING: A SURVEY	15
2.0 INTRODUCTION	15
2.1 COHERENCE	15
2.1.1 Cohesion and coherence	16
2.1.2 Inference in interpretation	20
2.1.3 Situational context	23
2.1.4 Receptor knowledge and degree of interpretability	24
2.1.5 Text strategies, text categories, patterns of exposition and argument	25
2.2 COHESION	26
2.2.1 Categories and subcategories of lexical cohesion	32
2.2.1.1 Repetition	33
2.2.1.2 Synonymy or near-synonymy	35
2.2.1.3 Opposition or contrast	38

2.2.1.4	Inclusion	39
2.2.1.5	Derivation	40
2.2.1.6	Collocation	42
2.2.1.7	Cohesion in first and last paragraphs	44
2.2.2	Quantitative studies of lexical cohesion	47
2.3	ERROR ANALYSIS	57
2.3.1	Description of errors	60
2.3.2	Quantitative studies of error analysis	65
2.4	CONCLUSION	68
CHAPTER 3: ANALYTICAL FRAMEWORK AND RESEARCH PROCEDURES		70
3.0	INTRODUCTION	70
3.1	RESEARCH DESIGN	70
3.2	ANALYTICAL FRAMEWORK	70
3.2.1	Coherence	70
3.2.2	Cohesion	72
3.2.3	Errors	75
3.2.4	Textual units	77
3.3	RESEARCH PROCEDURES	80
3.4	SAMPLE ANALYSES	85
3.4.1	Cohesion analysis	86
3.4.2	Analysis of first and last paragraphs	90
3.4.3	Error analysis	93
3.5	CONCLUSION	96
CHAPTER 4: RESEARCH FINDINGS		97
4.0	INTRODUCTION	97
4.1	RESULTS	97
4.1.1	Coherence Hypotheses	101
4.1.1.1	Cohesion Density - Coherence Hypothesis (General)	101
4.1.1.2	Cohesion Density - Coherence Hypothesis (Specific)	103
4.1.1.3	Error - Coherence Hypothesis	108
4.1.1.4	Cohesion of Opening Paragraphs - Coherence Hypothesis	110
4.1.1.5	Cohesion of Closing Paragraphs - Coherence Hypothesis	111

4.1.1.6	Review of Coherence Hypotheses	113
4.1.2	Academic Achievement Hypotheses	114
4.1.2.1	Cohesion Density - Academic Achievement Hypothesis (General)	114
4.1.2.2	Cohesion Density - Academic Achievement Hypothesis (Specific)	116
4.1.2.3	Error - Academic Achievement Hypothesis	120
4.1.2.4	Cohesion of Opening Paragraphs - Academic Achievement Hypothesis	121
4.1.2.5	Cohesion of Closing Paragraphs - Academic Achievement Hypothesis	123
4.1.2.6	Coherence - Academic Achievement Hypothesis	124
4.1.2.7	Review of Academic Achievement Hypotheses	125
4.2	RELATIONSHIP OF RESULTS TO AIMS	125
4.3	CONCLUSION	128
	Tables	129
	CHAPTER 5: CONCLUSION	131
5.0	INTRODUCTION	131
5.1	REVIEW	131
5.2	APPLIED LINGUISTIC IMPLICATIONS	133
5.2.1	Approaches to the teaching of writing	134
5.2.2	Implications for writing instruction of the findings of the study	138
5.3	LIMITATIONS AND SUGGESTIONS	144
5.4	CONCLUSION	149
	APPENDICES	150
	APPENDIX A: EXAM QUESTION	150
	APPENDIX B: BAMBERG'S (1984) HOLISTIC COHERENCE SCALE	152
	APPENDIX C: LIEBER'S (1979) CATEGORIES OF F-UNIT	155
	APPENDIX D: SAMPLE SCRIPTS AND TABLES	158
	BIBLIOGRAPHY	176

ACKNOWLEDGEMENTS

I would like to thank the following people:

Prof. Hilton Hubbard, my supervisor, for his invaluable guidance, patience and for providing detailed commentary on the thesis;

Joyce Jordaan and Leonie Venter of the Unisa Statistical Services Department for their efficient statistical testing of the data and for always being willing to help;

The Technikon SA library staff and especially Amanda Louw of the Department of Inter-library Loans for her helpful and effective service;

Zoë Bruckmann for her proofreading of earlier drafts of the thesis and for setting me along the right road in terms of style;

My friends and colleagues for their support and particularly Zien Kruger, Eleanor Botha, Moira Sampson and Marie Vosloo who took on extra duties towards the end of this research so that I could take time off to complete it;

My parents, Ann and Frans, and my husband, Deon, for their unfailing enthusiasm and support and Deon especially for his knowledge of the technical details required of the layout and printing of a manuscript of this size.

CHAPTER 1

INTRODUCTION

1.0 INTRODUCTION

This introductory chapter provides an overview of the problem central to this study, namely an exploration of the concept of coherence in terms of the contribution made by lexical cohesive relationships and lexical error to the perception of the coherence of texts of student academic writing. The aims of the study are presented and a preliminary perspective on the key concepts of the study is outlined. The research focus is also given and the chapter ends with a summary of the structure of the study.

1.1 STATEMENT OF PROBLEM

1.1.1 Research Questions

The central problem examined in the present study is to establish whether certain features of the lexis of a text contribute to the text being perceived as coherent. The focus of the study is limited to texts produced by students in an academic context, specifically students studying at university level where generally they are judged fit to pass an academic course based on the written answers they produce in an examination set by the lecturer of the course. The type of text examined in the study can thus be classified as student academic writing. This type of writing has been defined by Hubbard (1989:3) as "that sub-genre of expository writing that is required from students in the course of their study, ... that deals with the specific content particular to the relevant subjects" and that therefore incorporates "a much richer conceptual network" (1989:10) than does composition writing.

Assessment of academic achievement ratings of student writing has traditionally been done according to the overall quality (see for example Grobe, 1981), or in terms of the perceived coherence of the writing (cf. Bamberg, 1984). Hubbard (1989) revealed a significant relation between the achievement level of a text and the

coherence of the text (1989:262). To establish what part coherence plays in the assessment of the overall quality of writing, particularly persuasive writing, McCulley set up a study to determine the relationship between scores on primary-trait measures of coherence and writing quality (1985:270). His findings show that coherence is an important aspect of writing quality. Hubbard, who argues for a recognition of the distinction between the concepts of writing quality and coherence, maintains that although both concepts can be defined as pretheoretical in nature, “coherence’ is somewhat less vague and more amenable to investigation, and also more directly related to ... cohesion” (1989:12) which is one of the variables examined in the present study.

Accepting that coherence is an important measure of the assessment of the overall quality of student writing and that coherence is more readily studied than is writing quality, the present study evaluates student examination scripts for their coherence ratings and assesses the effect of the internal feature of lexical cohesion on whether the scripts are perceived as coherent as well as on their achievement levels in terms of the scores awarded them by the original examiners.

Lexical cohesion is “the cohesive effect achieved by the selection of vocabulary” (Halliday and Hasan, 1976:274) and is, therefore, the greatest resource writers have at their disposal for creating cohesion in texts. Evidence from certain empirical research suggests that lexical cohesion may be used by student writers in an effort to produce coherent writing of acceptable quality (cf. Engber, 1995; McCulley, 1985; Stotsky, 1986; Wessels, 1993; Witte and Faigley, 1981).

In the light of the above, the following research questions are central to the present study:

Does lexical cohesion contribute to the perceived coherence of a text of student academic writing?

Do lexical errors affect the perceived coherence of a text of student academic

writing?

1.1.2 Background considerations

“Writing is frequently a difficult skill for *any* language user, which is to say that writing represents a fairly challenging task for both native and nonnative speakers” (Kroll, 1990: 140). The writing expected of students in an academic context is especially challenging for a number of reasons. Firstly, there are differences between the acts of speaking and writing. Unlike when communicating orally, students who communicate in writing operate in a “context-reduced” situation and are cognitively involved in a highly demanding task (Starfield 1990:145 describing Cummins’ (1984) continua for addressing the range of contextual support and the degree of cognitive involvement in a task). Many students who are competent in face-to-face communicative situations experience difficulty in coping with written tasks (1990:146; cf. Stotsky, 1986:286, 287).

Secondly, the academic community communicates via a type of discourse for which the community itself has “established rules for the making and transmission of meaning” (van Zyl, 1993:52). These rules are adhered to unquestioningly by members of the community and are, therefore, not normally made explicit to the learners. It is partly as a result of this that it is very difficult for students to gain access to the academic discourse community. Bizzell (1986:297) suggests that it is by acquiring the language practices of the academic community that the learner will gain access to it and to its discourse practices and world view. As van Zyl (1993:54, citing Flower, Stein, Ackerman, Kanz, McCormick and Peck, 1990: 222) maintains, “students need to learn the textual conventions, the expectations, the habits of mind, and the methods of thought that allow one to operate in an academic conversation...”.

Starfield (1990) suggests that it is not only people from minority groups (cf. Bizell, 1986:294) who fail to cope in an academic environment. In her analysis of the student population who made use of the Academic Support Programme at Wits

University in the 1980s, she found that the majority of students who required support were black second language speakers of English. These students had been subjected to underqualified and poor schooling, amongst other things. However, many white students whose first language was English and who had not suffered these disadvantages, often also required support. Writing in 1990, she finds that the situation is much the same as it was in 1980. She accounts for this in terms of the lack of preparation for university study given to learners during their school careers.

In support of this, Johns (1986:251) claims that: "Many students at every level are unfamiliar with the conventions of English writing, which if well integrated, result in coherent prose". She goes on to say that the situation is exacerbated by the fact that there are no writing textbooks dealing with coherence features and, therefore, that not enough is taught about the coherence features necessary for acceptable writing.

It is clear, given the importance of academic writing in the assessment of student ability and the difficulty students have in producing this type of discourse coherently, evidenced by the high failure rates at university level (Hubbard, 1989; Wessels, 1993), that there is a need for ongoing research into the area of student academic writing.

1.1.3 Aims

The aims of the present study operate along three dimensions: theoretical, descriptive, and applied.

The first dimension is theoretical. This study seeks to assess the effect of lexical cohesion and lexical errors on the perceived coherence of a text with the ultimate aim of contributing to the body of knowledge seeking to explicate the concept of coherence, regarded by Hubbard (1993:55) for example, as "arguably the most central concept in text linguistics".

The theoretical aims are:

To determine whether and to what extent lexical cohesion contributes to the perceived coherence of a text of student academic writing.

To determine whether lexical errors affect the perceived coherence of a text of student academic writing.

At the descriptive level, the study seeks to describe the types of lexical cohesion relations in student academic texts and assess their effect on the academic achievement levels of these texts (here the marks awarded to the texts as answers to an examination question), and to consider possible connections between coherence and academic achievement. The aims can be written as follows:

To determine whether there is a relationship between lexical cohesion in a text of student academic writing and the academic achievement level of that text.

To determine whether there is a relationship between the density of lexical errors in a text of student academic writing and the academic achievement level of that text.

To determine whether there is a relationship between the perceived coherence of a text of student academic writing and the academic achievement level of that text.

The study also has applied linguistic aims in that it is hoped that insights gained from the aims of this study may be useful in the teaching of writing of an academic nature.

1.2 PRELIMINARY PERSPECTIVE ON KEY CONCEPTS OF THE STUDY

1.2.1 Coherence

Connor and Johns (1990) suggest that while coherence is becoming increasingly interesting to researchers (they cite as examples of research that done by de

Beaugrande, 1980; Carrell, 1982; Connor, 1984a; Enkvist, 1985; and Kintsch and van Dijk, 1978), there are not many practical applications available. The reasons they give for this are the difficulty of defining the concept of coherence and the fact that there are two opposing schools of thought regarding the interpretation of coherence. These approaches see coherence as text-based and reader-based (Johns, 1986:248 - 251) respectively. Enkvist (in Connor, 1987) labels these approaches to the study of coherence the sentence-based approach (which is concerned with the structural elements of a text) and the process-centred approach (which is more concerned with the procedural elements in a text).

Under the first approach at the level of what van Dyk calls "local" coherence (Bamberg, 1983:419), text linguists have traditionally been concerned with the connectedness of discourse (van Dyk, 1977:93), that is, the relations between sentences or propositions. One of the tools for revealing connectedness in discourse is the relation of cohesion of Halliday and Hasan (McCutcheon and Perfetti, 1982:115).

As part of the analysis of local coherence, van Dyk (1977) looks at the 'information distribution' (1977:94) in discourse. The continuity of information is usually discovered in the relation between old and new information which van Dyk labels the topic / comment relation and which is described by various other writers (Clark and Haviland, 1977; McCutcheon and Perfetti, 1982; Weissberg, 1984; Bardovi-Harlig, 1990). The concept *topic* is explained as "the main topic of the sentence" and the *comment* "what is being said about the topic" (Connor, 1987:683). From this approach to coherence have emerged topical structure analyses of text such as that of Witte (1983) (for a critique of this approach see Hubbard 1990).

Van Dyk (1977) points out that not all relationships at the level of local coherence are made explicit within the text. Many relations must be inferred by the reader who has to supply the missing links. The issue of inference in texts has been dealt with by Crothers (1978) and Shiro (1994) amongst others and relates to processes which the reader undergoes in an effort to interact with a text.

The process-centred approaches to coherence tend to focus on propositions rather than sentences with the emphasis being on the argument and predicate of the proposition (Connor, 1987:686). At the level of 'global' or overall coherence (van Dyk, 1977:95), van Dyk looks at the macro-structures of the text which include relations such as those between sets of propositions, and at the topic of discourse.

Another approach at the level of whole text analysis is that adopted by Winter who divides clausal relations in a text into the classes of Logical Sequence and Matching Relations (Hoey, 1983:19). Winter is "concerned to reveal the clause organisation of a passage as a whole without focussing on any one sentence in particular within it" (Hoey, 1983:17).

Apart from the above approaches to coherence, there are other approaches which focus on knowledge structures or schemata brought to the text by the reader (Cook, 1989) and those generated by the genre of the text (Swales, 1990). The concept of coherence and some of the issues which are associated with it will be examined in more detail in Chapter 2 of the present study.

It has become generally acknowledged that both the structural and procedural approaches are vital for a comprehensive description of the concept of coherence (Johns, 1986; Carstens, 1997; Connor, 1987; Hubbard, 1989; Wessels, 1993; see 2.1 below). At present, both approaches mentioned above are in use in writing instruction and some researchers have called for an integration of approaches. For example, Connor cites Phelps (1985) as arguing for "a unified theory in which the 'overarching process' is the cooperative enterprise whereby writers and reader construct meaning together" (Connor, 1987: 677 - 678).

1.2.2 Cohesion

Advocates of both approaches to coherence agree that certain internal features of a text contribute to its coherence:

"If a text is to be well formed, it must have semantic coherence as well as sufficient signals of surface cohesion to enable the reader to capture the coherence The general rule is that every sentence of a well-formed text must have a cross-reference to at least one other sentence of that text, and there has to be an overall coherence involving the text as a whole."

(Enkvist, 1978:126 as cited in Connor and Johns, 1990:1)

Halliday and Hasan in their ground breaking work on cohesion in text linguistics, define cohesion as a semantic concept which "refers to relations of meaning that exist within the text, and that define it as a text" (1976:4). For them:

"Cohesion occurs where the interpretation of some element in the discourse is dependent on that of another. The one presupposes the other, in the sense that it cannot be effectively decoded except by recourse to it. When this happens, a relation of cohesion is set up, and the two elements, the presupposing and the presupposed, are thereby at least potentially integrated into a text."

(Halliday and Hasan, 1976:4)

Although cohesive relations are semantic relations, they are realised through the lexicogrammatical system of a language. Halliday and Hasan (1976) distinguish between grammatical forms of cohesion and those forms which are revealed through vocabulary. The grammatical forms of cohesion, they list as: ellipsis, substitution, reference, and conjunction, although conjunction has in addition to this a lexical component to it. The form of cohesion realised through vocabulary is called lexical cohesion and is the focus of the present study.

As suggested above, cohesion is relational, with one instance of a cohesive relation labelled a tie. Specific texts can be characterised by the types and numbers of ties they display. Cohesive relations often weave through texts in the form of chains which occur when elements refer back to other elements which in turn refer back to others in the text, and so on. A cohesive item that presupposes an element which has occurred previously in the text participates in an anaphoric relation. By contrast, any item which points forward to a presupposed item participates in a relation of cataphora (Halliday and Hasan, 1976:17).

Halliday and Hasan (1976) divide lexical cohesion into the categories of collocation and reiteration which latter includes the sub-categories repetition, synonymy and general nouns. Over the years, these categories have been refined by various writers and researchers in order to distinguish more precisely those lexical relations used by writers in constructing different genres of texts.

A refinement of this model to account for the use of lexical cohesion in expository writing has been developed by Stotsky (1983). Stotsky used her model to reorganise Halliday and Hasan's (1976) categories into the categories semantically related items and collocation. The first category included lexical items related semantically and contained the following sub-categories: repetition, synonymy, opposition, inclusion, and derivation. The second, the category of collocation, contained words which are only related in terms of their potential to co-occur in similar texts.

Stotsky's sub-categories of synonymy, opposition and inclusion signify lexical or sense relations that hold between the vocabulary items of a language. These are apart from the sub-category of repetition where a relation is established by the repetition of a lexical item later in a text, and the sub-category of derivation where words are derived from the base word by the addition of a suffix or a prefix (Jackson, 1988:31). Leech (1981) and Lyons (1977b) (as cited in Malmkjaer, 1991:300) define synonymy, antonymy (or oppositeness) and hyponymy (an inclusion relation) as primitive semantic relations. These sub-categories are not always easily specifiable. The category of collocation was not examined in the present study.

1.2.2.1 Cohesive relations between student writing and the question prompt

In addition to the study of the relationship between cohesion and coherence in student academic writing, it is worth examining the extent to which student writers cohesively link their writing to the task instruction or question prompt to which their writing is a response. Part of the study of coherence consists of how texts are organised in terms of the presentation and development of ideas within the context of the task with which the writer is presented. What is of particular interest is whether

the establishment of both an introduction to the essay which links the essay to the specific instructional task, and a sense of closure which rounds off the essay by setting it once more within the context of the task, creates a perception of coherence for the reader of the essay. Thus as part of its overall investigation of lexical cohesion, the present study also focussed on the density of cohesive ties between the opening and closing paragraphs of the corpus of scripts under analysis and the question prompt to which they were a response and correlated this density with the coherence ratings of the scripts.

1.2.3 Lexical errors

A further feature which might be influential in determining whether a text of student academic writing is assessed as worthy of a high coherence rating is that of lexical error. A lexical error is defined as occurring when a lack of knowledge is exhibited about either the meaning or the form of a lexical item (Engber, 1995:146). There is still debate surrounding the influence on writing of lexical proficiency. Vocabulary does seem to play an important role in student writing. Engber cites Raimes (1985) and Spack (1984, as cited in Engber, 1995:140) who found that a lack of a strong vocabulary hindered student writing. A strong vocabulary is particularly important in the case of examination writing where student writers are judged on the written essays they produce within a limited time frame. Engber (1995:140 - 141) admits that knowledge about the relationship between lexical proficiency and the judgements of examination markers is still incomplete, but confirms that her own research has demonstrated that in fact lexical proficiency does affect reader judgement of the quality of examination essays. Citing Grabe (1985 in Engber, 1995:141), she concludes that: "Although no comprehensive model of what constitutes a communicative text has yet emerged, there is general agreement that the lexicon is a significant component in both the construction and interpretation of meaningful text". The present study set out to ascertain whether accuracy of word choice and form played a significant role in reader perception of coherence in student academic texts.

1.3 RESEARCH FOCUS

1.3.1 The corpus

The texts examined in the present study were the written examination answers to the following question:

In what ways do conditions on *Animal Farm* differ from those on *Manor Farm*?

(For the full question including the extract from *Animal Farm* provided as background to the question, see Appendix A, page 150.)

The question required the students to write an answer of an expository nature based on the discussion of a literary text, in this case *Animal Farm* by George Orwell.

The study looked at a random selection of texts, written by undergraduate university students studying the course: *Practical English* within the Department of English. The texts were analysed for examples of lexical cohesion and lexical errors by the researcher. They were assessed separately by three independent raters for their coherence level using Bamberg's (1984) four-point holistic coherence scale (see Appendix B, page 152). A correlation was made between the academic achievement of each text (in terms of the score each text was awarded by the examiner) and the coherence level of the text. The density of lexical cohesion devices was then correlated with the coherence ratings to assess the affect of lexical cohesion on coherence ratings in examples of student academic writing. Likewise, the errors in lexis were assessed and correlated with the coherence ratings of the texts. In addition, the densities of the lexical ties between the first and last paragraphs of the essays and the question prompts to which they were a response were correlated with the coherence ratings of the texts.

1.3.2 Hypotheses

The hypotheses, presented as non-directional, can be divided into two groups, one relating to coherence and the other to academic achievement. The hypotheses have been written non-directionally as, while there is some support in relevant literature to suggest that lexical cohesion may play a part in determining the coherence or in contributing to the academic achievement levels of texts of student writing, evidence from other research finds no support for these hypotheses. Likewise, evidence suggesting that errors affect the coherence of texts is conflicting. While there seems to be evidence suggesting support for the hypothesis that the perception of coherence in student writing correlates with academic achievement levels, the research base for this is not large and it has, therefore, been decided to present the hypothesis relating these two variables as a non-directional one.

Coherence Hypotheses:

H₁ Cohesion Density - Coherence Hypothesis (General)

There is a relationship between the density of lexical cohesion ties in a text and the holistic coherence rating of that text.

H₂ Cohesion Density - Coherence Hypothesis (Specific)

There is a relationship between the density of each category of lexical cohesion (repetition, synonymy, opposition, inclusion, derivation) in a text and the holistic coherence rating of that text.

H₃ Error - Coherence Hypothesis

There is a relationship between the density of lexical error in a text and the holistic coherence rating of the text.

H_{4a} Cohesion of Opening Paragraphs - Coherence Hypothesis

There is a relationship between the density of lexical cohesion ties between the opening paragraph of a text and its question prompt on the one hand and

the holistic coherence rating of that text on the other.

H_{4b} Cohesion of Closing Paragraphs - Coherence Hypothesis

There is a relationship between the density of lexical cohesion ties between the closing paragraph of a text and its question prompt on the one hand and the holistic coherence rating of that text on the other.

Academic Achievement Hypotheses:

H₅ Cohesion Density - Academic Achievement Hypothesis (General)

There is a relationship between the density of lexical cohesion in a text and the achievement level of that text.

H₆ Cohesion Density - Academic Achievement Hypothesis (Specific)

There is a relationship between the density of each category of lexical cohesion (repetition, synonymy, opposition, inclusion, derivation) in a text and the achievement level of that text.

H₇ Error - Academic Achievement Hypothesis

There is a relationship between the density of lexical errors in a text and the achievement level of that text.

H_{8a} Cohesion of Opening Paragraphs - Academic Achievement Hypothesis

There is a relationship between the density of lexical cohesion ties between the opening paragraph of a text and its question prompt on the one hand and the achievement level of that text on the other.

H_{8b} Cohesion of Closing Paragraphs - Academic Achievement Hypothesis

There is a relationship between the density of lexical cohesion ties between the closing paragraph of a text and its question prompt on the one hand and the achievement level of that text on the other.

H₉ Coherence - Academic Achievement Hypothesis

There is a relationship between the coherence of a text (HCR) and the academic achievement level of that text.

1.4 STRUCTURE OF THE STUDY

Chapter 2 presents conceptual definitions of the variables of the study and places them within the context of relevant literature. Chapter 3 outlines the analytical framework of the study and the research procedures adopted within which to carry out the study. Chapter 4 presents the findings of the study and provides an interpretation of these findings. Finally, Chapter 5 assesses the findings in terms of their applicability to writing instruction and reviews the limitations of the study while also making suggestions for possible future research.

CHAPTER 2

LEXICAL COHESION, LEXICAL ERRORS, COHERENCE AND ACADEMIC WRITING: A SURVEY

2.0 INTRODUCTION

This chapter looks at the conceptual definitions of the variables relevant to the present study. It looks at the major theories underlying these concepts and at the debates they give rise to and reviews some of the quantitative literature relevant to the study.

2.1 COHERENCE

Enkvist (in Connor and Johns eds, 1990:13) defines the concept of coherence as a hermeneutic phenomenon. For him the receptor of a text must be able to understand the text to accept it as being coherent. Enkvist (1990) introduces seven problems associated with discourse and the interpretation of it. All seven problems relate to the problem encompassed in the "analysis of the features of text, discourse, and communication that give the receptor an intuition of coherence and make the text interpretable" (1990:11). Enkvist approaches the problems related to textual coherence from the point of view of the interpreter and argues for a process approach to the study of coherence. However, many of the issues he raises are associated with the features of text that make it coherent and although these problems do not cover every issue associated with coherence, they are relevant to the issues dealt with in this study. They contain important aspects which will be highlighted in the following section.

2.1.1 Cohesion and coherence

The first problem Enkvist (1990) outlines is the difficulty of defining the relation between cohesion and coherence. He maintains that the positivist approach traditionally followed by linguists which has encouraged a focus on "concretely describable entities" (1990:11) has meant that text linguists and discourse analysts have tended to focus on formal cohesive relations as evidence of what combines sentences into texts.

Halliday and Hasan (1976) do not equate cohesion with texture (their term for what is usually referred to as coherence - Carrell, 1982:480), in that texture is made up of register and cohesion which together define a text (1976:23). At the same time they suggest that cohesive ties play an essential role in the identification of text in that "... cohesive ties ... are the only source of texture" (Halliday and Hasan, 1976:9). Brown and Yule (1983:195) identify an ambivalence in Halliday and Hasan (1976) towards cohesive relations which they identify as "semantic relations" (1976:13). It is the semantic relations in a text which are usually associated with the power to produce coherence (cf. Witte and Faigley, 1981:202). However, Halliday and Hasan (1976) seem to suggest that it is the explicit formal cohesive ties that create "texture" rather than the underlying semantic relations.

De Beaugrande and Dressler (1981) who operate from a procedural point of view, identify cohesion and coherence as two of seven standards of textuality necessary for a text to be defined as "communicative" (1981:3). Texts which do not communicate are regarded as "non-texts". Cohesion is defined as "the ways in which the components of the surface text, i.e. the actual words we hear or see, are mutually connected within a sequence" (1981:3). Cohesive devices operate according to the grammatical conventions of a text and include any function which is used to indicate a relationship.

Coherence “concerns the ways in which the components of the textual world, i.e. the configuration of concepts and relations which underlie the surface text, are mutually accessible and relevant (1981:4). De Beaugrande and Dressler (1981) maintain that relations of coherence are not always made explicit in the text in which case readers supply the relations they need to understand the text. This seems to suggest that formal cohesive relations are not always necessary for the identification of coherence relations.

Nevertheless, both Halliday and Hasan (1976) and De Beaugrande and Dressler (1981) seem to conclude that cohesion is an essential component of coherent text. This fact is disputed by amongst others Enkvist (1990) and Carstens (1987, 1997) who provide examples to show that cohesion is neither a necessary nor sufficient condition for a text to be interpreted as coherent (Carstens 1987:27, 1997:117 citing Brown and Yule, 1983: 194 - 199). Carstens demonstrates that a text which is linked by examples of various categories of lexical cohesion may nevertheless be incoherent if it displays no meaning relations (1987:28). Enkvist provides the following text which despite the presence of repetition links may be considered incoherent:

“My car is black. Black English was a controversial subject in the seventies. At seventy most people have retired. To re-tire means “to put new tires on a vehicle.” Some vehicles such as hovercraft have no wheels. Wheels go round.”

(Enkvist, 1990:12)

Although cohesive links are present in this text, there are no underlying semantic meaning relations and the reader, therefore, cannot derive a “consistent world picture” (1990:12) from the text.

Enkvist’s second example does not provide the reader with a consistent world picture either. In this case the cohesive ties in the text signal links between sentences that relate to a non-existent world.

"Suzie left the howling ice cube in a bitter bicycle and it melted. It soon tinkled merrily in her martini. Into her drink she then also poured the grand piano she had boiled in a textbook of mathematics the night before. She chewed the martini, read the olive and went to bed. But first she took her clothes off. She then took her clothes off."

(Enkvist, 1990:12)

Neither of the above texts can be summarised or interpreted for meaning. This suggests that cohesive links are not sufficient to produce a coherent text. It is a fact, though, that readers are inclined to interpret any text as coherent and will look for meaning relations which might suggest a plausible situation in what would at first glance be an incoherent text. Charolles (1985:2 - 3) points to the fact that in theory all texts are coherent in that it is always possible to conjure up a particular situation in which meaning is possible. While supporting this view in that one could envision certain apparently incoherent texts as examples of texts from the genre of dramatic literature, Carstens suggests that one would not succeed in ascribing coherence to such texts if they were to be interpreted as informative texts (1987:29, 1997:119).

Enkvist produces a text such as the following as a contrast to those above:

"The net bulged with the lightning shot. The referee blew his whistle and signalled. Smith had been offside. The two captains both muttered something. The goalkeeper sighed for relief."

(Enkvist, 1990:12)

Enkvist maintains that the difference between the above three texts cannot be explored syntactically. They are only distinguished in terms of meaning. The third text is interpretable in that a reader with a prior knowledge of soccer can create a text world which makes sense. The lexical items such as "net", "shot" and "offside" form collocational links in that they belong to the same semantic field.

Widdowson (1978:29 as cited in Brown and Yule, 1983:228) provides an example

from a conversational sequence which appears to have no formal cohesion at all but which is readily interpretable within the genre of conversation:

A: That's the telephone.

B: I'm in the bath.

A: O.K.

Widdowson maintains that this sequence is interpretable as a conventional sequence of utterances each performing an action. In terms of a conversational sequence one might imagine a plausible situation in which such a conversation might occur where A hears the telephone ring and asks B to answer it. B explains that she or he cannot and gives the reason for it and A then agrees to answer it. The fact that texts such as the one above can be interpreted with little or no evidence of formal cohesion suggests that cohesion is not a necessary condition for coherence.

Carrell (1982) who operates from a schema-theoretic orientation maintains that the reader interacts with the text using background knowledge in conjunction with the content and structure of the text to arrive at its meaning. She cites Steffensen (1981, as cited in Carrell, 1982:485) who maintains that "textual cohesion represents only a potential which can be fully realized only when the reader appropriately identifies the schema underlying a text". Shiro (1994:175) suggests that an important part of creating the coherence of a text is derived from the reader's ability to provide inferences to build up a textual world. The text can never be totally explicit and the reader in assuming that the text is coherent looks for the writer's intention and makes an effort to find meaning in the text by supplying the missing links in the text. Shiro (1994:176) maintains that "... the reader's interpretation results from decoding the linguistic signs that appear in the text combined with other processing strategies based on his/her world knowledge and other cognitive abilities". Brown and Yule (1983) argue that far from being

textually explicit:

“ ... it is typically the case that the texts which a reader will normally encounter will show a minimal amount of formal cohesion, assume massive amounts of existing background knowledge, and normally require the reader to make whatever inferences he feels willing to work for in order to reach an understanding of what is being conveyed.”

(Brown and Yule, 1983: 269 - 270)

If cohesion is neither a necessary nor sufficient condition for coherence it is worth asking whether it can be usefully related to the concept of coherence. Hoey (1991:12) who defines coherence as an aspect of the reader's evaluation of a text, suggests that cohesive devices are objective, easily recognisable features of text which can be assessed in terms of the extent to which they influence the reader in finding the text coherent. Carstens (1997) adds that cohesion is a useful aid in the creation of texts. Most texts have a clear cohesive structure and it is suggested that this can strengthen coherence (Weideman, 1984:73 as cited in Carstens, 1997:121; cf. 2.2.2 which details the results of empirical studies relating certain types of cohesion to the coherence of student texts in particular). Cohesion can also contribute to compactness of a text (Donnelly, 1994:111 in Carstens, 1997:121; Yde and Spoelders, 1985) thereby making it easier to read. The present study followed the tradition of empirical research which has correlated various categories of cohesion with coherence in texts of student academic writing in an effort to establish whether and to what extent specific cohesive relations have an effect on the coherence of these texts. It is hoped by doing this to explicate more precisely the concept of coherence.

2.1.2 Inference in Interpretation

The study of inference as a source of coherence in text focuses on the contribution made by the reader to supply information not provided in the text but which is understood to have been part of the intended message of the writer

(Brown and Yule, 1983; Enkvist, 1990; Shiro, 1994). Inferences are made when readers make connections or bridge gaps in texts (Brown and Yule, 1983:257; de Beaugrande and Dressler, 1981:101; Sanford and Garrod, 1981).

The notion of inference in natural language is a rather loose one in comparison with that studied in fields such as mathematics and logic, and is one based on socio-cultural knowledge (Brown and Yule, 1983:35). The study of inference in language has primarily been the domain of cognitive psychologists as part of the study of human comprehension generally. Natural inferences, which are inferences based on word knowledge, have been studied experimentally (Crothers, 1978:51) by both cognitive and educational psychologists. This type of study appears to have been largely theoretical, however, using the specially constructed "two-sentence text" as representative of the type of language encountered by the language user (Brown and Yule, 1983:262) rather than authentic texts. Inference has not been dealt with in much detail by text linguistic theories possibly because, as Enkvist has stated, inference "refuses to yield to assays by traditional linguistic methods" (1990:17) and because inference is better approached intuitively rather than analytically (Brown and Yule, 1983:269; but cf. Crothers, 1978, for a proposed typology of inferences). However, Brown and Yule (1983:269) criticise the 'intuitive approach' taken by psychologists towards the study of inferences as dealing in "simplistic analytic metaphors" and recommend that techniques be developed to enable one to draw conclusions about the way people understand texts in "real-life" situations.

There have been very few empirical studies dealing with inference in authentic texts. One such study has been that done by Shiro (1994) who conducted an experiment in which three groups of readers analysed a journal article for inferences. As a result of her experiment, Shiro concluded that inferences are usually formed by making meaning relations beyond the level of the sentence (Shiro, 1994:177) either with other parts of the text or by relating information to

previous reader knowledge, and that much textual meaning is constructed through the use of inferences.

Accounting for how inferences are first recognised by the reader and then made is a controversial issue. Apart from the inaccessibility to measurement of the mental processes used by the reader in recognising and making inferences, other problems include the impossible task of trying to delimit "all of the world knowledge that could ever be tapped to execute a text inference - that is to say, all knowledge" (Crowthers, 1978:57). There is also the problem of establishing how readers continuously bridge gaps as they proceed through a text. Brown and Yule maintain that in the process of reading, readers make inferences which they later reject as they receive new information from the text (1983:34; cf. Hubbard, 1989:35). Shiro's research (1994:172) showed that readers can hold two contradictory inferences at the same time, one of which they will eventually reject depending on what information is presented later in the text. To prevent themselves from being involved in a process of continual inferencing, Enkvist has argued that readers have a "built-in stopping mechanism" (1990:18) which prevents them from making further inferences once a particular gap has been bridged.

The issue of establishing how readers recognise when an inference is required is equally difficult. Shiro (1994) maintains that inferences are elusive in that they are only identified as such if readers have to make an effort to bridge a gap in the text. The subjects of her study found it difficult to recognise the statement *he settled down with a book* as requiring the inference 'he read the book' (1994:172). It also appears that inferred propositions are stored in the memory with the same status as are stated propositions (Kintsch, 1972, as cited in Markels, 1983:455). A further contributory factor to the difficulty of recognising inferences is the fact that by definition they are not present in the text (Shiro, 1994:167). Shiro suggests that in order to distinguish between what is stated in a text and what

requires inferencing, one should look at the difference between explicit and implicit information. She cites Sperber and Wilson (1986 as cited in Shiro, p. 169 - 170) who maintain that language is never completely explicit and that both explicit and implicit information require inferencing in order to be understood. They do, however, distinguish between “explicatures” (assumptions derived from explicit information) and “implicatures” (interpretations made on the basis of what the reader expects the writer’s intentions to be). They conclude that there are degrees of explicitness in information.

The reason for readers making inferences appears to be the desire to create coherence (Crothers, 1978:56; Hubbard, 1989:35) or cohesion (Markels, 1983:460) from what they read. Evidence that favours an intuitive approach to the study of inferences suggests that how readers carry out inferences is by drawing upon mental processes (Crothers, 1978; Sanford and Garrod, 1981; Shiro, 1994). It is generally agreed that the process of inferencing depends on the interactional relationship between text writer and reader, the background knowledge of the reader, the situational context and the text itself (Brown and Yule, 1983; Crothers, 1978; Enkvist, 1990). To these Shiro adds the aspects of reader variability, which includes experience in reading, and reading purpose which determines the depth of text processing (1994:177-178).

2.1.3 Situational context

For Enkvist (1990:19) a text does not contain a single unchangeable meaning but depends for its interpretation on the receiver’s expectations of it and on the situation in which it is produced. Together the receiver and the situation form the context in which the text is understood. Readers interpret texts on the basis of their knowledge of the situational context and on their experience beyond the immediate context (Enkvist, 1990:18). When readers attempt to interpret texts, they activate relevant knowledge which they use to supply missing links between

the propositions of the text. Enkvist (1990:20, citing Joos, 1972) explains that readers always choose to interpret texts using the most likely and least informative interpretations available to them. To evaluate the possibility of other interpretations, readers rely on what Enkvist (1990:20) calls “situational relevance”, choosing what is relevant from the situation to interpret the text.

Situational relevance exists at a number of levels, the most important of which for the purposes of the present study are the intratextual and the intertextual levels. At the intratextual level the reader uses textual cues such as those provided by cohesive relations within the text to interpret the text. Situational relevance is also appropriate for the reader at the intertextual level where the reader’s interpretation of a text is based on knowledge of similar texts they have encountered in the past. The present study analysed examination answers written in response to a specific question prompt. The question prompt creates certain expectations in the reader’s mind which are fulfilled to a lesser or greater extent by the text which follows. Of particular interest to this study, is the way writers organise their texts in response to the question prompt. Hence, the study adopted an analysis of the way writers introduce and conclude their writing by referring to the question prompt using lexical cohesive ties.

2.1.4 Receptor knowledge and degree of interpretability

Enkvist (1990:21) argues that the previous knowledge of the interpreter plays a crucial role in interpreting the message and metamessage of the text. He defines the metamessage as the “symptomatic meaning” of the text which reveals “something about the state, condition, and intentions of its producer” (Enkvist, 1990:14).

In the process of comprehending, “the reader is confronted with words and sentences on the page, and by applying the appropriate mental machinery ends

up with something in his mind which captures the gist of what has been written" (Sanford and Garrod, 1981:3). This mental machinery includes the reader's ability to make inferences about the text. It also includes the ability to activate appropriate background knowledge from organised knowledge deposits known as schemata (Cook, 1989; Enkvist, 1990:21) in order to access the text. It is the task of the writer to predict what the reader needs in order to interpret the text and to provide adequate cues for the reader to use in understanding the text.

2.1.5 Text strategies, text categories, patterns of exposition and argument

Enkvist (1990:22, citing Lindeberg, 1985b) argues that successful writing is judged by the way the writer combines propositions into arguments rather than by the number of cohesive ties in the text. He states that propositions are arranged hierarchically into specific patterns or "macrostructures" (Enkvist, 1990:22) according to text type. The writer employs a text strategy which is defined as a principle used to structure how the text is organised from its macropatterns right down to syntactical organisation. Enkvist (1990) suggests that many texts have conventional patterns and that writers employ conventional strategies in creating these patterns. Hence, "optimal coherence results from conforming to the optimal strategy" (1990:23).

The corpus of the present study consisted of scripts within the genre of student academic writing. In this case the scripts were structured in accordance with constraints provided by an examination question which required an expository analysis of a literary text. The text strategies employed by the writers would, therefore, centre around structures associated with expository argumentation but would also have to incorporate the patterns associated with narrative writing as the writers integrate a narration of relevant parts of the literary text into their writing.

Where the present study differs from the argument of Enkvist (1990) outlined above, is that it does not consider macrostructures to be the only indicators of coherence within the genre of student academic writing. It attempts to establish to what extent densities of lexical cohesive relations in student texts contribute to the perceived coherence of these texts and whether they are judged successful in terms of academic score.

For Enkvist (1990) text strategies are necessary principles used in the process of text construction. He argues that the product approach to text, which regards language as structure describable in terms of rules, is inadequate in that it cannot account for the type of language produced in poetry and advertising which is able to communicate despite not being structurally well-formed (Enkvist, 1990:25). He suggests that the production and reception of text must take into account the writer, the reader and the environment and that a coherence model should be a process model which "is sensitive to situation and context including the world knowledge of the communication partners" (Enkvist, 1990:26).

Coherence is, therefore, made up of many more determinants than textual features. However, it seems as if cohesion has a role to play as a useful device in the construction of coherent texts (cf. Carstens, 1997; see 2.1.1 above). The next section will outline some important classificatory models of cohesion and some of the empirical studies which have been undertaken to examine the influence of cohesion densities on the perceived coherence of student texts.

2.2 COHESION

In Chapter 1, the concept of cohesion was examined and Halliday and Hasan's (1976) model outlined. This section deals with the concept of lexical cohesion and the various theories and literature related to it.

Halliday and Hasan (1976) see lexical cohesion as being formed of two aspects: reiteration and collocation. Under reiteration, they include: repetition, synonym or near-synonym, superordinate, and general word (1976:278). Typically, a reiterated item in a tie will share the same referent as the item it presupposes (1976:319), a relationship that will be revealed by the determiner *'the'*. However, according to them, reiterated items are cohesive regardless of whether there is identical reference; cohesion is achieved as a relation of the lexical forms themselves (1976:284) and every lexical item has the potential to form a cohesive relationship (1976:288). Apart from the relation of identical reference, Halliday and Hasan, therefore, make provision for cohesive relations where the lexical elements are related by inclusion, and exclusion, or where they are unrelated.

Collocation is achieved through the association of lexical items that occur in the same environment regularly (1976:284). Under collocation, Halliday and Hasan (1976) include relationships of oppositeness, ordered and unordered lexical sets including co-hyponyms, and lexical items whose cohesive relationships are probabilistic rather than systematic (Carter, 1987:73). For Halliday and Hasan (1976) all cohesive relations are specific to the text in which they occur.

Although Halliday and Hasan's (1976) approach to lexical cohesion is the most well known, there are other approaches to the topic which provide useful insights particularly within the context of student expository writing.

Exploring whether "good writing" implies 'coherent meaning' and "coherence is expressed partially through linguistic cohesion" (Hartnett, 1986:143), Hartnett reorganises Halliday and Hasan's (1976) classification of cohesive relations to examine their effect on coherence. Hartnett divides cohesive relations into the two subcategories of static and dynamic ties according to their function in creating discourse. Static ties are those ties that hold attention on a topic (1986:142), whereas dynamic ties are used to manipulate ideas rhetorically.

Under static ties, Hartnett includes several cohesive devices, amongst them the following lexical relations:

- repetition of the same lexical item
- synonyms, near-synonyms, antonyms and collocations (which she defines as items expected in the environment of a given lexical item)

Dynamic ties include the following lexical ties:

- superordinates (which identify high-level logical relationships such as those expressed in expository definitions)
- hyponyms (which identify low-level relationships such as those expressed by illustrations of concepts. Hartnett points out that in both inductive and deductive writing superordinates and hyponyms contribute to logical coherence among main arguments.)
- comparative and superlative forms of adjectives and adverbs (textual signals of comparison and contrast)

Dynamic ties are usually non-essential for the writer and sparsely used as compared to static ties which are essential and less difficult for the writer to construct. Dynamic ties may also have the effect of complicating text to the extent that it makes it seem dense and possibly incoherent. Hartnett suggests that cohesive devices in texts of student writing should be defined in terms of their function in that writing.

One of the most important revisions of Halliday and Hasan's original model has been made by Hasan (1984) who expands the category of lexical reiteration to include the subcategories of antonymy, hyponymy, and meronymy. She excludes collocation on the grounds of the potential for "inter-subjective reliability" (1984:194), but adds the category of instantial cohesive devices to account for

certain relations the writer sets up specific to a particular text as opposed to the general relations the writer displays which are part of the language system. She believes that the counting of numbers of ties in text is not a productive approach to textual analysis and advocates "cohesive harmony" which involves the analysis of text for chains of lexical and grammatical cohesion (1984:205).

Another important adaptation of the original approach is that developed by Stotsky (1983). Acknowledging the importance of Halliday and Hasan's (1976) approach to the analysis of lexical cohesion in text, Stotsky (1983) finds two major flaws in their approach which make it inadequate for the analysis of texts of academic expository writing which she maintains appear to be "characterized by a large, diverse, and highly literate vocabulary and by a richness of cohesive ties established through its vocabulary" (1983:440).

The first major flaw she finds concerns the subcategories posited by Halliday and Hasan (1976) under the category of reiteration. They do not acknowledge the cohesive relationships created through derivations or derivational elements. For them a lexical item is not bound "to a particular category, or to a particular morphological form" (Halliday and Hasan, 1976:291), and thus they would consider, for example, the words *differ*, *different* and *difference* as examples of the same lexical items, (see 3.2 for a discussion of the concept of lexical item). Halliday and Hasan subsume relationships of derivation under repetition and miss potentially important lexical relationships (particularly in expository writing) within the text.

Stotsky argues that there are important stylistic reasons for examining relations of derivation in expository writing. In an earlier work she suggests that academic writing is marked by the use of inflected or suffixed words (Stotsky, 1981:20), partly due to a preponderance of Latinate words (Latin being the language in which scholarly writing took place for hundreds of years). Academic writing also

has a tendency towards nominalisation. Derivatives allow writers of academic prose the flexibility to create a nominalisation style which allows them a precise focus on the subject of their writing and an increase in the density of ideas required for concise expression.

The repetition of derivational elements gives the writer a further stylistic advantage in that it gives already related lexical elements an added source of cohesion. Multiple ties strengthen the cohesion of a text and provide an extra source of readability for the reader.

Apart from the absence of a subcategory of derivational cohesion, Stotsky finds the subcategory of superordinate inadequate in that it only allows for the possibility of a tie formed by a superordinate item following a subordinate item whereas it does not acknowledge the possibility of a subordinate word following a superordinate item. Halliday and Hasan define the latter as a relation of hyponymy which they view as a type of synonymy (1976:338, as cited in Lieber, 1979:142). Stotsky considers both types of tie to be equally possible in expository writing and believes that these should therefore be accorded equal status as cohesive devices (see 3.2).

The second major flaw that Stotsky finds in the Halliday / Hasan (1976) approach is the lack of a consistent principle on which to base their division into the categories: reiteration and collocation.

Stotsky questions the grounds on which items are grouped together under the category of reiteration, particularly as these items do not necessarily share the same referent. She finds the category of collocation to be particularly confused, containing as it does items related only in terms of their probability of frequent co-occurrence as well as items which are systematically related as members of ordered or unordered sets, through opposition or as co-hyponyms.

Stotsky suggests a scheme which reorganises and expands Halliday and Hasan's approach to make it more appropriate for analysing the lexical cohesion of expository writing. She proposes a framework consisting of two categories; one containing semantically related words and the other containing collocationally related words. Words that fall under the first category are those which are related to each other systematically in that they have a stable relationship in the language, as well as specifically (what Hasan, 1984, would call instantially) within the context of the text (1986:280). Under this category she includes the subcategories: repetition, synonymy or near-synonymy, opposition, inclusion as a coordinate, superordinate, or subordinate member in an ordered or unordered set and derivation. Stotsky points out that these categories are well known in vocabulary instructional materials where they are used to teach how words are related to each other (1983:446, footnote 8).

The above category she distinguishes from the category containing collocationally related items which she defines as items that are related, only in so far as they tend to appear together in similar texts. Stotsky suggests that her framework might provide a more comprehensive approach from which to analyse lexical cohesion not only in expository writing, but also in other types of discourse particularly as her classification contains the original Halliday and Hasan subcategories. The present study adopted Stotsky's (1983) classification of lexical cohesion to assess the effect of the density of lexical cohesive ties on scripts of student academic writing. Stotsky's (1983) subcategories of reiteration are potentially better able to discriminate between the types of lexical relations that typically occur in expository writing. The fact that she distinguishes relations of derivation, inclusion, and opposites as separate categories allows a close analysis to be made of the way writers connect semantic units (Stotsky, 1986) to signal the development of their ideas. The study only examined relationships of reiteration, excluding collocational relationships from the analysis on the basis of the potential difficulty of establishing the reliability between raters' identification of

collocational ties (cf. Hasan, 1984:194; see 2.2 above and 2.2.1.6 below).

Another classificatory model of lexical cohesion has been developed by Hoey (1991). This classification has been the basis of an empirical study undertaken by Wessels (1993) on the effect of lexical cohesion on perceived coherence in student academic writing. Wessels (1993) found that many features contained in Hoey's (1991) model were important contributors of perceived coherence in the student academic writing she analysed. The present study is closely related to Wessels' (1993) study in terms of its aims and subjects. The study seeks to assess the effect of the density of lexical cohesion ties on the holistic coherence of texts of student academic writing. The present study however attempts to build on Wessels' (1993; 1994) research by analysing another well known classification of lexical cohesion, namely that of Stotsky (1983) and by using as data the scripts of a slightly different sub-genre, that of student expository analysis of literary text.

The next section will be devoted to looking at each of Stotsky's (1983) categories and subcategories in more detail.

2.2.1 Categories and subcategories of lexical cohesion

Carter (1987:72) argues that the study of lexis in discourse accounts for:

“... more obvious features of cohesion such as repetition of items as well as more complex relations of collocation and of structural-semantic sense connections across sentence boundaries.”

One approach towards accounting for the structural-semantic sense relations between lexical items in text is that of field theory which holds that vocabulary is organised into lexical or conceptual fields (Lehrer, 1969:39). Carter (1987:50) defines field (or field of discourse) as “the particular activity, cultural feature,

social institution or topic for which a particular set of ideationally related lexical items is often evolved or adapted". Chapman's approach is that lexical fields are groups of words which cluster around a particular concept (Chapman, 1983:97). Chapman cites the examples of the semantic fields of colour names given by Fillenbaum and Rapaport (1971:53, as cited in Chapman, 1983:98) who maintain that "the meaning of a lexical item is a function of the set of meaning relations which hold between that item and the other items in the same domain" (1971:3, as cited in Chapman, 1983:99). Although Lehrer states that only some words are able to be easily described in terms of fields (as cited in Carter and McCarthy, 1988: 21), she does find that the vocabulary of each field is highly organised (see her study on cooking terms - 1969:39). She maintains that the meaning of a lexical item can be described in terms of its sense relations of synonymy, antonymy, incompatibility and hyponymy (1969:140).

The following sections will deal with the types of structural-semantic relationships that were analysed in the present study as well as the devices of repetition and derivation.

2.2.1.1 Repetition

Repetition is the most frequently used form of lexical cohesion. A problem for the identification of ties of repetition is that it depends on the researcher's definition of a lexical item. (See 3.2 for a discussion of how this problem was resolved in the present study.) Repetition may be of an item, a phrase or a bigger textual structure (Lieber 1979:139). Low-frequency items in a language such as *power* and *leaders* contribute to the cohesion of a text whereas high-frequency items such as *take* and *say* do not carry cohesive power unless something in the text suggests that they should (Gutwinski 1976: 81 as cited in Lieber 1979:139; Stotsky, 1983:433). A second instance of the repetition of a phrase may have slightly different phrasing but still contain the force of repetition. Lieber gives as

examples: "plans of construction" repeated as "construction plans" (1979:140).

Hoey (1991) who is primarily concerned with repetition cohesion, bases his work on insights gained from Hasan's (1984) cohesive harmony, Winter's (1974, 1979) repetition-replacement relations and Phillips' (1985) long-distance organisation (cited in Hoey, 1991:25). Hoey regards lexical relations as context dependent (Hoey, 1991:63; Wessels, 1993:38) unlike Stotsky (1986) who distinguishes between general ties which are systematic to the language they occur in and ties which are instantial in that they are dependent on their context for recognition. Hoey (1991) lists the following types of repetition:

Simple lexical repetition: This occurs when a lexical item is repeated identically with the only changes being grammatical in nature. This is similar to Stotsky's definition of repetition which allows for the repetition of an identical word or word varying in terms of inflectional or comparative ending (Stotsky, 1986:280; see 3.3 below).

Complex lexical repetition: This occurs when two lexical items share a lexical morpheme but are not formally identical. It can also occur when two lexical items are identical in form but have different functions. Stotsky (1986) identifies these types of relations as derivational. Identical lexical items with different functions are related as instances of zero-derivation or derivation by zero morpheme.

Simple paraphrase: This occurs "whenever a lexical item may substitute for another in context without loss or gain in specificity and with no discernible change in meaning" (Hoey, 1991:62). This corresponds to Hasan's (1984) category of synonymy except that Hasan defines synonymy as a general relation whereas Hoey sees it as context dependent (1991:63). This would also correspond to Stotsky's (1986) category of synonymy or near synonymy which includes the potential for both general and instantial relations.

Complex paraphrase: This is said to occur when one item includes the other to which it is related although the two items do not share a lexical morpheme. Hoey restricts the application of the complex paraphrase to three situations which are beyond the scope of the present study.

Superordinate, hyponymic, and co-reference repetition: Hoey includes under the general category of repetition instances of links between general words which follow more specific words on the grounds that general words do not provide more information than specific words. If, however, a more specific word follows a general word, a relation of hyponymic repetition is set up where the superordinate item precedes its hyponym. It is crucial that a common referent is established between the items for a link to be established between them. Hoey (1991) also identifies a relation of co-reference repetition where different lexical items share the same reference. Stotsky (1986) includes relations of hyponymy under her subcategory of inclusion.

It is clear that many of Hoey's (1991) and Stotsky's (1986) categories and subcategories of lexical cohesion overlap although each model emphasises certain features not necessarily highlighted by the other. These approaches to the study of lexical cohesion, therefore, complement one another in their explications of how cohesive devices contribute to textual organisation.

2.2.1.2 Synonymy or near-synonymy

The meaning of the term synonymy may be accounted for as two or more forms associated with one meaning (Lyons, 1968:405) or as items which are related through a fairly close semantic equivalence (Lieber, 1979:141). The above definitions show that there are differing conceptions of what constitutes synonymy and of how close in meaning lexical items should be before they may be termed synonymous. Cruse (1986) proposes two ways of establishing synonymy: first in

terms of “necessary resemblances and permissible differences” (1986:266) between lexical items, and secondly in terms of their contextual relations. He concludes that synonyms are:

“Lexical items whose senses are identical in respect of ‘central’ semantic traits, but differ, if at all, only in respect of ... ‘minor’ or ‘peripheral’ traits”

(Cruse, 1986:267)

In terms of this definition lexical items such as ‘boar’ and ‘sow’ which have a large semantic overlap in that they both refer to the concept ‘pig’, would not be considered synonymous in that they embody the complementary opposition: male / female.

Cruse uses his second means of establishing synonymy, that is, examining lexical items contextually, to point out that there are very few cases of absolute (Cruse, 1986:268; Lyons, 1981:148) or true synonymy (where words are 100% interchangeable) (Carter and McCarthy, 1988:28). According to Cruse, for pairs or groups of lexical items to be considered absolute synonyms, they should be completely synonymous in every context in which they occur. To this Lyons (1981:148) adds that absolute synonyms should have the same distribution and should be ‘completely synonymous’ (that is, they should have the same expressive, descriptive and social meaning) in all contexts. (One should bear in mind, however, that Lyons addresses lexemes rather than lexical items (Cruse, 1986:293).) Determining cases of absolute synonymy contextually proves impossible as the number of contexts in which pairs of synonyms can occur is potentially infinite. On the other hand, it only takes one context in which a pair of lexical items is shown to be non-synonymous to disprove their potential for absolute synonymy. Cruse (1986:270) maintains that, in fact, there is no reason to imagine the need for words with identical meaning in any language, and that if such a pair were to occur, the identical relationship would be unstable and one

item would probably fall away or adopt a different semantic function.

Cruse (1986) sees absolute synonymy as the end point of a scale of synonymy which moves through various degrees of synonymy to non-synonymy. He justifies the existence of such a scale on the basis of the intuition that some pairs of lexical items are more synonymous than others (1986:265). Meaning can be expressive, descriptive or social (Cruse, 1986; Lyons, 1981).

To be classified as synonymous, lexical items need not be identical in word class. Halliday and Hasan (1976), for example, classify *odious* and *dislike* as synonyms in one of their sample texts (1976:353, as cited in Lieber, 1979:141). In the present study, there were many synonymous ties which were easily identifiable as they originated from the same word class. These were ties which linked items such as *leader* and *ruler* or *care* and *welfare*. Ties linking items such as *kill* and *done away with* and *ill-treated* and *cruelty is done on them* were also accepted as synonymous pairs although they were not part of the same word class.

Collinson (1939, in Ullman 1962:141, as cited in Carter and McCarthy 1988:28) developed a set of nine criteria for distinguishing synonyms. These include such principles as: "One term is more general than another: *refuse* - *reject*" and "One term may imply approbation or censure where another is neutral: *thrifty* - *economical*". To this, Carter and McCarthy (1988:29) add the principle that synonyms do not always collocate identically. However, they do allow for the existence of "local" synonymy where words are used interchangeably in specific contexts (cf. Stotsky's, 1983, text specific semantically related items, and Hasan's, 1984, instantial relations). Writers exploit this type of synonymy in their writing. In the present study, synonymous pairs were identified according to their meaning in context. If two lexical items are synonymous in language generally but were not used synonymously by the writer, they were not counted as forming a tie of synonymy (cf. Lieber, 1979:141).

Synonyms are used by writers with the idea of “re-entering” important topic words into their writing at a later stage with the aim of foregrounding them again (McCarthy, 1991: 66). Writers weave synonyms into their texts to focus their readers’ attention on the topic and to provide them with a more inclusive understanding of the topic (see 2.2.1.7 below). One of the aims of the present study was to establish whether better writers were able to develop ideas using this type of lexical cohesive device (cf. Witte and Faigley, 1981).

2.2.1.3 Opposition or contrast

Both Lyons and Leech agree that antonymy (opposition) is a “basic or ‘primitive’ semantic relation(s) between words” (Carter and McCarthy 1988:22). Lyons sees incompatibility as indicating items in sets (1977:288 in Carter and McCarthy 1988:23). He points to four main types of oppositeness (that is, binary contrast - for him contrast does not specify how many words are being contrasted).

These are:

antonymy - which refers to gradable antonyms

complementarity - which refers to ungradable antonyms such as alive / dead

converseness - which refers to a relationship such as parent / child

directionality which points in a direction such as up / down

(Lyons 1977: 279 as cited in Carter and McCarthy 1988: 24).

Cruse (1986) provides greater detail in terms of the classification of types of opposites. Initially these categories were incorporated into the study. However, it became clear after early data analysis that Cruse’s (1986) distinctions were too finely tuned for the types of opposites encountered in the corpus and his distinctions were not used.

Instantial opposites which exist within the context of the text were also considered important in the identification of opposites in the student scripts. McCarthy (1988:23) claims that the antonym of a word is never fixed but always depends on the discourse context. The present study followed Stotsky's classification of lexical cohesion which distinguished between systematic ties and ties which depend on discourse context for their interpretation or instancial ties. Both types of ties were identified in the corpus and counted (see 3.2.2 below).

2.2.1.4 Inclusion

Inclusion is a relationship in which specific terms in the vocabulary are covered by more general terms (Carter and McCarthy, 1988:25), for example, *animal* is the superordinate term for *pigs* and *dogs*. It also includes lexical items related as coordinate members of ordered or unordered sets (Stotsky, 1986:279).

A relationship that falls under the category of inclusion is that of general noun. For example, within the context of *Animal Farm*, the general noun *things* encompasses the items *files*, *reports*, *minutes* and *memoranda*. Halliday and Hasan (1976) include under reiteration the category of general noun. They consider the class of general noun to exist on the border between lexical items and grammatical items. They state that general nouns are superordinate members of major lexical sets which operate anaphorically as a type of synonym (1976:275). Grammatically, they operate with a determiner in a similar way to a reference item. Usually general nouns are marked attitudinally, normally conveying a feeling of familiarity, particularly with regard to human nouns where someone could be referred to as an *idiot* or a *dear*. Carter (1987) cites a discussion by Francis (1985, as cited in Carter, 1987:79) of anaphoric nouns or A-nouns which include items marked attitudinally for a range of items including but not exclusive to human beings. Carter (1987:79) states:

"A-nouns operate as organisational signals, as it were. They serve to label a preceding stretch of discourse, integrate and align it with the ongoing argument and thus represent a position which the writer hopes to have established with the reader."

Carter considers the classification of A-nouns important in relation to the discourse of argumentation which is an important type of student academic writing. A-nouns allow for evaluative elements in text which are very often instantial (p. 81). Although Halliday and Hasan only allow for anaphoric relations of general nouns, this has been disputed by Stotsky (1983; and see 3.2 of this study).

In the present study ties created by the use of general nouns were counted because, although they could be considered weak ties in that they exist on the border between lexical and grammatical items and are often high frequency items, they appeared to have a deliberate text organising function in the scripts analysed.

2.2.1.5 Derivation

Stotsky (1986: 280) follows Marchand's definition of derivation in her analysis of cohesive relations. Marchand (1969:1) distinguishes between "free forms" and "free morphemes"; only "free morphemes" are considered words. He maintains that the word is a two facet sign based on the significant / significate (from the signifiant / signifie relation posited by Saussure). In word formation only meaningful elements can create new formations. "Word-formation is that branch of the science of language which studies the patterns on which a language forms new lexical units, i.e. words" (1969:2). Words formed as grammatical syntagmas include compounding, prefixation, suffixation, derivation by zero morpheme, and back derivation (1969:3). New coining of words is based on a synchronic relationship between morphemes.

Prefixation and suffixation form part of the process of coining new words by derivation. Prefixes are bound morphemes (which Marchand calls determinants) which are preposed to free morphemes (determinatum). Prefixes function as the determinants of the words to which they are prefixed (1969:129). Suffixes are also bound morphemes. The suffix is the determinatum of a syntagma whose determinant is a simple or composite free morpheme such as *father/hood* or *colour-blind/ness*. Marchand does not count formatives such as *-er* or *-est* which express degrees of comparison as suffixes.

Marchand also does not find inflectional morphemes relevant to word formation in that they form an inflected form of the same form, not a new lexical unit. For example, *talk*, *talking*, and *talks* are all inflected versions of the same form. One can speak of derivation only when a word changes its word class or its lexical class. (Jackson writes of a change from one kind of a word to another without changing word class, giving as an example the affix *-hood* which changes a 'concrete' noun to an 'abstract' noun such as *child* to *childhood* (1988:32)).

If a derivative element is not overtly expressed in the derivative, we speak of derivation by a zero morpheme (Marchand, 1969:3). This type of derivation has been called zero-derivation, or conversion (Jackson, 1988:32), being a process that converts a word from one word class to another.

Marchand defines zero-derivation as:

"The use of a word as a determinant in a syntagma whose determinatum is not expressed in phoric form but understood to be present in content thanks to an association with other syntagmas where the element of content has its counterpart on the plane of phoric expression."

(Marchand, 1969:359)

If one function of derivational elements is to change word classes: for example -

ation derives a noun from a verb, e.g. *computation* from *compute* (Jackson, 1988:32), another is to change the meaning of the word, for example: *archbishop* from *bishop* (*arch-* meaning *chief* or *supreme* - Jackson, 1988:32).

In the present study derivation was counted as a type of lexical cohesion device. Lexical items that were linked in terms of the one item being a derivation of the other were counted as derivational ties, for example *democracy / democratically, leader / leadership*. Ties of derivational elements such as *cracy* in *autocracy / democracy* were also counted as derivational ties in that it was felt that this type of tie created an extra dimension of cohesion, often in addition to the tie that might already be established by these items, in this case the tie of opposition. Derivational relations were considered to be potentially important features of cohesion in student academic writing as derivations have been identified as prominent stylistic features of academic writing in general (Stotsky, 1981; see 2.2 above).

2.2.1.6 Collocation

Halliday and Hasan (1976:284) define collocation which they describe as the most "problematical part of lexical cohesion" as "cohesion that is achieved through the association of lexical items that regularly co-occur". Under collocation they include items related as part of ordered or unordered sets as well as those items related in terms of opposition and those items which co-occur frequently in similar texts (Halliday and Hasan, 1976: 285 - 287).

Hoey (1991:7) finds Halliday and Hasan's (1976) definition of collocation problematic in that it includes a variety of lexical relations many of which, he states, have no readily available name. There are other approaches, however, which have tried to systematise the relationships defined as collocational. Reviewing collocation, Carter (1987) suggests that the notion of collocation is

often associated with systemic linguistics as an independent level of language. Words which collocate "are grouped into lexical sets as a series of semantically related options from which a coherent text can be constructed" (1987:50). He maintains that lexical sets can be examined in conjunction with lexical fields which are usually concerned with a specific topic around which clusters vocabulary that relates to that topic. Carstens (1997:324) cites De Stadler (1989:98 - 99) who lists lexical relationships formed by lexical items from particular semantic fields according to whether they form paradigmatic or syntagmatic relationships. De Stadler (1989:98 - 99, as cited in Carstens, 1997:324) maintains that synonyms, opposites, hyponyms and members of both ordered and unordered sets are related paradigmatically and that lexical items relate to each other syntagmatically where words are usually associated with each other, for example, *blond* is usually associated with *hair* (Carstens, 1997:327).

A further difficulty of studying collocation as defined by Halliday and Hasan (1976) is that collocational relationships often have to be identified intuitively (Hoey, 1991:8). In terms of collocational restrictions there is often no reason why certain words should collocate. For example, there is no provable reason why *green* should collocate with *envy* (Carter, 1987:53).

This relates to a further difficulty cited by Hasan (1984:194): that of "inter-subjective reliability". Words used cohesively in academic discourse are often abstract and complex and "the cohesive potential of the content words in academic discourse depends heavily upon the reader's previous experiences with them." (Stotsky, 1983:438 & 439). Because of this element of subjectivity, collocation was not examined in the present study. Many of the relationships identified by De Stadler (as cited in Carstens, 1997) above have been included by Stotsky (1983) as subcategories of semantically related words. For Stotsky (1983; see 2.2 above) items falling under this category are systematically related

to each other in that they have a stable relationship in the language.

2.2.1.7 Cohesion in first and last paragraphs

The structure of a written text is determined by the context in which it is written. Although writers of academic texts do not have the same contextual support they would have in spoken situations, they are constrained in their writing by the communicative context in which they have to produce the text. In the case of the present study, the writing was produced in the context of an examination situation in response to an instruction given in the form of an examination question.

The question demanded of the writers that they produce an expository essay about a narrative set work they had studied. Eiler (1983:170) writes that the register "*writing about literature* has its own generic structure, in this case a narrative typology involving participants, objects, and events". Writers must integrate the narrative text into their expository writing and to do so must adopt various roles including those of reader of the literary text, writer and critic of the literary text, writer for a particular audience and writer of an expository text which follows certain traditional structures such as the provision of introduction, body and conclusion.

It is the structure and conventions of the expository text of this nature that the writer must keep in mind when writing. de Beaugrande and Dressler (1981:182) claim that these are embodied in the *text type*. Certain traditional text types can be defined according to their functions. Narrative texts are used to order actions and events sequentially (1981:184), whereas argumentative texts are used to evaluate or advocate certain beliefs or ideas (1981:184). The fact that both the writer and reader of the text depend on their knowledge of another text (in this case a particular text type) to interact with the text is part of the study of intertextuality (1981:182). The choices the writer makes about structuring the

text depends on their knowledge of the text type they are producing.

Another key consideration for the writer is that of the audience for whom they write. Writers write specifically for the audience they perceive will read their texts and make decisions regarding the structuring of their texts based on the degree to which they believe they share information about the text world with their reader. In the context of student academic writing the audience "is non-intrusive, nonpersonal, silent, and anonymous and ... shares knowledge of literature as verbal art" (Eiler 1983:171).

Brandt maintains that "successful writers plan their text around unwritten conditions for meaning that direct lexical and syntactic choices having mutual significance for their readers" (1986:95) and suggests that there are three important ways in which they do this, namely the use of exophoric reference, cohesive relations (cf. Eiler, 1983:173), and thematic structure.

Exophoric referents refer the reader to something outside the text either to the "real" context (Brandt, 1986:95) where the writer refers to "I" or "you", or to the context of the text world, in this case the world of the work of literature: *Animal Farm*. The reader accesses the exophoric referents by mediating their beliefs and goals with regard to the text world into the communicative situation (de Beaugrande and Dressler, 1981:182). In the case of the present study little mediation was required as the examination question provided a background passage from *Animal Farm* to which the writer and reader could refer.

The writer uses cohesive ties to present an explicit context for the reader. Writers choose which cohesive ties to use based on how much knowledge they assume the readers to have of what they are writing, and how they can make their writing more understandable. Writers use repetition, for example, to present the topic or theme of the writing. Repetition limits the interpretation of a concept.

The use of synonyms, on the other hand, allows the reader a more inclusive interpretation of the concept and a richer understanding of the topic of the text.

The theme of a clause is usually written in the subject position. The theme announces the topic of the text, the focus of the text at that point. The rheme is the rest of the clause and provides information about the theme. Clause or sentence themes relate to larger topics or macrothemes (Brandt, 1986:98) but of themselves are a good indication of what the writer chooses to focus on at each point in the text. Theme and rheme may form patterns which can be traced through a text, with the rheme of a particular sentence becoming the theme of the next sentence (see 1.2.2 - studies on given and new information or topic and comment). Themes may originate in the text world rather than in the text itself and may demand of the reader an application of their knowledge of the context of the communicative situation.

The present study focussed on the use of cohesive references within the context of answering a particular examination question. It was anticipated that the use of lexical items in the introductory and concluding statements of the essays would be influenced by the question to which the essays were a response. Brandt analysed three passages by the same student writer with the aim of examining the use of exophoric and cohesive devices and thematic patterns as signals of the semantic constraints provided by the contexts in which the essays were written (Brandt, 1986:99). She found that when responding to a written instruction from the teacher, the opening paragraph of the student's essay revealed the "highly specialised and semantically constrained conditions under which the student composed" (1986:101). She found that only a small range of linguistic features were chosen for the introduction and that these related to the teacher's instruction. She found that the choice of what lexical item the writer used depended on the audience they were writing for. When writing for the teacher, the terms used could be found in the lexicon of the teacher - the writer

seemed to be attempting to align his meaning with the teacher's meaning. Where the audience was the teacher, the writer could rely on a high degree of shared knowledge and thus used repetition and exophoric items extensively. When the writer was given a task that involved writing for a more general audience, establishing a shared understanding with the audience forced him to use a wide variety of lexical items in the form of synonyms and to place a greater emphasis on cohesion than on exophora.

Brandt concludes that:

"How a writer chooses to refer and represent seems particularly constrained by such factors as the nature of previous discourse present in the context, the level of shared reader - writer knowledge which preexists a text, and the discourse role a writer plays or is expected to play in the text."

(Brand, 1986:104)

The present study examined the writers' use of cohesive ties in the first and last paragraphs of their texts in relation to the question which they answered to establish whether any meaningful patterns could be discerned from this part of the context in which they wrote.

2.2.2 Quantitative studies of lexical cohesion

There have been many empirical studies conducted to establish the effect of cohesion on writing quality or coherence in student writing. These studies have often had different specific aims and areas of focus. Most of the studies listed here have aimed to provide insight into the writing quality of student composition writing. There is, however, a growing body of research dealing with the relationship between cohesion and coherence in student academic writing (cf. Hubbard, 1989; Wessels, 1993). A small body of literature correlates coherence

with writing quality (McCulley, 1985; Hubbard, 1989; see 1.1.1) which means that it is expected that literature which assesses the influence of cohesion on writing quality is relevant to this study. The following section details the quantitative literature considered to be the most relevant to the present study.

Eiler (1983) conducted a study to determine the effect of the type and frequency of cohesive relations on the development of writing skills in the genre of writing about narrative fiction. Working on the assumption that text is “ a continuous process of semantic choice” (Halliday, 1979:195 as cited in Eiler, 1983:222), Eiler maintains that a practical way to answer such questions as “what choices, what ongoing selections, constitute a ‘successful’ production of text of literary analysis, and how can we tap these choices in an explicit linguistic analysis of text” (Eiler, 1983:173) is to consider the cohesive relations in text.

Expository essays written every three months over one academic year (Tasks I, II, III) by fifteen ninth grade high-achieving students were analysed by Eiler. This represented 45 texts dealing with narrative fiction. The Tasks III of the students were scored by four English teachers and ranked low, middle, and high after which they were analysed linguistically. The scale used to score the essays included a section used to record literary response under which Eiler assessed the effect of lexical cohesive ties on the development of writing. Apart from lexical cohesion, Eiler also looked at relations of reference and conjunction. The cohesive relations were then examined in the context of:

register theory

cohesive relations sub-categorized in association with:

the field

the tenor

the mode of the discourse

A contextual configuration (Hasan 1978, as cited in Eiler, 1983:170), which addresses the genre of the text as well as the roles assumed by the writer discussing a literary text, was then developed.

Initially the groups of essays rated high, middle and low were studied for lexical cohesion in total without differentiating between the different types. Eiler found that lexical cohesion accounted for in most cases 75% of all the cohesive ties in the expository texts that she studied and that the number of lexical cohesive ties outnumbered any other type of cohesive relation in the texts of expository writing about narrative literary writing. She also found that, over time, the manifestation of lexical cohesion increased for the low, middle and high groups.

Eiler then subcategorised the lexical ties in order to establish what constituted specific evidence of literary response. She specified lexical types to include:

- (a) systematic relations
- (b) fieldbound relations including lexical items from the literary text being discussed (often proper nouns) and lexical items involved in the analysis of the literary work, and
- (c) instancial relations, that is, those representing the writers' critical interpretation of the literary work.

She found that in terms of developmental trends, measurements of lexical field-bound and instancial relations revealed that the low and middle groups seemed to improve the most in terms of writing about literature analytically and interpretatively. However, Eiler was unable to prove that the undifferentiated frequencies of lexical relations had an effect on the high marks awarded by the raters and concluded that there is need for specifying lexical semantic relations at higher levels of delicacy (Eiler, 1983:219). Although Eiler's (1983) research had a quantitative component to it in that she correlated frequencies of lexical

relations with teacher ratings, she suggests that the figures given are meant only to be taken as suggestive of "semantic movement and processes of text production" (Eiler, 1983:221). She states that her aim was to provide a procedure for describing qualitatively cohesive strategies for text-formation and suggests that quantitative research be undertaken to determine the frequencies of the different types of cohesion in the context of the other types.

Eiler's (1983) research suggests that lexical ties show an increase as writers develop their ability to produce acceptable academic texts. It is hoped that the present study will go some way towards assessing whether quantitative counts of types of lexical density have an impact on the scores awarded by raters.

Witte and Faigley (1981) in an effort to explore what internal characteristics of student essays distinguish those ranked high from those ranked low in terms of writing quality, applied Halliday and Hasan's (1976) cohesion taxonomy to five high-rated and five low-rated essays. They discovered that high-rated essays contained a higher density of cohesion than did the lower-rated ones and that two thirds of the cohesive ties used in all the essays were lexical. They found that lexical collocation ties best distinguish writing ability. Witte and Faigley (1981:199) conclude that writing quality incorporates factors other than cohesion density. These factors which are also conditions of coherence include the writer's purpose, the reader's knowledge and expectations and the information to be conveyed (1981:202).

Tierney and Mosenthal (1983) studied the relationship between Halliday and Hasan's (1976) cohesive devices and the coherence ratings awarded the essays of twelfth grade students from two advanced rhetoric classes writing on two different topic types. They discovered that no causal relationship could be statistically established between cohesive ties and coherence ratings. They did discover, however, that reference and lexical cohesive ties played a part in

distinguishing between text topics and that it appears as if topic affects the types of cohesive ties manifest in a text (1983:221). Despite this, they conclude that the ties which are most pervasive in a text, i.e. reference, conjunction and lexical cohesion have the effect of diminishing the usefulness of cohesion as a measure of coherence. Rather than statistically assessing the affect of cohesive ties on coherence, they suggest that the function of cohesive ties in a text should be seen as the result of the coherence of the argument of the text. Coherence, they believe, is a product of the structure of the text, which is a function of the argument of the text, and the relationships between propositions indicated in different ways, of which cohesion is only one (1983:227).

McCulley (1985) conducted a study to explore the relationships among certain of Halliday and Hasan's (1976) cohesive devices and primary-trait assessments of writing quality and coherence to establish whether specific textual features of writing affect judgements of coherence and quality of writing. To do this he analysed 493 persuasive essays written by 17 year-olds for an authoritative audience. He found that the lexical cohesion relations of synonym, hyponym and collocation contributed most to the judgements of coherence and writing quality (1985:278), but concluded that textual features do not entirely account for coherence judgements. Coherence consists of more than relationships of cohesion.

Yde and Spoelders (1985) carried out an investigation to determine, amongst other things, the degree of cohesiveness shown in texts of beginning and advanced writers and the kind of cohesive devices used by beginning and advanced writers. Focussing on narrative writing, they assessed the discourse production of Dutch-speaking students who were enrolled at elementary school. The first sample they selected consisted of 14 children aged between 8 and 9 years old, the second sample of 14 children aged between 10 and 11 years old where they anticipated that the children would be more aware of the difference

between speech and writing (1985:409). They used Halliday and Hasan's (1976) classification of cohesive devices, listing under lexical cohesion both reiteration and collocation. Their findings showed that although lexical cohesion was the predominant type of cohesive device used in the texts of both the younger and older children, the majority of ties were of repetition. Synonyms were rarely used. They nevertheless conclude that cohesive skills are important skills in writing and that they can be taught to improve writing. Despite this, they discovered that the older children used cohesive ties more effectively to produce more compact writing.

As part of a study to explore what constitutes growth in writing about ideas and how this growth affects writing quality, Stotsky (1986) analysed the semantic relations in the examination essays of 12 Grade Ten high school students. Using the lexical subcategories she adapted from Halliday and Hasan's (1976) cohesive classification, she analysed six high-rated and six low-rated essays for cohesive ties. She found that writers of high-rated texts created semantic units with a variety of cohesive relations which operated throughout the text to link sentences in all parts of the text together, thereby creating a rich texture (1986:287). In contrast to this, the low-rated writers not only created a sparse network of lexical ties most of which were repetition, but also exhibited a limited vocabulary and a tendency to use their vocabulary in a limiting way.

In terms of the development of ideas within their essays, she discovered that writers of high-rated essays had clear introductions which referred to the topic. The writer took up a position in relation to the topic and developed that position throughout the essay using examples to illustrate their arguments. These writers also produced clear concluding statements. On the other hand, the writers of low-rated papers did not have explicit introductory or concluding statements.

Stotsky concluded from the study that a writer's growth in the ability to produce

expository writing can be measured in the writer's increasing capacity to use semantic networks to link semantic units and thus present and develop ideas. She maintains that this may explain why a strong and varied vocabulary indicates growth or writing quality (1986:289).

To test whether dynamic ties are more related to effective writing than are static ties (see 2.2 above for an explanation of these two terms), and to establish whether dynamic ties are more difficult for writers, Hartnett (1986) analysed 316 expository and persuasive essays written by students in a developmental writing course at a community college in Texas. In her analysis Hartnett used 17 subtypes of cohesive devices derived from Halliday and Hasan's (1976) classification.

She found that students did have more difficulty constructing dynamic ties than static ties and that analysts, therefore, had difficulty in recognising them. However, the presence of static and dynamic ties did not seem to ensure success in student writing although badly constructed ties did seem to have an effect on the quality of the writing which suggests that better writers use ties accurately to organise their texts. Hartnett explains that Halliday and Hasan do not consider the analysis of cohesion as a means to interpret or evaluate text (1976:328 as cited in Hartnett, 1986:151). She concludes that while cohesive "features and their distinct uses can help us to describe how readers understand and writers control the textual structures that express rhetorical development in written discourse" (1986:152), there are many features other than cohesive devices which contribute to successful writing.

Neuner (1987), who cites the studies of many previous researchers in his work, expanded on and developed their studies (Neuner 1987:94 - 97 as cited in Wessels 1993:29). He examined 20 good and 20 poor freshman college essays in order to compare the use of cohesive ties and chains between the two groups.

He used 18 types of cohesive devices derived from Halliday and Hasan's (1976) scheme but found that no specific type of cohesive tie was used more frequently by the writers of the good essays and that cohesive ties did not distinguish skilled and less skilled writing. He did find that cohesive chains, "a series of lexical collocations, reiterations, synonyms, and superordinates and their reference pronouns all semantically related to one another" (1987:96), have several features that affect the texture of a piece of writing. He suggests that "longer chains, greater variety of words, and greater maturity of word choice" (Neuner 1987:101) are signs of good writing, and this has been interpreted by Wessels to suggest that "skilled writers employ lexical cohesive devices more effectively than do less skilled writers" (Wessels 1993:29).

In her own study, Wessels (1993:9) investigated the effect of certain aspects of lexical cohesion such as bonding (Hoey 1991), keywords and lexical collocation on coherence in student academic writing. Taking as an assumption that coherence is one of the major problems in student writing (1993:5), her study aimed to contribute to an explication of coherence and to use this knowledge to teach students to write coherently.

She studied 40 examination history essays written by second year university students which were rated impressionistically for overall coherence by two raters using Bamberg's Holistic Coherence Scale (1983; 1984). The essays were then divided into a group containing the 13 highest rating texts and a group containing the 13 lowest rating texts. The lexical devices in the texts were correlated with the coherence ratings.

Wessels discovered that the more coherent texts displayed a greater amount of bonding, and a greater density of bonds per text and bonds per sentence (1993:83). While the presence of keywords in a text could not be related to coherence, she found a relation between repetition of keywords in the text and

coherence, as more coherent texts contained more keyword links (1993:88). She also found that the more coherent texts had a greater variety of lexical items and more keywords that formed links and bonds in the text (1993:93).

She was able to conclude along with such other researchers as Witte and Faigley (1981), Stotsky (1986) and Neuner (1987) that "skilled writers thus seem to have a far more extensive vocabulary" (1993:93) than unskilled writers do. In fact, Wessels' research suggests that this difference between skilled and unskilled writers is even more significant than has been suggested by previous researchers. Skilled writers use repetition more often and in better ways than do unskilled writers and have a far greater variety of lexical items in their texts (1993:98).

At the other end of the scale, Patricio (1993) found a correlation between low coherence ratings and ineffective use of cohesion in student writing. Using as her sample 20 examination essays of Standard Ten second language pupils in a rural black school, Patricio calculated the number of cohesive devices in each piece of writing using Halliday and Hasan's (1976) classification of cohesive elements. She then had four experienced teachers rate the essays in terms of their holistic coherence.

She found that the correlation between the cohesion and coherence of the essays was close (0,73), although the percentage of cohesive devices was higher in each essay than the percentage awarded for coherence (1993:67) which showed that the cohesive devices were often incorrectly or unnecessarily used (1993:69). She concluded from this that the writing was at a low level (the writers also showed undeveloped language usage, composing skills and analytic abilities (1993:68) where cohesion and coherence would be expected to correlate because they would affect each other. Her study revealed that lexical cohesion was the second highest device used, (pronominal reference being the highest)

but that writers had a limited scope of vocabulary and a narrow range of lexical cohesive devices, a sign of serious problems in student writing.

Hubbard (1993) in a study designed to begin to provide an explication of the concept of coherence, analysed the effect of the densities of the categories of reference and conjunctive cohesion on the coherence of student academic texts. The texts, examination answers written by university students in response to two questions each representative of a different discipline, were assessed holistically for coherence by three raters. Hubbard (1993) used Halliday and Hasan's (1976) categories of reference and conjunctive cohesion but included new subcategories under each category. He found that the relationship of indirect reference had a strongly significant positive relation to coherence. Indirect reference can be defined as occurring when a presupposing reference expression is co-referential with but is not equivalent to the presupposed item (Hubbard, 1993:61). Indirect reference is of the same type of relationship advocated by Hoey (1991) under his category of superordinate, hyponymic and co-reference repetition (see 2.2.1.1) and obviously contains a lexical cohesive element. Hubbard (1993:73) concludes that semantically defined subcategories are more likely to correlate with coherence than are formally defined subcategories.

As Hubbard (1993:59) has stated, "differences in aims, research designs, subjects and the classification of items" makes the direct comparison of studies difficult. It seems, however, that there are some common findings in the above literature that are relevant to the present study. It appears firstly that lexical cohesion features prominently in student writing but that it is the more effective use of lexical cohesion, including the use of semantically related words such as synonyms and hyponyms that distinguish scripts rated higher from those rated lower in terms of writing quality and coherence measures. Lexical cohesion is also an important indicator of the writer's development in the ability to produce acceptable texts (cf. Eiler, 1983; Stotsky, 1986) and that the use of cohesive ties

contributes to compact writing which as Stotsky (1981) has stated enables more efficient text processing on the part of the reader. The present study, then, continues the tradition of analysing cohesion in student writing with the intention of contributing to this field in that it applies a particular classification model to the analysis of lexical cohesion (including a specific focus on first and final paragraph cohesion) and investigates how these features influence the perception of coherence in the writing.

The following section examines another aspect of the present study, namely the effect of lexical errors on student academic writing.

2.3 ERROR ANALYSIS

The present study looked at lexical errors with a view to establishing whether they influenced the coherence of the writing in which they occurred. Corder (1974, as cited in Ellis 1994:48) has identified the following steps in error analysis research:

1. Collection of a sample of learner language
2. Identification of errors
3. Description of errors
4. Explanation of errors
5. Evaluation of errors

This study is focussed on the identification, description and evaluation of errors. Step 1 is recognised as a preliminary step to carrying out any empirical research, not specific to error analysis (Lennon, 1991:181). The sample of learner language contained in this study constituted student examination scripts analysed for lexical cohesion. Step 4, the problem of identifying the source of learners' errors, is difficult (Sonaiya, 1988:28) and is outside the realm of this study which

is concerned to see what type of lexical difficulties learners experience in writing.

The identification of what constitutes an error has been the subject of a certain amount of debate. One of the controversies within this debate is that of determining what is a mistake (defined as occurring when learners fail to perform their competence) and what is an error (a deviation caused by lack of knowledge on the part of the learner) and therefore worth studying (Corder, as cited in Ellis, 1994:51). To determine when a writer has made an error and when a mistake, it has been suggested that the writer be interviewed to establish intended meaning (Corder, as cited in Ellis, 1994:52 - 53). Bartholomae (1994) advocates in addition to this a procedure which encourages the learner to read out loud while the reader / teacher notes the differences between the expected response and the actual observed response. Bartholomae (1994) notes that in his own research he has found that learners while reading often substitute the correct forms for the written incorrect forms, although they are not necessarily aware that they have done this. He concludes that most error difficulties are rooted in learners' performance and not in their competence. However, Bartholomae's subjects are *basic writers*, a term normally associated with first language speakers (cf. Kroll, 1990), whom he is concerned to teach the "second language [of] ... formal, written discourse" (Bartholomae, 1994:344). For second language speakers errors may be a reflection of competence rather than performance.

Other research has established that some types of errors are a reflection of the learner's competence (Sonaiya finds that lexical errors are usually the result of a lack of knowledge about contextual appropriacy - 1988:53). The learner's competence, however, is actually far from stable, particularly with regard to the use of vocabulary. The learner's knowledge of certain lexical items may be limited to specific contexts. Procedures such as those advocated by Bartholomae, rooted in the specific context applied by the learner, are not likely to prove useful in determining the extent of the learner's knowledge of certain lexical

items.

While the types of procedures advocated by researchers such as Corder and Bartholomae (cf. also Krashen, 1977, 1981, as cited in Sonaiya, 1988:51 - 52) yield insights into identifying errors as opposed to mistakes in learners' writing, researchers / teachers are in practice rarely able to carry them out. They must rely on their own interpretation of the text as written product (often produced under examination conditions), to ascertain the learner's errors. For the purpose of this study, therefore, no distinction was made between errors and mistakes.

A further issue in the identification of errors that needs to be reviewed, is the distinction made between errors of first language speakers and those made by second language speakers. Laufer (1991) distinguishes between lexical disruptions made by native speakers of a language and lexical confusions made by second language speakers. She suggests that lexical disruptions, such as the "slip of the tongue" phenomenon, are "involuntary deviations from the speaker's current intention" (Laufer, 1991:318) which are accidental, can be corrected by the speaker and are therefore performance errors. By contrast, lexical confusions made by second language speakers are errors of competence. This might be taken to suggest that the study of errors in student writing should be limited to the study of errors made by second language speakers. However, Laufer (1990) does acknowledge that the acquisition of lexis is a gradual process and that both native speakers and second language learners are likely to show lexical confusion while acquiring new words. The subjects in this study included both native speakers as well as second language speakers of English. As it was assumed that both of these groups are learners being initiated into the realm of academic discourse, no distinction was made between first and second language errors. All lexical faults (Hubbard, 1989:157) were taken to be examples of error, it being not practicable to distinguish between error and mistake.

2.3.1 Description of errors

This study focussed on lexical errors in student academic writing and in the light of this, it is interesting to look at some of the procedures that have been established for describing lexical errors in writing.

In an effort to establish a “viable” test of writing, Martin Nystrand (1982) undertook a pilot study with the Trent Valley Centre of the Ontario Institute for Studies in Education in 1975. Prior to undertaking the study, members of the project assumed that writing assessment involved the examination of texts for their strengths and weaknesses. However, acknowledging that establishing the worth of a text results as much from reading as from the text itself, Nystrand began to explore the idea that “many of the salient features of written communication lie ... in the interaction between writer and reader by way of the text” (Nystrand, 1982:70). Thus the focus of the pilot study shifted from the study of text to the process of interaction between writer and reader. Working from the premise that communication requires writers to establish predictions (1982:64) their readers will probably make, and noting that the type of audience puts particular constraints on writers, he sought to describe and classify the variances between the writers’ texts and the guesses of readers, in order to note the ways in which written communication can fail.

He developed an initial taxonomy of areas of constraint on writing:

1. Graphic constraints
2. Syntactic constraints
3. Semantic constraints (which deal with the presuppositions brought to the text by the reader)
4. Textual constraints (which deal largely with cohesion)
5. Contextual constraints (which deal with the contextual situation of a

text)

He called the variances between what writers say and what readers expect from them, misconstraints (as opposed to error which assumes that the variance is bad which is not always the case).

The study hypothesised three types of variance:

1. Simple misconstraint
2. Impaction: where there is a density of details. The reader needs the main idea to sort out all the details.
3. Rarefaction: where the reader needs more detail to sort out the main idea.

Nystrand (1982) applies the above model at the various constraint levels, most interestingly at the level of the lexicon, problems of internal cohesive relations in the text and problems within the situational context of the text. He characterises writing as having three levels of functional relations. At the graphic level, he posits relations of legibility, at the syntactic and lexical levels, relations of readability, and at the textual and contextual level, relations of lucidity (1982:72). Thus, he suggests, in order to be lucid, the writer should accurately employ textual and contextual relations that are meaningful to the reader.

The assumption of the above model of error description is that the writer interacts with the reader and that errors arise from writers making the wrong assumptions about their readers. As Flower(1979) says, much writing done by students entering tertiary education is self-centred, with very little focus on the audience. What is interesting about this model is that it examines discrepancies or errors that occur above the sentence level.

Another approach to error description is the lexical approach advocated by Sonaiya (1988). She claims that not much has been done with regard to the study of error in the lexicon, despite a general acknowledgment made by teachers and learners alike, of the importance of vocabulary as part of language learning. The main focus of linguistic research in the past has been grammar. The lexicon has been regarded as the "repository of all those idiosyncratic facts about language that cannot be handled by the grammar" (Sonaiya, 1988:5). This focus on grammar has failed to realise that the lexicon is a highly organised system. Summarising a quotation from Bailmer and Brennenstuhl (1981, as cited in Sonaiya, 1988:74), she maintains that:

"...to claim knowledge of a word involves knowing more than the usual meaning ascribed to it; it means knowing its relationships to other words (cf. the notions of semantic networks and lexical fields .. Lyons 1977, Lehrer 1983), the implications and presuppositions that the word usually carries, the kinds of contexts in which it is usually employed, etc."

Sonaiya defines errors in lexis in terms of lack of knowledge. Lexical errors are made because writers do not have the knowledge of how words they should have used are different to the words that they have used.

In developing her approach to error analysis, Sonaiya collected errors from written assignments of American and Nigerian students studying French at Cornell University and the University of Ife respectively. Sonaiya examines each error and the corresponding correct target word in terms of their shared semantic space to see how the items are related and how they differ. She posits the theory that where learners lack knowledge, they rely on what they already know. (This is borne out by Rwigema's study (1996), see 2.3.2 below.) The errors are also examined in terms of their distance from the target word. A lexical approach to the examination of errors is a useful one in that it looks at the relations between lexical items. The appropriate use of synonymous, hyponymous, and antonymous lexical items contributes to lexical variation which has been found to

be a reliable measure of lexical richness in student writing (Engber, 1995). This in turn has been posited as necessary for the development of invention strategies students need to generate topics for writing (Spack, 1984; cited in Engber, 1995:140).

Laufer (1990; 1991) posits a type of lexical error made by second language speakers, that of confusions of words that are similar in form. She refers to those pairs or groups of words that can be confused with each other as "synforms" or "similar lexical forms"(1991:321).

Laufer lists 10 categories of synforms. Each synform category has aspects of similarity between the target word and the error.

For example: Category 3 contains synforms that differ in that a suffix is present in one synform but not in the other - historic / historical. (Laufer, 1991:321)

Laufer designed two tests to show whether learners commonly confused synforms. She also tried to establish whether there were salient features of words which learners retained even when they confused the target lexical item with another form. She tested 321 adult foreign learners of English at a level equivalent to the Cambridge First Certificate of English and 207 English-speaking 12 year-old children from Edinburgh. She acted on the assumption that the children, while able to cope with comprehension and production of a relatively simple written text, were still learning their native language.

She discovered that the salient features of the lexical items for both first and second language speakers were: the grammatical category of items, the stress pattern of the items, their initial elements and their consonants. She found that these features were correctly produced in cases where the fact that the item had

been stored wrongly or insecurely in the learner's lexicon, resulted in the item being retrieved incorrectly by the learner.

Initially the scripts in the present study were tested for confusions in synforms, but it was discovered that very few of the learners made errors that could be ascribed to this type of confusion, the learners being more inclined to produce inappropriate but correctly formed lexical items, or incorrectly formed lexical items. This is nevertheless an interesting and useful procedure for identifying this particular type of lexical error.

Engber (1995) has developed a model for looking at lexical errors in student essays which focuses on the role of the writer of the text and thus falls within the framework of discourse / text-analysis as discussed under 2.1 above. Engber (1995) is concerned to determine the effect of lexical error on the quality of student composition writing. Using as the criterion for identifying a lexical error the premise that "knowing a word (lexical item) means knowing both its meaning and its form" (1995:146), she groups lexical errors under the broad headings: lexical choice and lexical form (see 3.2.3 below). Under lexical choice, she distinguishes between individual lexical items, and combinations of lexical items which include: two lexical items, phrases, and multiple errors involving core lexical items.

Under the category lexical form, she groups derivational errors, errors of verb form, words which are phonetically similar, but semantically unrelated to the target word, and words distorted by major spelling errors. (For the study of lexical errors using Engber's error categories, see 2.3.2 below.) Engber's classification of errors was used in the present study.

2.3.2 Quantitative studies of error analysis

There are two types of study of error analysis that are of relevance to the present study. The first is the importance of lexical errors within the study of language errors and the second is the effect of lexical errors on the perceived quality of the writing in which they occur.

With regard to the place of lexical errors within the study of learners' errors, Sonaiya (1988) reports on a study undertaken by Meara (1984, as cited in Sonaiya, 1988:76) in which an analysis of learners' errors was carried out at Utrecht University. The study found that the incidence of lexical errors was greater by three or four times than that of grammatical errors (1988:76).

Rwigema (1996) in a study to establish what kind of errors interfere with the communicative strategies of ESL learners, analysed end-of-year examination scripts of students doing the Diploma of Education at University of Transkei over a period of five years, from 1990 - 1995. She found that word choice errors formed the highest percentage of errors: 32% of the total error count. Of these 40% was the use of the wrong lexical item. She concluded that the high incidence of lexical error was the result of a poor vocabulary on the part of the students. The errors showed evidence of students trying to create new words by (amongst other devices) applying derivational rules, although, as a result of a poor lexico-functional grammar, doing so inappropriately.

Along similar lines, Roets (1991) carried out an error analysis of entrance test scripts to Rand Afrikaans University written by second language speakers of Afrikaans. This was to determine their level of proficiency in Afrikaans as a means of establishing whether they needed language support in the form of a course in Afrikaans for Academic purposes.

In 1989, 112 scripts were analysed and in 1990, 312 scripts were analysed. The tests, which took the form of cloze exercises, evaluated the learners' performance with regard to preposition usage, vocabulary and nexus. The results of the 1989 test revealed that most errors related to vocabulary. In the 1990 test, most of the errors that occurred were in preposition usage but were generally semantic in nature. Thus the analysis of both tests revealed that errors were largely caused by a limited vocabulary on the part of the subjects.

Sonaiya (1988) also claims that lexical errors are judged as being more serious than grammatical errors by native speakers. In support of her claim, she cites studies by Johanssen (1978) and Politzer (1978, as cited in Sonaiya, 1988:76). In an experiment where native speakers were asked to evaluate 60 pairs of German sentences containing error, Politzer noted that the subjects rated errors in terms of gravity in the following order:

Type of error % judged worse than another type

Vocabulary	77
Verb morphology	55
Word order	54
Gender confusion	51
Phonology	36
Case ending	26

[Table 4.1 Error Gravity, Politzer 1978, in Sonaiya 1988: 77]

The above findings are supported by the results of an investigation conducted by Santos (1988, as cited in Engber, 1995:140) in which the reactions of professors to writing errors of second language speaking students, were examined. Santos found that readers were able to evaluate content and language separately,

except in the case of lexical errors. She found that the use of the wrong lexical item, although a language error, obscures the meaning of content. This negatively affects reader judgement and the readers in her study rated lexical errors as the most serious of language errors.

In contrast to the above studies, Lieber (1979:203) reported few lexical cohesive errors in the tertiary level, second language students' essays she studied. Errors only consisted of 2% - 6% of the total number of lexical devices used.

Superordinate and general items which contribute least to lexical cohesion, were the major problems in the students' writing. This suggests that the students were able to choose words appropriate to their essay topics and portray the sense relations of lexical items fairly accurately; an important aspect of knowing a word.

With regard to the second type of error analysis mentioned above, that is the effect of lexical errors on the perceived quality of writing, quantitative studies show varied results. Kroll (1990) studied whether there is a connection between the level of syntactic accuracy (under which she includes word choice and other lexical errors) and the overall success of the student essay. She examined 100 compositions written by 25 advanced English second language students at freshman university level. She carried out two analyses: one at sentence level where she identified every syntactic, morphological, and lexical error and one at a global level where readers were asked to rate the essays according to a holistic scale as if the essays were free of errors. The syntactic analysis revealed that lexical errors constituted 32% of the total number of errors in the corpus, a far higher number than any other syntactic category. Kroll then correlated the syntactic analysis with the global analysis. Her findings were that the two scores did not correlate statistically. Essays with low holistic scores had accuracy ratios (the ratio of the number of errors to the total number of words in the essay) ranging from very low to very high. She concluded that there was not necessarily a relationship between the syntactic accuracy including lexical accuracy and the

rhetorical competency of the essays in her sample, and that the level of proficiency in one area could not be used to predict the level of proficiency in another.

Engber (1995) examined whether lexical errors influenced the quality score given to student compositions. She had 66 timed, student placement essays assessed holistically by 10 teachers. She then examined the essays for lexical errors and correlated the results. Her results showed that the readers were negatively affected by lexical errors only moderately. Engber suggests that this is the result of the intermediate to advanced level of proficiency of the student writers. At this level the students did not make large numbers of errors. Engber predicts that readers are tolerant of a small number of errors. Essays that are full of errors are likely to receive a lower rating. Interestingly, she cites a study by Linnerud (1986) who compared compositions written by 17 year-old Swedish students who had studied English for nine years with compositions produced by students who were native speakers of English. These students were also advanced writers of English. Linnerud found that the correlation between lexical error and the quality score was low and non-significant. Engber suggests that this was the result not only of a smaller number of errors, but also that the kinds of errors were easily interpretable by the readers within the broader context. At the intermediate level Engber notes, some types of errors made by the writers would have been difficult or uninterpretable within the context.

2.4 CONCLUSION

This chapter focussed on the concepts of coherence, lexical cohesion and error in the applied linguistic context of student writing. Descriptive models classifying types of cohesion and errors were presented and various studies which centred on the application of these models to student writing were summarised. It appears from the findings of these studies that while both lexical cohesion and

error types in some instances seem to influence the perception of coherence or the awarding of ratings to student essays, it is not possible to say with certainty that this is generally the case. There have been studies which report negligible or no correlation between cohesion or error and coherence or writing quality. It is for this reason that the hypotheses upon which the analytical design outlined in the next chapter is based, were drawn up as non-directional.

CHAPTER 3

ANALYTICAL FRAMEWORK AND RESEARCH PROCEDURES

3.0 INTRODUCTION

This Chapter outlines the analytical framework used to assess the data in the study as well as the research procedures adopted.

3.1 RESEARCH DESIGN

In terms of research design the present study can be classified as operating within a quantitative paradigm. It has a quasi-experimental design in which quantitative data are chosen randomly and analysed statistically for the occurrence of certain features which have been defined operationally (Nunan, 1992:5 - 6).

3.2 ANALYTICAL FRAMEWORK

3.2.1 Coherence

The study defined coherence operationally by using Bamberg's four point holistic coherence scale (Bamberg, 1984) against which to measure the scripts under analysis. Myers (1980:1) claims that one of the most productive ways of assessing pieces of writing is to score them holistically using raters' impressions to grade the essays on a numerical scale. He suggests that it is better to list characteristics of good writing such as clarity and coherence rather than to use counts such as those of T-units, cohesive ties or error counts where an assumption is made that the total number of errors reflects an essay's worth. He also finds that inter-rater reliability and intra-rater consistency is high within a holistic marking framework, citing Charles Cooper as saying:

"... When raters are from similar backgrounds and when they are training with a holistic scoring-guide - either one they can borrow or devise for themselves

on the spot – they can achieve nearly perfect agreement in choosing the better pair of essays ... ”

(1977:19, as cited in Myers, 1980:3)

Bamberg developed her holistic coherence scale to address the need for assessing coherence based on current linguistic theory and discourse analysis (1984:306). Believing that coherence is essential for writing to communicate, and that it should be an important component of writing instruction, she recognised that “... to teach coherence more effectively, we need a better understanding of the linguistic features and rhetorical structures that create coherence ... ” (1984:306). To develop her scale, she therefore drew up as an initial framework, a list of linguistic features and discourse structures taken primarily from van Dijk (1977; 1980) and Halliday and Hasan (1976). She rejected the notion of cohesive ties as sole markers of coherence, and accepted van Dijk’s (1980) claim that to assess global coherence, one should take into account the structure, plan and schema that order propositions (van Dijk, 1980, as cited in Bamberg, 1984:307). Bamberg concluded that:

“Meaning and coherence are not inscribed in a text, but are constructed by readers who are guided by textual cues and by their own knowledge and expectations to bridge gaps and to fill in assumed information.”

(Bamberg, 1984:307)

Bamberg then assessed a subgroup of essays and divided them into four groups depending on their level of coherence. She finalised her four-point rubric by comparing the features in each group of essays with her list of features (1984:309). Her scale was then assessed for construct validity and revised, after which it was implemented in a pilot study which proved it reliable.

Bamberg’s scale has various strengths. It assesses coherence holistically by looking at the whole essay and in terms of a list of features that are believed to create both global and local coherence. In addition, it regards coherence as being a relative quality in that essays are seen to achieve various degrees of coherence rather than being coherent or not coherent. One limitation the scale has, however, is that it does not give insight into the strengths and weaknesses of a group of essays. To counteract this, Bamberg suggests that a feature analysis also be done of the essays

(cf. Myers, 1980:58). The present study undertook an analysis of the specific feature of lexical cohesion and its effect on the coherence of student academic writing.

3.2.2 Cohesion

In the analysis of lexical cohesion, the present study used Stotsky's adaptation of Halliday and Hasan's (1976) categories, as her categories of lexical cohesion were specifically derived to analyse expository essays as opposed to Halliday and Hasan's classic model which was based on conversational or literary discourse (Stotsky, 1983:431; 1986:279; see 2.2 above).

Stotsky reclassifies Halliday and Hasan's (1976) categories of lexical cohesion as follows:

- | | |
|----|---|
| i. | <p>Semantically related words: a type of cohesion in which a lexical element (or group of lexical elements) is systematically related to a previous element (or group of elements) through:</p> <ol style="list-style-type: none"> 1. repetition: e.g., <i>test / test; high school diploma / high school diploma</i> 2. synonymy or near-synonymy: e.g., <i>difficult / hard</i> 3. opposition or contrast: e.g., <i>simple / difficult</i> 4. inclusion as a superordinate: e.g. <i>animals</i>, subordinate: e.g. <i>pigs and dogs</i>, or coordinate member in an ordered or unordered set of general or specific terms: e.g., the days of the week constitute coordinate members of an ordered set; the names of occupations, such as <i>teacher / janitor / cook</i>, constitute coordinate members of an unordered set. 5. derivation or repetition of a derivational element: e.g., <i>benefit / beneficial; employer / worker</i> |
| ii | <p>Collocationally related words: a type of cohesion in which one lexical element is related to another only through frequent co-occurrence in similar contexts: e.g., <i>worker / skills / job</i></p> |

(Adapted from: Stotsky, 1986:279)

It was decided only to focus on semantically related words in the present study in order to avoid the problems associated with studying collocation empirically (cf.

Hasan, 1984; see 2.2.1.6 above).

As part of the procedure for counting ties of repetition, it was necessary to determine what would count as an instance of repetition. Following Stotsky's (1986) procedures (see 3.2 below), it was decided that apart from identical words the following would count as ties of repetition:

- phrases and fixed expressions, e.g.: *the fruits of their labour*
- words having different inflectional endings, e.g.: *exploiting / exploit, leave / left*
- words having different comparative endings, e.g.: *working harder / work hard*

Synonymy can be thought of as occurring along a continuum (cf. Cruse, 1986; see 2.2.1.2) from synonyms identifiable as part of the language system in general to local synonyms identifiable as synonymous only within the text in which they occur. All tokens of both systematic (general) and instantial synonyms were included in the analysis. Thus, pairs of synonyms such as the following were counted in the study: *more prosperous / richer, supervision / rule; benefit / advantage; cunning / cleverer, working / labours*.

Several subcategories of opposition have been posited (Cruse, 1986; Lyons, 1968; see 2.2.1.3). For the purposes of the study these subcategories were broadly used to determine what ties counted as instances of opposition although text-specific instances of opposition were also noted. Ultimately, these subcategories were not specified in tallying the oppositional ties, there being too few instances of these ties to warrant it.

Under instances of inclusion, ties between superordinate and subordinate words such as *abuse of power / corruption; horses / foals* and *animals / pigs and dogs*, and between co-hyponyms such as *hens / sheep* were included. Ties formed with general words, although identified as weak relations, were also included as they seemed generally to have been used with specific intent by the writers.

Derivatives and derivational elements were considered important sources of potential lexical cohesion within the context of any study of student academic writing (Stotsky, 1981). Derivatives were identified as those words which had been created by the addition of a prefix or suffix to the base word, and as those words which had been created by zero-derivation or conversion (see 2.2.1.5). Derivational pairs such as the following were counted as ties: *administration / administrating; organized / organizing; Animalism / animal; change (n) / change (v)*. Derivational elements such as *ruthlessly / basically* were also counted on the grounds that they bring an added cohesion to the text and allow the writer a certain stylistic flexibility (Stotsky, 1983). Although the identification of derivational elements is subjective, certain principles were established in order to make the identification process more systematic. The derivational elements had to be part of low frequency words that would be considered to have the potential for forming cohesive relations, for example, *democracy / autocracy*. They had to occur at a distance near enough to one another for the researcher to associate the presupposing item with that which it presupposed. Derivational pairs had to appear to have been intentionally chosen by the writer to produce, for example, the effect of parallelism.

As regards identifying ties between the first and last paragraphs respectively and the question prompt, it proved impossible to count numbers of each token of lexical cohesion, as this produced too many null values for the appropriate statistical tests to be implemented. As a result, only the total number of ties for each script was counted. In many cases the essays had not been divided into discernable paragraphs and it was necessary to estimate by assessing the content of the script where the introduction ended and where the concluding statement began. To eliminate the effect of the variation in paragraph length, the total number of ties was divided by the total number of F-units (see 3.2.4 below) contained in what was estimated to be the writer's opening or closing statement. For some scripts no indication of an attempt at an opening or closing statement could be found. These scripts could not be analysed and this is reflected in the number of observations for the statistical tests performed on these variables which was lower than the maximum number of observations possible.

3.2.3 Errors

The present study adopted Engber's (1995) error classification as a means of assessing the errors found in the corpus of scripts under analysis. Of the descriptions of error examined (see 2.3.1 above), Engber's subcategories of error which included both meaning and form were deemed most appropriate as a classification of the types of errors made in these scripts. Operating within the context of discourse analysis, Engber developed her categories of error in order to relate measures of lexical error to the quality of written compositions of second language learners of English from mixed language backgrounds. In the same way, the present study attempted to determine the effect of lexical errors on the holistic coherence of scripts.

Below is a list of Engber's subcategories with examples taken from the scripts of the present study:

- I. **Lexical Choice**
- A. Individual lexical items
1. Incorrect - semantically unrelated
 - a. Pigs' *attitudes of getting drunk*. (tendency towards)
 2. Incorrect - semantically close
 - a. There is a windmill, threshing machine, and an *elevator* of its own. (Hay elevator)
- B. Combinations
1. Two lexical items
 - a. The animals were able to work harder *than before when they were in Manor Farm*. (during the time that)
 2. Phrases
 - a. until *every atom of strength* has been sapped. (ounce)
 3. Multiple errors involving core lexical items
 - a. Major has this dream of overthrowing the human and *look for selfgoverning*. (and working towards self government)
- II **Lexical Form**
1. Derivational errors
 - a. Squealer justify and say that the *provide* is very high. (provision)
 2. Verb forms
 - a. The pigs are taught and *educate* so that they can be the rulers on the farm. (educated)
 3. Phonetically similar, semantically unrelated
 - a. The animals is tired of being *explore* by Jones. (exploited)
 4. Word distorted - major spelling error
 - a. The conditions were better because every animal received his share of diet at a *stibelated* time. (stipulated)

(From: Engber, C. A., 1995:146, Table 3)

The present study largely followed Engber's procedures for counting errors, but deviated from her scheme where the aims and limitations of the study made it necessary to follow a different procedure. For instance, in line with Engber's classification, the present study did not count grammatical and syntactic errors. Thus, concord errors and inappropriate tense changes were not included in the analysis. However, all instances of lexical error were counted (Engber counted instances of identical error only once) as the study aimed to examine the density of error in student academic writing. Errors in the scripts were analysed according to Engber's subcategories of error but because the frequency of lexical errors was found to be low, only the total error count per script was tested statistically. The density of lexical error was then tested statistically for each script and correlated with the holistic coherence rating given to the script.

3.2.4 Textual units

In order to measure both the length of a text and the density of cohesive ties in that text, one requires a unit of measurement within which to do this. As Halliday (1989: 66) notes:

“... words are not packed inside words; they are packaged in larger grammatical units - sentences and their component parts. It is this packaging into larger grammatical structures that really determines the informational density of a passage of text.”

Practically, texts are segmented into these “larger grammatical structures” or textual units in order to count the number of cohesive ties per unit and thereby to determine the density of cohesive relations in the texts. Various grammatical structures have been used for studies of feature analyses of this nature, most commonly the sentence, clause and T-unit. The structure deemed most suitable for the present study was the functional unit of discourse or F-unit, developed by Lieber (1979) as a basis for segmenting the expository essays which formed the corpus of her study.

As a unit of analysis, the F-unit is more discriminating as a textual unit than the sentence which Halliday and Hasan made use of in their 1976 study. Although they

acknowledge that any grammatical structure is irrelevant to cohesive relations and that cohesive relations occur both within and across sentence boundaries, Halliday and Hasan examine only cohesive relations which occur across sentences:

"... since the effect is more striking and the meaning is more obvious: cohesive ties between sentences stand out more clearly because they are the ONLY source of texture, whereas within the sentence there are the structural relations as well."

(Halliday and Hasan, 1976:9)

While the sentence is still used as a unit of analysis in studies today (cf. Hoey, 1991; Sinclair's, 1993, structural cohesion involving the encapsulation of sentences), it was considered unsatisfactory as a unit within which to analyse texts of student academic writing in a study operating in an applied linguistic context. The reason for this is that, as Markels (1983:453) suggests, within a sentence there are "no syntactic limits to its length or complexity once the minimum requirements of subject and predicate have been met ...". This is the case in student writing in particular. In sentences such as the ones below where the writers have produced long, complex sentences, there are elements of lexical cohesion which would have to be disregarded as textual relations, because they fall within the boundaries of the sentence.

- [1] The pigs get more food rations because Squealer convinced the other animals that since the pigs are the most intelligent on the farm, and are the most hard working on things called "files", "reports", "minutes" and memoranda, they have to get more food rations.
- [2] In fact, during the time when the farm was called Manor Farm, the conditions were better because every animal received his share of diet at a stibelated time, there was no hunger (except the time when Mr. Jones got drunk), there was no misunderstanding between the animals, no animal ruled over the others.

Hunt (1970:189 as cited in Hubbard, 1989:62) defines the T-unit as "one main clause plus whatever subordinate clauses are attached to that main clause". Both the T-unit

and the clause, while more discriminating than the sentence, are too large to account for certain structures which function rhetorically in furthering discourse but are realised in the surface structure of sentences as phrases with nonfinite verbs or reduced phrases with no explicit verb (Lieber, 1979:58 & 59).

In order to include in her analysis of cohesion instances of ties between all those units isolated as minimum units of structure, Lieber's F-unit encompasses "conjoined coordinate clauses, subordinate clauses, zero-verb clauses, and phrasal clause-equivalents" (Lieber, 1979:99). To distinguish between clauses and clause equivalents which promote rhetorical development within an essay from those which only serve to fill syntactic slots, Lieber uses the concept in transformational grammar of the root sentence (see Lieber, 1979:58 - 83 for a full discussion of this approach). A structure which is dominated by an S-node will make up a base sentence in underlying structure but may be realised in surface structure as a phrase such as a reduced, non-restrictive relative clause or reduced appositive. These structures are accepted as F-units. For example:

- [3] During the time of Snowball // the brain behind the windmill, // the animals were told that the windmill was to provide things like electricity, and hot and cold water.

A non-root sentence by contrast falls under the domain of a structure such as a noun phrase or a verb phrase. A non-root sentence (such as a restrictive relative) cannot be removed from the matrix clause without the clause becoming incomplete as a result. Non-root sentences are, therefore, not accepted as F-units. For example:

- [4] The humans are the only animals that consume and never produce.

Hubbard (1989:120) modifies the F-unit to include temporal and locative clauses (excluded by Lieber on the grounds that they function in the same way as do adverbial conjunctions, although as Hubbard points out, she does accept conditional clauses introduced by *when*). The present study adopted Hubbard's modification,

particularly as temporal clauses are used extensively in narration (a structure incorporated into the expository writing of the corpus of scripts analysed in the present study). Both Hubbard (1989) and Maringa (1995) have adopted the F-unit successfully as the unit of analysis in their studies of cohesion.

The subcategories of the F-unit, with examples from the present study, are listed in Appendix C on page 155.

3.3 RESEARCH PROCEDURES

Initially 110 scripts were randomly selected from the larger population of scripts from the end-of-year examination for the subject Practical English, which incorporates both language and literature components and which is offered at undergraduate level at university. Of these 110 scripts, ultimately only 70 were selected based on restrictions set on text length and format. It was decided to accept only scripts of 18 F-units or more in order to reduce the possibility of text length influencing rater perceptions. In addition, it was decided only to accept texts written in paragraph or point form. Texts which had been set out in columns, comparing point-by-point conditions on *Animal Farm* with conditions on *Manor Farm*, were excluded on the grounds that this particular structure might negatively influence raters' perceptions. The home language of the writers was not taken into consideration as the aim of the study was to ascertain the effect of cohesive relations on the coherence of student academic writing in general. Markers of student examination essays do not usually have access to information regarding the home language of the candidate and, therefore, do not take into consideration whether the candidate is a first or second language speaker of English when awarding a mark for the essay.

The question to which the scripts were a response is presented in full in Appendix A, page 150. The question provides a background passage from the literature set work with which candidates had been required to familiarise themselves as part of their course of study. The question required candidates to compare two situations presented in the set work. Reid (1990, as cited in Engber, 1995:143) found that

comparison/contrast topics resulted in more content words being used by the writer than did essays requiring descriptions or visual interpretations. Engber specifically selected a topic requiring comparison and contrast for her 1995 study in order to ensure that the essays would contain enough concrete vocabulary for a lexical study. It was felt that analysing texts that conformed to a topic requiring candidates to form comparisons and contrasts would promote the use of lexical relations such as opposites and synonyms. This would be apart from the relations of derivation usually expected of academic writing and the relations of repetition and inclusion usually found in such essays.

Having selected the 70 scripts for analysis, it was discovered that the text sample was negatively skewed as the scripts were better than expected. The mark allocation of the scripts ranged as follows (the highest mark obtainable was 15):

number of scripts	mark
1	14
6	13
5	12
19	11
18	10
8	9
5	8
8	7

Three independent raters were chosen to assess the scripts in terms of Bamberg's holistic coherence scale. The raters were all teachers of English at upper secondary school level. Raters were asked to read through the texts quickly and rate them according to their first impression in line with the four categories of the scale (cf.

Myers, 1980:42). The ratings were then added up to give a final holistic coherence rating (HCR) for each script. Inter-rater reliability was found to be highly significant according to the Spearman-rank order correlation test with the moderately good correlations of .46 between raters 1 and 2, .41 between raters 1 and 3, but the fairly low correlation of .29 between raters 2 and 3.

The study followed Stotsky's guidelines for identifying, analysing and counting cohesive ties. One of the initial procedures was to decide what constitutes a lexical item. Halliday and Hasan maintain that:

"The concept of the lexical item is ... not totally clearcut; like most linguistic categories, although clearly defined in the ideal, it presents many indeterminacies in application to actual instances"

(Halliday and Hasan, 1976:292).

In the text samples at the end of their 1976 work on cohesion, they include words, phrases and clauses as cohesive items (Stotsky, 1986:290). In fact, Carter (1987:49) suggests that it is difficult to demarcate what might be the upper limit of a lexical unit. Following Halliday and Hasan (1976), Stotsky counted phrases, clauses and independent words as single lexical items when they enter into lexical cohesive relations with other words, phrases and clauses (1986:280).

In line with Stotsky's procedures, the present study counted as an instance of repetition only an identical word or a word varying in inflectional or comparative ending (Stotsky, 1986:280). Words which are morphologically related but have a different form such as *differ / different* and *animal / Animalism* were classified as derivatives.

Any phrase or word that was repeated, such as *hard working* or *the rest of the other animals* or any idiomatic expression was counted as only one token of repetition. For example, the name *Animal Farm* was counted as one instance of repetition rather than as consisting of two words. The identification of idiomatic or fixed expressions is a subjective process which presented difficulties in the analysis of the corpus. Carter (1987:63 - 64) has devised a list of criteria for formal linguistic recognition of

fixed expressions. These criteria are: collocational restriction; syntactic structure; semantic opacity. Each criterion operates on a cline from less fixed, for example, *take a look / a holiday / a rest*, to more fixed, for example, *dead drunk, the more the merrier, right on*. Fixed expressions are generally not changeable and relate by repetition. Although Carter's scheme was used to identify fixed expressions in the present study, it was found that very few fixed expressions could be identified in the scripts. On the occasion that the writers attempted to use fixed expressions, they often produced them incorrectly, for example, *every atom of strength* instead of *every ounce of strength*.

Stotsky counted both inter-sentence ties (textual or nonstructural ties) and intra-sentence ties (structural ties) but coded them separately. In doing this she followed Eiler (1979) whose study of cohesion in the expository writing of ninth graders (as cited in Stotsky, 1986:290) focussed on intra- and inter-sentential ties. Halliday and Hasan (1976) do not allow for structural ties in the analysis of cohesive relations. As Hasan (1984:184) points out, structure, by definition, is coherent and structural units (for Hasan the basic structural unit is the clause) which are internally coherent, do not necessarily form a text when grouped together. Texture is provided by semantic relations which are reflected in nonstructural cohesive chains (1984:184). The focus of the present study was specifically on textual ties, the study being interested in the broader connections made by students across F-unit boundaries. However, because the textual unit is not strictly adhered to by researchers (cf. Stotsky, 1986; and Eiler, 1983), it was decided to examine both aspects of cohesion and, therefore, both structural and textual cohesive relations as defined in terms of F-unit boundaries were counted. In the corpus examined, instances of structural ties were largely limited to instantial ties of opposition contrasting initially the humans, and then later the pigs, with the other animals. It was felt that it was important to include these ties in the total count as they were indicative of the way writers were using cohesive ties to distinguish between the states they were describing. Structural ties were distinguished from textual ties by the use of a superscript *w* (Stotsky, 1986, uses "^w" to indicate that the tie is found within the sentence). The analysis revealed that there were very few cases of structural ties (only 5% of total ties) and, as a result, they

were not given a separate tally from that of the textual ties.

Like Stotsky, the present study counted relations systematic to the language (general ties) as well as those systematic to the texts (instantial ties). Text-specific systematic ties or instantial ties were counted with the other systematic ties under the appropriate categories of semantically related items. Systematic ties and instantial ties were not tested separately, however, as Stotsky's findings suggest that the distinction between these types of ties has no developmental significance in student writing (1986:284).

The study included multiple ties where individual words within a group of words that formed part of a tie, formed ties with other previous elements within the text (Stotsky, 1986:281). For example, *Napoleon and his clansmen* has an opposite relation with *the animals*. At the same time *Napoleon* relates to a reference to *Napoleon* earlier on in the text in a relation of repetition, and *clansmen* relates to *family* in a relation of synonymy. Although the present study included such multiple ties, unlike in Stotsky's study it did not label them as such.

As with Stotsky, the present study only counted one tie when it was possible for a lexical item to enter into more than one tie at once. Like Stotsky, a textual tie was counted before a structural tie. Although Stotsky counted examples of both systematic ties and instantial ties, she maintains that it is theoretically more defensible to distinguish ties with stable relations in language from those which are related only through frequent co-occurrence in a text (1986:279).

The study counted both ties that pointed backward and those that pointed forward. The most common direction of lexical ties is anaphoric where the presupposed item precedes the presupposing item. What is presupposed may be in the immediately preceding sentence, in a much earlier sentence or may be the whole of a longer, preceding passage (Halliday and Hasan, 1976:14).

According to Halliday and Hasan:

"If the cohesion is lexical ... the second occurrence must take its interpretation from the first; the first can never be said to point forward to the second."

(Halliday and Hasan, 1976:17)

Stotsky, however, disputes this point in terms of relations of inclusion where she claims that it is just as likely for a superordinate item to appear before a subordinate item(s) which appears later in the text, as it is for the superordinate word to follow the subordinate one. In the present study it was found that although most ties were anaphoric, there were a few cataphoric ties of inclusion where a superordinate item pointed forward to subordinate items in the text. This type of cataphoric tie, however, occurred very rarely.

The relation of exophora which accounts for any element which points to the non-linguistic context outside the text was not examined in the study. Although other studies of cohesion have included exophoric relations (cf. Lieber, 1979; Hubbard, 1989), only anaphoric and cataphoric relations cohesively bind elements of a text together and it was deemed more useful to limit the study to relations of this nature. In analysing the lexical relations between the opening and closing paragraphs of the student writing and the question which prompted the writing, the study specifically chose to look at intertextual ties. These ties are not examples of exophoric ties as they do not refer to a text world that has to be inferred by the reader (Brandt, 1986:95), but rather point anaphorically to the relevant examination question which is a text directly accessible to the reader and which can be linked cohesively to the essay texts via cohesive relations.

3.4 SAMPLE ANALYSES

Samples of the analysis of cohesive ties, errors, and relations between first and last paragraphs and the question have been included below to provide details of how the analyses were undertaken, what decisions were made with regard to analytical procedures, what difficulties were encountered and how they were solved.

The sample texts and analytical tables that follow them are provided in Appendix D from pages 158 to 175. The texts chosen include the highest and lowest scripts in the corpus in terms of score or holistic coherence rating (HCR). Script 58 has an HCR of 5 (the lowest in the corpus) and a score of 7; script 49 has an HCR of 12 and a score of 14 which is the highest score obtained by any script. Two other scripts were chosen as containing interesting examples of error (script 17) and first and last paragraphs (script 6) respectively.

3.4.1 Cohesion analysis

In the analysis below, certain cohesive relations from all four sample texts have been chosen for exemplification.

Each tie is presented with the presupposing item first followed in brackets by its F-unit number. This is then followed by the category of tie according to Stotsky (1986). Lastly the presupposed item is presented followed by its F-unit in brackets.

- (R) Repetition
- (S) Synonymy
- (O) Opposition
- (I) Inclusion
- (D) Derivation

A superscript *w* indicates the presence of a structural tie.

A superscript *k* indicates the presence of a cataphoric tie.

The determiner of each noun phrase has been included in brackets. The determiner is not counted as part of the tie, but may sometimes play a role in determining the identity of reference between the two elements of the tie.

Script 58 (page 158)

(the) pigs(7) (O) Mr Jones(6)

This is a text-specific systematic (instantial) tie of opposition. The pigs are in this instance being compared to Mr Jones as agents running the farm. The tie is textual; "but" signals the beginning of a new F-unit which takes the form of a non-verbal phrase (see Appendix D, page 158).

Mr Jones (21) (O)* (The) pigs (21)

This is a text-specific relation of opposition again comparing the pigs with Mr Jones. This relation, however, occurs within the same F-unit and has, therefore, been labelled a structural tie. It has been counted as a cohesive tie, because it exhibits the same relationship as that of the tie between the items in F-unit 7 and 6 (see 3.2).

(The) pigs and the dogs (17) (O) (The) animals (15)

This tie has been classified as a text-specific relation of opposition. It appears that for most of the scripts analysed the pigs and dogs are set in opposition to the rest of the animals and, therefore, ties of this nature have been defined as opposites rather than hyponyms.

Script 17 (page 160)

(the two) farm (15) (R) (the) farm (11)

The farm exists under two different management systems at different points in the novel. It is the writer's task to compare these states. Here the writer explains that the "two farms" contain the same conditions. The "two farms" in F-unit 15 relates back to "the farm" in F-unit 11 where the farm under the first management system is described. Later in the text F-unit 18 brings up the farm again. This time the

reference is to the farm under the control of the animals. This reference to the farm relates via the reference in F-unit 15 to the reference in F-unit 11, but it is clear that the references in F-unit 11 and in F-unit 18 and those following it are exclusive.

(all the) work (26) (D) work (hard) (7)

This tie has been classified as a relation of zero-derivation or conversion where each element in the tie comes from a different word class but there is no change in word form.

Script 6 (pages 165)

things (19) (I)^{wk} electricity and hot and cold water (19)

This has been defined as a relation of inclusion. The lexeme "things" is an example of a general noun which although they are defined as lexical items can have grammatical functions, and, as examples of high frequency words, contribute little to lexical cohesion (Lieber, 1979; Halliday, 1989:64 - 65). In this case, however, it has been counted as a tie as the cohesive relation is reinforced by the tie formed by "items" in F-unit 21 which relates back to "electricity and hot and cold water" in F-unit 19 as a relation of inclusion. This tie is also both cataphoric and structural. However, it appears as if it has the function of creating a list and that it should, therefore, be counted as a cohesive relation.

Animalism (22) (D) (the) animals (19)

This has been classified as a text-specific tie of derivation. "Animalism", in the context of the novel to which the texts make reference, denotes the philosophy adopted by "the animals" to promote their cause. "Animalism" is derived from "animals" by the addition of the suffix "ism" which is used to indicate a form of doctrine (Shorter Oxford English Dictionary, 1972:68).

management (28) (S) administering (4)

This has been classified as a tie of synonymy although the elements do not come from the same word class (cf. Halliday and Hasan's 1976 sample texts).

Ironically (36) (D) basically (33)

This tie is formed as a result of the repetition of derivational elements which can add to the cohesiveness of otherwise unrelated words and create stylistic flexibility on the part of the writer (Stotsky, 1983:434).

brief essay (42) (S) short exposition (1)

shows (42) (S) (will try to) reveal (2)

These are ties of synonymy which the writer uses effectively to link the last paragraph to the first paragraph in order to create a sense of coherence within the essay.

Script 49 (pages 169)

useful (3) (D)^w used (3)

This has been counted as a structural tie of derivation.

slaughtered (5) (O) see out their natural life (4)

This has been defined as a text-specific tie of opposition. The writer seems to make a link between these elements by using the word *rather* which begins F-unit 5. (Markels, 1983, has posited that syntactic elements allow the interpretation of semantic chains.)

(the) Rebellion (16) (S) (the) uprising (2)

This is a systematic tie of synonymy. The elements of the tie relate to each other systematically within the English language as types of synonyms and so they have been counted as part of a synonymous tie here.

(All) animals (18) (R) (the) animals (1)

Here the determiners indicate that there is a relation of inclusion of reference between the elements of this tie. The animals on Animal Farm are included in the set of all animals within the context of the novel who are declared equal. Halliday and Hasan's (1976) coding of the type of ties indicates the type of identity of reference between the elements of the tie.

left (23) (R) leave (15)

This has been classified as a tie of repetition following Halliday and Hasan's (1976) and Stotsky's (1986) definition of a lexical item which includes different tenses as part of the same lexical item.

3.4.2 Analysis of first and last paragraphs

The tables of analysis for the first and last paragraphs for each of the script below are included in Appendix D, page 158.

The writers primarily used ties of repetition to link their essays to the question prompt. This confirms Brandt's finding that students use only a narrow range of linguistic features to introduce and close their essays based on what they think the teacher (in this case the examiner) expects of them (Brandt, 1986:101). The students occasionally used derivatives to link the verb *differ* in the question to the noun *differences* or adjective *different* in their opening or closing paragraphs. Ties of synonymy, opposition or inclusion were very rarely used.

Script 58

There is no real sense of introduction or closure in this essay in terms of paragraph structure or signals of cohesive relations with the question prompt. The essay is only 25 F-units long and has been written as one paragraph. As part of the question the essay compares the two farms only at the point in the novel where the background passage occurs. Therefore the progress of the narrative and the changes that take place are not delineated. It is possible to accept the first 8 F-units as the opening paragraph as here the writer begins to describe the state of the farm under the pigs. The next section which could be understood to be the body of the essay begins at F-unit 9 and the closing paragraph at F-unit 18. The fact that the writer seems to start new topics at these junctures suggests that these F-units could be the beginning of new sections. The essay is not directly related to the question except for the repetition of "Manor Farm" (19) in the last section.

Script 17:

The essay has been divided into separate paragraphs by the writer and at first appears to have a coherent text structure. The first paragraph is 12 F-units long and the last paragraph 4 F-units long. Each paragraph deals with a different topic. The first and last paragraphs however do not provide the reader with a definite sense of introduction and closure respectively and relate to the question minimally in terms of lexical ties. There is only one lexical element in the first paragraph which refers back to the question prompt. This is the repetition of "Animal Farm" in F-unit one which the writer uses to contextualise the writing by referring to the work of literature on which the question is set. "Animal Farm" here is also used to refer to the farm discussed in the opening paragraph as no further mention is made of the farm on which the animals experience the conditions listed. The writer seems to assume that the reader will understand that the essay details the conditions on each farm. Examples of the conditions are given with no explicit reference to the word "conditions" until F-unit 49 in the penultimate paragraph. The essay lacks sufficient

text organisation for the reader to follow the writer's argument easily. The absence of lexical signals, such as appropriate references to the names of the farm, add to the processing time the reader must spend interpreting the content of the essay.

Script 6:

There are clear opening and closing paragraphs both of which relate to the question. The opening paragraph relates the essay to the question prompt by answering the question directly, (1) "In many respects conditions on Animal Farm greatly differs from that of the old Manor Farm of Mr Jones". Thus the writer uses several ties of repetition to link the first paragraph to the question prompt. The second sentence in the paragraph states the aim of the writer in writing the essay. The sentence repeats the topic of the essay, "the conditions of the Animal Farm", which once again links the essay to the question prompt through repetition of lexical items. The third sentence introduces a narration of the events on the farm which bring about changes to the conditions on Animal Farm and thus illustrate the central argument of the essay. The closing paragraph consists of only one sentence. This sentence serves merely to sum up the writer's contention " ... that conditions on the Animal Farm completely differ from those on Manor Farm". This sentence refers back to the question prompt via repetition relations. At the same time it links back to the second and organising sentence of the essay with "This brief essay" which relates synonymously to "This short exposition". Thus, the anaphoric links between the first and last paragraphs of the essay and the question prompt reveal the ability on the part of the writer to structure and organise the essay and to signal this organisation lexically.

Script 49:

The writer's first paragraph is 15 F-units long. It lists all the conditions on Manor Farm prior to the rebellion which it ends by describing. The first 5 F-units have been taken as the opening paragraph for analysis as they give a brief overview of the conditions - "miserable, laborious and unacceptable". They then highlight the first

problem; that animals are slaughtered before their time. The rest of the paragraph deals with rations and “the fruits of their produce”.

Although the writer has not provided a clear closing paragraph, the final 9 F-units of the essay were analysed as a conclusion. The writer sums up the answer to the question in F-unit 38 which reads: “In actual fact, the conditions on Manor Farm do not nearly differ from those on Animal Farm in respect of the animals”. The essay is then set in a universal context and the rebellion on Animal Farm is compared with the process of rebellion in modern countries around the world. The result of introducing a new topic at this stage, is that the writer brings in a new set of vocabulary items which do not necessarily tie to the question or opening paragraph. However, the writer does provide several repetition links to the question prompt in F-unit 38 which presents a statement of the writer’s answer to the question and concludes the argument presented in the essay. The F-units which follow are used to expand on the writer’s statement by providing the reader with evidence for it and creating an analogy which broadens the significance of the story by placing it within the broader context of human political action. The reader is thus presented with a clear organisational text structure which is clearly demarcated by lexical signals.

3.4.3 Error analysis

The errors in script 17 have been listed below and classified according to Engber's subcategories. To ascertain why the lexical item in question caused the perception of error, each error was examined in the context of the F-unit in which it occurred.

(6) The animals is tired of being **exploire** by Jones.

This is an error of lexical form where the student appears to have mistaken *exploire* for *exploit* which is phonetically similar but semantically unrelated.

(9) Major has this dream of overthrowing the human and **look for selfgoverning**.

Although much of this sentence appears to contain errors, for example, *this dream* has a colloquial feel to it and *the human* is definitely an error, it was decided that the only lexical errors occurred in the phrase *look for selfgoverning*. This phrase in itself contained different types of error and was thus classified as an error involving key elements.

(11) and will enjoy the **fruits and profits** of the farm.

The writer appears to have been trying to write a phrase based on the idiom: *to enjoy the fruit of one's labour*. The addition of the word *profits* appears to cause the writer to abandon the idiom and the phrase becomes ambiguous as it appears that *fruits* may refer to the types of fruit produced on the farm.

(22) The pigs are taught (23) and **educate** (24) so that they can be the rulers on the farm.

The error here is with the form of the verb. It appears from the context that it is the pigs which are educated so that they can rule the farm. The verb *educate*, however, suggests that the pigs are doing the educating and this results in ambiguity for the reader.

(34) Animals are **ration** (35) and do not have enough to eat, ...

This is also an error of verb form. The reader must use the context provided by the next F-unit to establish what the writer means. This could be classified as an "extraction" error in that it requires extra processing time on the reader's part, although the reader is probably ultimately able with the help of contextual clues, to "extract" the meaning intended by the writer (Hubbard, 1994).

(36) Squealer justify (37) and say that the **provide** is very high.

This is an example of a derivational error. The writer intends to write *provision* but

does not know how to derive the noun form of this item from the verb form.

(38) **Trate** business is establish with the neighbouring farmers.

This is a spelling error which has caused a word distortion. The word appears at the beginning of a sentence which allows it to be interpreted as the name of a business or as some type of business. Readers must draw on their knowledge of the text world to understand that the writer intends them to understand that a system of trade was set up between neighbouring farmers.

(39) The hens are ask to do sacrifices (40) and when they **receive** (41) they are killed.

The writer appears to have substituted *receive* for *refuse*. These words are phonetically similar although semantically unrelated.

(42) Animals has to confess and **executed**.

This is another example of an error in verb form. The writer has omitted part of the verb which makes the meaning of the proposition less clear.

(44) and the Seven Commandments are altered to **soothe** the pigs.

Again the writer has substituted a phonetically similar word for the one intended which in this case is probably *suit*. The words have no semantic relation and it is only through the context that the reader can work out the writer's intended meaning.

(46) he uses language (47) to **manupilate** the animals.

Manupilate is an example of an error made as a result of a major spelling error or word distortion.

(49) but comrades was informed that things **where** better than before.

The writer here has confused two potentially similar sounding words which do not share the same meaning but which also share similar spelling patterns. The error in this case may be a case of the writer confusing the two words due to the pressure of writing under the time constraints imposed by the examination. Alternatively, the words *were* and *where* sound similar to the writer in terms of his or her phonological system.

3.5 CONCLUSION

Chapter 3 outlined the research design and analytical framework for the study. The concept of coherence was defined for operational purposes in terms of Bamberg's (1984) Holistic Coherence Scale. The lexical cohesion sub-categories were delineated and procedures for identifying types of lexical cohesive items established. The general research procedures involved in analysing the data and testing for coherence were then outlined. Lastly, samples of the methods of analysis were given.

CHAPTER 4

RESEARCH FINDINGS

4.0 INTRODUCTION

This chapter presents the findings of the study. It presents the statistical procedures followed and details the results of each hypothesis in terms of the statistical tests performed on the relevant data. The results are then interpreted and related to the aims of the study with the intention of determining whether the aims have been realised. The results are summarised in Table 4.1 and Table 4.2 at the end of the chapter.

4.1 RESULTS

The results and a short interpretation in respect of each hypothesis are presented below. Initially, however, an overview of the statistical procedures followed for testing the hypotheses will be given.

As the scores of the sample scripts seemed unexpectedly high, the sample was tested for skewness using the Statistical Package for the Social Sciences (SPSS) statistical computer software. Skewness refers to the "asymmetry of a distribution" (Mulder, 1986:43). A normal distribution is expected to be symmetrical. A distribution is skewed when the scores are "scrunched up" (Popham, 1981, as cited in Brown, 1988:89) toward higher scores, in which case the distribution is negatively skewed, or toward lower scores in which case it is positively skewed. The distribution was found to be negatively skewed and a Kurtosis Test revealed that the distributional curve was steeped or peaked (Brown, 1988:91), suggesting that the distributional curve of the sample was not normal either.

Mulder (1986:49) maintains that the distribution of scores taken from a small sample will not be normal. He suggests that for a normal distribution, the sample size should be at least 100 (1986:53). In the case of the present study, an initial sample of 110

scripts was chosen. However, of this sample only 70 scripts were ultimately chosen for analysis on the basis of their conforming to a particular text length (18 or more F-units) and form (paragraph or list of sentences). This had the effect of reducing the sample size to one of fewer than 100 scripts as well as of possibly excluding shorter scripts which might have obtained lower scores.

A skewed distribution is undesirable as it will have the effect of depressing the values of any correlation coefficients obtained (Brown, 1988:145). As a remedy, rather than transforming the data, it was decided to use a non-parametric test which does not require normality. Parametric tests are more powerful than non-parametric tests but have a set of assumptions, one of which is that the distribution of the sample should be normal. Seliger and Shohamy advise the use of a non-parametric test when any of the assumptions of parametric tests are violated (1989:204). The statistical test chosen to evaluate the hypotheses of the study was the Spearman-rank order correlation coefficient (Rho) which does not assume a normal distribution in a sample.

Using the Spearman-rank order correlation involves the study in a correlational design. Correlational techniques are used to determine the degree of the relationship between two or more variables (Hatch and Farhady, 1982; Seliger and Shohamy, 1989). In the present study two sets of scores were obtained and correlated. One set of scores consisted of the holistic coherence ratings awarded the texts. The other set of measurements included the densities of both lexical cohesion expressions and lexical errors. The densities were obtained by dividing the number of ties by the number of F-units for each script. The set of scores obtained by the scripts was also correlated with the same densities of cohesion and error.

In general, high scores on one variable associated with high scores on the other will reveal a positive relationship between the two variables. A high score on one variable associated with a low score on the other will reveal a negative relationship between the two variables. If there is no relation between the two scores, there will be no systematic pattern between the high and low scores (Hatch and Farhady,

1982:195).

A correlational result is expressed in a correlation coefficient. A +1.00 correlation represents a perfect positive relation and a -1.00 correlation a perfect negative correlation. Correlation studies cannot be used to establish a causal relation between variables. Correlation coefficients are only used to show the degree of relationship between variables and the interpretations given these coefficients are usually cautious ones.

The Spearman-rank order correlation is used to measure variables on an ordinal or ranking scale. The coefficient reveals how the rankings of the scores on the variables are related (Hatch and Farhady, 1982:205). A weakness of the Spearman correlation elucidated by Hubbard (1989:242) is that the results are less valid where there are large numbers of null values. Fearing that some of the variables in the present study might present this problem, the null values of each variable were counted. The numbers of null values (presented below) did not seem great enough to prevent statistical testing given the sample size, although the ten null values in the *ties in closing paragraph* measure were considered fairly high.

Variable	Null values
total number of cohesive ties	0
ties of repetition	0
ties of synonymy	4
ties of opposition	1
ties of inclusion	6
ties of derivation	5
ties in opening paragraphs	5
ties in closing paragraphs	10
errors	2

Traditionally, when formulating hypotheses for the study, the researcher establishes the null hypothesis which states that there will be no significant correlation between the variables under examination. An alternative hypothesis which states that there is a significant relation between variables is posited but is not directly tested statistically. The researcher must test the null hypothesis in order to determine whether it is likely that the correlation is "real", in which case the alternative hypothesis is supported, or a result of chance, in which case the null hypothesis is supported (Brown, 1988:116).

To determine whether the null hypothesis can be rejected, it must be established at what level it can be accepted that a correlational relationship is not a matter of chance. This level is called the level of significance or alpha level. In language studies (Brown, 1988:166), it is customary to set the level at 0,05 which means that there is a 5% likelihood that a correlation happened by chance, or 0,01 which establishes that a correlation would have occurred by chance once out of 100 times or 1%. The levels are represented as $p \leq ,05$ and $p \leq ,01$ respectively. In the present study a correlation of $p \leq ,05$ was taken as significant and a correlation at the level of $p \leq ,01$ as highly significant, as is customary.

A statistical result should be significant in order to be considered meaningful. However, a significant result should not automatically be considered meaningful. Researchers are cautioned to examine correlational results carefully before assuming that statistically significant results are meaningful from a linguistic or applied linguistic point of view (Brown, 1988; Seliger and Shohamy, 1989). Seliger and Shohamy advise that one should look at the theoretical basis for suggesting a correlational relation between variables. If there is no theoretical basis for assuming a relation, then, they say, a correlation of, for example, 0,45 might be considered high. Where a high correlation was predicted, a correlation of 0,45 will not be considered high (Seliger and Shohamy, 1989:220 - 221). However, Brown (1988:142) points out that under certain circumstances even a low correlation of ,09 might be considered interesting.

In the present study the complexities of interpreting correlations will thus be borne in mind, but use will be made of Mulder's labels when interpreting positive correlations, as follows:

1,00	- perfect correlation;
0,80 to 0,99	- very high correlation;
0,60 to 0,79	- high correlation;
0,40 to 0,59	- moderate correlation;
0,20 to 0,39	- low correlation;
0,01 to 0,19	- very low correlation;
0,00	- no correlation.

(Mulder, 1986:73)

The following section details the results and interpretations in support of the hypotheses tested.

4.1.1 Coherence Hypotheses

The coherence hypotheses have been formulated to test to what extent the density of lexical cohesive ties and the density of lexical errors each contributes to the perceived coherence of texts of student academic writing. Specific sub-hypotheses have been established to test the relationship between each subcategory of lexical cohesion and the coherence ratings of the texts. In addition, hypotheses have been formulated to determine the extent to which densities of ties relating the opening and closing paragraphs to the question prompt have an effect on the perceived coherence of the texts.

4.1.1.1 H_1 Cohesion Density - Coherence Hypothesis (General)

Null hypothesis

$H_{0.1}$: There is no relationship between the density of lexical cohesion ties in a text

and the holistic coherence rating of that text.

The density of cohesive ties was established by dividing the number of ties by the number of F-units in each of the 70 scripts analysed. The figure in the table below thus reveals that the scripts contain on average 1,51 ties per F-unit. The holistic coherence rating is calculated out of a maximum of 12. The mean coherence rating of 9,67 can be seen to be fairly high indicating that the sample is negatively skewed not only in terms of academic score, but also in terms of coherence ratings.

Statistics	Density of cohesive ties	Holistic coherence rating
n	70	70
S.D.	0,32	1,95
Mean	1,51	9,67
r	0,3241	
Significance	0,006	

Result

The Spearman-rank order correlation reveals a highly significant ($p \leq 0,01$) relationship between the variables of cohesion density and coherence rating. The null hypothesis can therefore be rejected and the alternative hypothesis, H_1 , can be accepted.

Interpretation

While the correlation between lexical cohesion and coherence is highly significant, the correlation coefficient is low (see Mulder at 4.1 above). The results as expressed in the literature are varied, with Neuner (1987) and Tierney and Mosenthal's (1983) studies revealing no connection between number of lexical ties and coherence ratings, while Stotsky's (1986) findings indicate a positive relationship. Given that this study made use of Stotsky's subcategories of lexical cohesion, the results here might have been expected to emulate hers and the low correlation could be considered disappointing. However, it is important to take into consideration the fact

that the sample is negatively skewed with fewer poorer scripts forming part of the sample. It is possible that within a more normally distributed sample the correlation between the cohesion and coherence variables might have been greater.

4.1.1.2 H₂ Cohesion Density - Coherence Hypothesis (Specific)

Null hypothesis

H₀₋₂: There is no relationship between the density of any category of lexical cohesion (repetition, synonymy, opposition, inclusion, derivation) in a text and the holistic coherence rating of that text.

Each of the lexical cohesive expressions mentioned in the above hypothesis was tested. Their results are presented below:

(a) Repetition

Statistics	Density of repetition ties	Holistic coherence rating
n	70	70
S.D.	0,27	1,95
Mean	1,09	9,67
r	0,2144	
Significance	0,075	

Result

The result narrowly misses being significant at the $p \leq 0,05$ level. The null hypothesis therefore cannot be rejected and the alternative hypothesis, H₂, cannot be accepted.

Interpretation

This result reveals that number of repetition ties is not significant in contributing to

high coherence ratings which appears to confirm the findings of previous researchers with regard to repetition. Repetition devices are probably those lexical ties which are most pervasive in texts (cf. Tierney and Mosenthal, 1983: 227 - 228). Hartnett (1986) defines repetition as a type of static tie which she says appears frequently in student texts (1986:146). In their research, Witte and Faigley (1981) found that less skilled writers tended to repeat lexical items redundantly, while Stotsky (1986) found that repeated items were found in both high- and low-rated texts although the writers of high-rated texts produced proportionally more ties of all types. One would thus not necessarily expect ties of repetition to distinguish significantly between skilled and less skilled writing, as this result confirms. The analysis of the corpus of scripts in the study did in fact show that the majority of the scripts contained large numbers of ties of repetition and that ties of repetition far outnumbered any other type of tie.

(b) Synonymy or near-synonymy

Statistics	Density of synonym ties	Holistic coherence rating
n	70	70
S.D.	0,06	1,95
Mean	0,10	9,67
r	0,2122	
Significance	0,078	

Result

The result narrowly misses being significant at the $p \leq 0,05$ level. The null hypothesis is therefore not rejected and the alternative hypothesis, H_2 , cannot be accepted.

Interpretation

The non-significant result is in this case disappointing in that previous research seems to indicate that synonymy plays a significant role in contributing towards high coherence ratings in student texts (cf. McCulley, 1986; Neuner, 1987; Stotsky, 1986).

It is also disappointing in that the question topic was chosen on the basis that it might produce lexical ties of opposition and synonymy in the essays written in response to it. Tierney and Mosenthal's findings reveal that reference and lexical cohesion ties discriminate most strongly in cohesive patterning between topics (1983:219). The analysis of the scripts showed that the writers did not use large numbers of synonyms to compare the conditions on the two farms. This finding may suggest that ties of synonymy are not a good discriminating device in terms of coherence ratings for a topic of comparison and contrast.

(c) Opposition or contrast

Statistics	Density of opposition ties	Holistic coherence rating
n	70	70
S.D.	0,06	1,95
Mean	0,14	9,67
r	0,0595	
Significance	0,624	

Result

The result is not significant at the $p < 0,05$ level. The null hypothesis cannot therefore be rejected and the alternative hypothesis, H_2 , cannot be accepted.

Interpretation

This result confirms that for this sample, ties of opposition or contrast do not have a significant effect on the building of a coherent text. The fact that the scripts under analysis were a response to a topic that required the writers to compare and contrast two situations might have led one to expect writers to use many instances of ties of opposition and contrast to develop their ideas, whereas in fact the scripts under examination did not reveal large numbers of oppositional ties. However, it has been demonstrated (cf. Hartnett, 1986:145) that texts may signal contrasts using

adversative conjunctions such as *but* and indicate comparison using comparative and superlative forms of adjectives and adverbs. In the present study, comparative forms of lexical items were counted as instances of repetition. Denials are also used in written texts to express comparison or contrast (Pagano, 1994:258) such as in the following example:

- [1] *They believe they would be able to relax, which was not the case on Manor Farm.*

This sentence contains an implicit comparison between Animal Farm as the animals envisage it and Manor Farm. However, no explicit cohesive tie is used to signal this comparison. It appears that the writers may often not have used ties of lexical opposition to further the development of their ideas, choosing instead to use other means to do so.

(d) Inclusion

Statistics	Density of inclusion ties	Holistic coherence rating
n	70	70
S.D.	0,06	1,95
Mean	0,08	9,67
r	0,0752	
Significance	0,536	

Result

The result is not significant at the $p \leq 0,05$ level and the null hypothesis is therefore not rejected. The alternative hypothesis, H_2 , cannot be accepted.

Interpretation

The result of this test suggests that inclusion relations do not affect the coherence ratings of the texts of this sample. Lieber (1979; see 2.3.2) maintains that

superordinate and general items contribute least to lexical cohesion, a position which is confirmed by Halliday (1989:64 - 65) who finds that as general items are part of a class of highly frequent words, they contribute little to lexical density. On the other hand, Hartnett (1986:145) suggests that superordinates and hyponyms develop discourse logically and provide a source of logical coherence among arguments. As the essays under analysis included a narrative structure within an expository base, the ties of inclusion consisted mainly of general words and superordinate hierarchies which served to define characters and their actions within the text world rather than develop the writers' arguments logically.

(e) Derivation

Statistics	Density of derivation ties	Holistic coherence rating
n	70	70
S.D.	0,06	1,95
Mean	0,09	9,67
r	0,3767	
Significance	0,001	

Result

The result is significant at the $p \leq 0,01$ level. The null hypothesis can therefore be rejected and this part of the alternative hypothesis, H_2 , can be accepted.

Interpretation

This result suggests that there is a highly significant relationship between ties of derivation and coherence ratings in the scripts analysed. The correlation coefficient borders on the upper limits of the low category towards the moderate category according to Mulder's (1988:73; see 4.1 above) interpretation. This result is nevertheless interesting as derivation does not constitute a cohesive device in terms of Halliday and Hasan's (1976) classification scheme and has, therefore, not been included in the studies of cohesion such as those undertaken by McCulley (1985),

Neuner (1987), and others. Derivation as a subcategory of lexical cohesion has been posited by Stotsky (1981;1983) as an important feature of academic writing. She notes that "expository writing depends heavily ... on the use of affixed words for succinctness" (1981:324). The use of derivational affixes allows the writer to explore different shades of meaning and abstractness (1981:319) and to adopt a concise style of writing which enables them to include more information in shorter stretches of language. The reader is thus able to comprehend the message more quickly and efficiently and this may well contribute to the perceived coherence of the text.

Rulon Wells (1979, as cited in Stotsky, 1981:321) lists the effects of a nominalising style which is a tendency of expository writing and which is facilitated by the use of affixed words. A nominalising style creates a greater density of text and a greater average length of word. Halliday points out that there are several pressures in modern text structure towards nominalisation. One is that certain things, particularly in the register of science and technology can only be expressed in nominal constructions (Halliday, 1989:73). Another pressure towards nominalisation exists in the way the presentation of information is structured. The theme and rheme, or given and new information, usually occur in the subject and object position respectively (1989:75). Stotsky's (1986) study confirmed that the use of concrete or abstract lexical items in subject position distinguished the high-rated papers from the low-rated papers where the writers tended to use personal pronouns as the subjects of their sentences and used lexical items in places where they were not the centre of focus. The result in the present study goes some way to confirming that the use of derivational ties is one of the features of coherent writing.

4.1.1.3 H₃ Error - Coherence Hypothesis

Null hypothesis

H₀₋₃: There is no relationship between the density of lexical error in a text and the holistic coherence rating of the text.

Statistics	Error density	Holistic coherence rating
n	70	70
S.D.	0,10	1,95
Mean	0,14	9,67
r	-0,3669	
Significance	0,002	

Result

The result is significant at the $p \leq 0,01$ level. The null hypothesis can therefore be rejected and the alternative hypothesis, H_3 , can be accepted.

Interpretation

This is a negative correlation which indicates that the lower the density of errors in a text, the higher the holistic coherence rating awarded that text. The correlation coefficient lies in the low band of coefficients which appears to confirm Engber's (1995) findings that readers were only moderately affected by lexical errors. Engber's results in fact revealed a moderate negative correlation of $-.43$ between lexical error and quality score at the significance level of $p \leq 0,01$, a finding which is slightly higher than that of the present study. The significance of the results suggests that the writers are still in the process of learning active vocabulary (Engber, 1995:150). It is anticipated that students at undergraduate university level are at an intermediate level of proficiency where they are experimenting with and adding to their vocabulary. This, as confirmed by Engber's findings, may result in higher lexical variation which contributes to higher quality writing (1995:151) but may also result in the errors of word choice and form made by the writers of the texts in the present study.

4.1.1.4 H_{4a} Cohesion of Opening Paragraphs - Coherence Hypothesis

Null hypothesis

H_{0-4a} There is no relationship between the density of lexical cohesion ties between the opening paragraph of a text and its question prompt on the one hand and the holistic coherence rating of that text on the other.

Statistics	Cohesion of opening paragraph	Holistic coherence rating
n	69	70
S.D.	1,03	1,95
Mean	0,93	9,67
r	0,1067	
Significance	0,383	

Result

The result is not significant at the $p < 0,05$ level. The null hypothesis therefore cannot be rejected and the alternative hypothesis, H_{4a}, cannot be accepted.

Interpretation

The result of this test is non-significant. It appears as if the links the writers make between their opening paragraphs and the question prompt has no significant effect on the coherence ratings of their texts. This is an interesting finding given the body of literature that describes the importance of some form of introduction of topic or theme (cf. Brandt, 1986; Eiler, 1983; Stotsky, 1986) and that Bamberg's holistic coherence scale lists an opening statement which places the topic in context as one of the criteria for coherence. However, as Brandt (1986:101) has pointed out, students have only a small number of linguistic features available to them in terms of the instruction or question with which to introduce their essays. There is evidence that some student writers from the present sample used a different set of vocabulary

items to introduce the purpose of their essays. These items may have formed cohesive relations with other lexical items in the essays but did not relate to the question, as these introductory statements from one of the scripts reveals:

- [2] Animal Farm is a satiric animal fable that is filled with a lot of irony. / The purpose of Orwell's writing really, is to expose how ideals can be inverted by those who have the prerogative of exercising power.

While these statements relate minimally to the question prompt, the writer uses them to place the essay in the broader context of the literature set work and its themes. The essay thus shows an attempt at textual organisation on the part of the writer. Essays such as the one from which the example above has been taken might have achieved high coherence ratings but not necessarily high densities of ties linking the essays to the question.

4.1.1.5 H_{4b} Cohesion of Closing Paragraphs - Coherence Hypothesis

Null hypothesis

H_{0-4b} : There is no relationship between the density of lexical cohesion ties between the closing paragraph of a text and its question prompt on the one hand and the holistic coherence rating of that text on the other.

Statistics	Cohesion of closing paragraph	Holistic coherence rating
n	65	70
S.D.	0,66	1,95
Mean	0,62	9,67
r	0,3843	
Significance	0,002	

Result

The result is significant at the $p \leq 0,01$ level. This means that the null hypothesis can be rejected and the alternative hypothesis, H_{4b} , can be accepted.

Interpretation

In contrast to the correlation between ties in the opening paragraph and coherence, there appears to be a highly significant relationship between ties in the closing paragraph and coherence. Although the correlation coefficient is fairly low, there is still reason to consider this an interesting finding as this is an additional variable to the intra-textual cohesion analysed by previous research. There has been no similar research undertaken which might predict whether this would be a significant variable or not.

The finding is nevertheless difficult to explain in that the correlation between ties in the opening paragraph and coherence is not significant. Both introductions and conclusions are emphasised as part of Bamberg's (1984) Holistic Coherence Scale. An explanation for the difference in results might be found in the proximity of the paragraphs to the question. Quirk (1978:30 -31) maintains that beginnings are always continuations in that they assume common knowledge, established norms and expectations on the part of the participants. In this case, the reader would come across the essay directly below the question and would anticipate that the essay would provide an answer to the question. In other words the reader would assume coherence between the essay and the question.

However, when the writer does not round off the arguments developed in the essay and thus provide a definite sense of closure, the reader might experience a sense of incompleteness which detracts from the perception of coherence in the essay. Johns (1986:249) identifies an important part of coherence as that of "sticking to the point".

One of the features which provides this coherence is "an information structure imposed on the text to guide the reader in understanding the theme or intent of the writer" (Grabe, 1985, as cited in Johns, 1986:250). Grabe includes the use of cohesive devices as part of this information structure. In the corpus under analysis,

cohesive ties between the closing paragraph and the question might be expected to have the effect of reminding the reader of the theme of the essay and of summarising the writer's argument. Where these links are absent the reader is not provided with a sense of closure. In the present sample, two of the raters wrote that a particular script was incomplete because it had ended mid-sentence. In no other case did a rater make a written comment about a script.

4.1.1.6 Review of Coherence Hypotheses

There are four significant findings within this set of hypotheses. It appears that lexical cohesion in general, cohesive ties of derivation in particular, error density and the ties that relate the last paragraph to the question prompt all correlate with perceived coherence in a highly significant way. Although the relation between lexical cohesion and coherence is not particularly interesting given the large body of literature which has found stronger correlations between these two variables, the other correlations reveal more interesting findings.

The category of derivational relationship is a fairly unexamined one and the fact that this category correlates with both coherence and academic achievement (see 4.1.2 below) is particularly noteworthy. Bhatia (1993) and Stotsky (1983;1986) both suggest that academic discourse is marked by the tendency toward nominalisation and the use of derivations. That the more coherent texts show evidence of a significantly high density of derivational ties indicates that the writers of these texts are beginning to identify and use distinguishing features of the type of writing expected of them by the academic discourse community. These writers are able to use derivations in ties to create to signal the development of their ideas.

The relationship between the density of lexical errors and coherence confirms previous research into lexical error (cf. Engber, 1995). Research has shown that errors in other types of cohesion relation do not necessarily affect coherence ratings of student writing significantly at various levels (cf. Hubbard, 1989; Maringa, 1995) and thus it is interesting that lexical errors do at least moderately influence the

perception of coherence in a text.

The relationship between the ties of closing paragraphs and coherence is especially interesting as this is a variable which has been introduced into this study. There is therefore no research which might be used to predict what this relationship might be. It appears that the use of lexical ties relating the closing paragraph to the question prompt might act as a mechanism for summarising and concluding the writer's argument and, therefore, bring a sense of closure to the essay.

It appears that all four variables mentioned above relate to the coherence of student academic texts of this nature. Whether they would contribute to student texts of another type, written under different constraints might prove a topic for further research.

4.1.2 Academic Achievement Hypotheses

The academic achievement hypotheses have been formulated to ascertain the degree to which the academic achievement scores of the corpus are influenced by densities of lexical cohesion ties in general, each subcategory of lexical cohesion in particular and lexical error. Hypotheses have also been established to test to what extent the densities of ties of lexical cohesion relating the first and last paragraphs respectively to the question prompt contribute to the academic achievement scores of the texts under analysis. A final hypothesis tests whether there is a relationship between academic achievement scores and holistic coherence ratings.

4.1.2.1 H₀, Cohesion Density - Academic Achievement Hypothesis

Null hypothesis

H_{0.5}: There is no relationship between the density of lexical cohesion in a text and the achievement level of that text.

Statistics	Cohesion density	Achievement level (score)
n	70	70
S.D.	0,32	2,08
Mean	1,51	9,94
r	0,1692	
Significance	0,161	

Result

This result is not significant at the $p \leq 0,05$ level. The null hypothesis therefore cannot be rejected and the alternative hypothesis, H_1 , cannot be accepted.

Interpretation

This is an unexpected finding, given that the correlation between cohesion density and coherence rating is highly significant and that the correlation between coherence and achievement level is also highly significant (see below). It is perhaps explained by the different emphases of the readers of the texts. Whereas the readers rating the texts for coherence were looking solely for specific signals of coherence as delineated by the holistic coherence scale drawn up by Bamberg (cf. Bamberg 1984), the examiner was assessing whether the writers had mastered the content of the work of literature on which the question was based and whether they could present that content in the form of coherent expository essays. The examiner was more interested in the students' ability to provide accurate intertextual references to the set work studied (a background passage was also provided from which the writers could draw). The students were also expected to establish accurate exophoric references to the text world of *Animal Farm*. Thus an error such as confusing the names of the two farms (which some writers appeared to have done) would not have influenced the coherence rating of the script but might well have negatively influenced the awarding of an examination score.

4.1.2.2 H_0 Cohesion Density - Academic Achievement Hypothesis (Specific)

Null hypothesis

$H_{0.6}$: There is no relationship between the density of any category of lexical cohesion (repetition, synonymy, opposition, inclusion, derivation) in a text and the achievement level of that text.

(a) Repetition

Statistics	Density of repetition ties	Achievement level (score)
n	70	70
S.D.	0,27	2,08
Mean	1,09	9,94
r	0,0507	
Significance	0,174	

Result

The result is not significant at the $p \leq 0,05$ level. The null hypothesis cannot therefore be rejected and the alternative hypothesis, H_6 , cannot be accepted.

Interpretation

This result indicates that there is no significant relationship between repetition relations in a text and the academic achievement level of that text. While the relation between repetition and coherence narrowly misses being significant, the relation between repetition and academic achievement is clearly insignificant. This suggests that the use of repetition is of even less importance in distinguishing well written from poorly written texts in the awarding of academic scores.

(b) Synonymy

Statistics	Density of synonym ties	Achievement level (score)
n	70	70
S.D.	0,06	2,08
Mean	0,10	9,94
r	0,1642	
Significance	0,174	

Result

The result is not significant at the $p \leq 0,05$ level. The null hypothesis therefore cannot be rejected and the alternative hypothesis, H_6 , cannot be accepted.

Interpretation

There appears to be no relationship between synonymy and the achievement level of the text. Brandt (1986) has posited that when student writers write for an audience they perceive as sharing a high degree of knowledge about the topic with them, they do not necessarily use many synonyms to establish an understanding of the field about which they are writing. The corollary of this may be that an examiner marking a set of essays about a given topic with which they expect the writers to be very familiar may not look for writing that assumes that the topic needs to be explained to them. In this case the use of synonyms to define the range of the topic is irrelevant to them.

(c) Opposition

Statistics	Density of opposition ties	Coherence ratings
n	70	70
S.D.	0,06	2,08
Mean	0,14	9,94
r	-0,0374	
Significance	0,759	

Result

The result is not significant at the $p < 0,05$ level. The null hypothesis therefore cannot be rejected and the alternative hypothesis, H_0 , cannot be accepted.

Interpretation

The correlation coefficient in this case is a low, negative one. The result comparing ties of opposition to coherence also reveals a low insignificant relation. It appears as if neither the raters for coherence nor the examiner was looking for relations of opposition to indicate the successful development of a coherent topic. This unexpected finding suggests that a possible future area of research might incorporate the investigation into the effect of different topics and types of task instruction on use of particular types of cohesive tie in essays of student writing.

(d) Inclusion

Statistics	Density of inclusion ties	Achievement level (score)
n	70	70
S.D.	0,06	2,08
Mean	0,08	9,94
r	0,2328	
Significance	0,053	

Result

The result very narrowly misses being significant at the $p \leq 0,05$ level. The null hypothesis cannot therefore be rejected and the alternative hypothesis, H_a , cannot be accepted.

Interpretation

This result suggests that there is a near relation of significance between density of inclusion devices and the achievement level of any text in this sample. This is an unexpected result given that there is no significant result correlating inclusion relations with coherence ratings in the scripts. Student writers write within a particular context with a particular audience in mind. Within the text world of *Animal Farm* there are many hierarchies of animals and people that contribute to the development of ideas in the text as they express power relationships which change as the story progresses. These are expressed as part of inclusive relationships. It is possible that the examiner who is interested in the writer's expression of the mastery of the topic will find the use of inclusive devices slightly indicative of text development.

(e) Derivation

Statistics	Density of derivation ties	Coherence ratings
n	70	70
S.D.	0,06	2,08
Mean	0,09	9,94
r	0,3476	
Significance	0,003	

Result

The result is significant at the $p \leq 0,01$ level. The null hypothesis can thus be rejected and the alternative hypothesis, H_a , can be accepted.

Interpretation

Of all the specific lexical devices, only derivation has a highly significant relation to the achievement level of the texts. This confirms Stotsky's (1983; 1986) thesis that derivatives are an important feature of expository text (see 4.1.1 above). Stotsky claims that the use of derivatives to condense meaning into smaller language units contributes towards greater efficiency of communication. The fact that a text communicates efficiently should contribute to its success and might lead to its receiving a higher score.

4.1.2.3 H₇ Error - Academic Achievement Hypothesis

Null hypothesis

H_{0.7}: There is no relationship between the density of lexical errors in a text and the achievement level of that text.

Statistics	Error density	Achievement level (score)
n	70	70
S.D.	0,10	2,08
Mean	0,14	9,94
r	-0,4715	
Significance	0,000	

Result

The result is significant at the $p \leq 0,01$ level. The null hypothesis can therefore be rejected and the alternative hypothesis, H₇, can be accepted.

Interpretation

This result reveals a highly significant relationship between density of errors in a text and the achievement level of that text. The correlation coefficient reveals a moderate negative correlation between errors and score as is the correlation between errors and coherence ratings. It appears as if errors affect the

achievement level of a text more than they do the coherence of the text in that the correlation coefficient for coherence was lower. There is a body of literature that suggests that readers of texts of student writing are influenced negatively by lexical errors. Santos (1988, as cited in Engber, 1995) found that an error in lexical choice affected the meaning of the writing and affected the ability of readers to separate content and language in their evaluation of them. Sonaiya also cites research to show that errors of vocabulary are judged as being more serious than other errors. Rwigema's (1996) study showed that errors were made by writers attempting to create new words. One of the ways in which they did this was to apply derivational rules although doing so incorrectly as a result of an inadequate lexico-functional grammar. In the present study it seemed as if there were frequent errors of derivation, particularly with regard to the word *differ* which formed part of the question. As successful ties of derivation seem to contribute positively to the achievement level of a text, it may be that unsuccessful ties of derivation contribute negatively to it. However, unsuccessful ties of derivation were not examined in the study and so there are no provable grounds on which to base this assumption. In addition to this the examiner might have had reason to view errors as being detrimental to the success of a text in that this was an examination of the subject *Practical English*, one of the aims of which was to improve the students' language skills.

4.1.2.4 H_{8a} Cohesion of Opening Paragraphs - Academic Achievement Hypothesis

Null hypothesis

H_{0-8a} : There is no relationship between the density of lexical cohesion ties between the opening paragraph of a text and its question prompt on the one hand and the achievement level of that text on the other.

Statistics	Cohesion of opening paragraph	Achievement level (score)
n	69	70
S.D.	1,03	2,08
Mean	0,93	9,94
r	-0,0720	
Significance	0,557	

Result

The result is not significant at the $p \leq 0,05$ level. The null hypothesis therefore cannot be rejected and the alternative hypothesis, H_{0a} , cannot be accepted.

Interpretation

The correlation coefficient reveals a very low negative correlation between the ties of opening paragraphs and the achievement level of the text. This finding confirms the finding that there is no relationship of significance between the ties of opening paragraphs with the question prompt and the coherence ratings of the texts. In the latter case it was speculated that the use of lexical ties to link the opening paragraph to the question prompt is insignificant in distinguishing more coherent from less coherent writing because the close proximity of the question prompt makes establishing the context unnecessary. It was also suggested that students do not have a wide range of introductions available to them if they wish to align their essay with what they perceive the reader to require of them from the instructional task. These factors may also relate to the finding in this case.

4.1.2.5 H_{bb} Cohesion of Closing Paragraphs - Academic Achievement Hypothesis

Null hypothesis

H_{0-bb} : There is no relationship between the density of lexical cohesion ties between the closing paragraph of a text and its question prompt on the one hand and the achievement level of that text on the other.

Statistics	Cohesion of closing paragraph	Achievement level (score)
n	65	70
S.D.	0,66	2,08
Mean	0,62	9,94
r	0,1325	
Significance	0,293	

Result

The result is not significant at the $p \leq 0,05$ level. The null hypothesis therefore cannot be rejected and the alternative hypothesis, H_{bb} , cannot be accepted.

Interpretation

This result is not significant. This is an unexpected finding in view of the fact that the relationship between lexical ties linking the closing paragraphs with the question prompt and the coherence ratings of the scripts proved to be highly significant. Possibly this is the result of the examiner's interest in the content of the writing taking presidency over form. If the essays under examination are shown to deal with all the issues required of them from the instruction task, the examiner may not expect the writers to produce a sense of closure to their essays. The essays were written under the constraints provided by an examination and they might, therefore, not have been expected to conform to the structural constraints expected of an assignment essay.

4.1.2.6 H₀ Coherence - Academic Achievement Hypothesis

Null hypothesis

H₀₋₉: There is no relationship between the coherence of a text (hcr) and the academic achievement level of that text.

Statistics	Holistic coherence rating	Achievement level (score)
n	70	70
S.D.	1,95	2,08
Mean	9,67	9,94
r	0,6790	
Significance	0,000	

Result

The result is significant at the $p \leq 0,01$ level. The null hypothesis is therefore rejected and the alternative hypothesis, H_a, can be accepted.

Interpretation

This result shows a highly significant relationship between coherence and achievement level for the texts of the present sample. The correlation coefficient reveals a high correlation between the two variables. This confirms research done by Hubbard (1989), McCulley (1985) and Witte and Faigley (1981). Witte and Faigley (1981), who conducted their study to find out what internal characteristics distinguish high-rated texts from low-rated texts, define writing quality in terms of which they claim texts are rated as "the 'fit' of a particular text to its context, which includes such factors as the writer's purpose, the discourse medium, and the audience's knowledge and interest in the subject" (1981:199). Their study found that lexical collocation was the best indicator of writing ability because it provides cues which the reader can access via their knowledge of text conventions and understanding of the text world. It is coherence conditions that provide a "pragmatic

unity", a unity between the text and the reader's world (1981:201). They maintain that a description of writing quality must include a variety of coherence conditions both within and outside the text (1981:201).

4.1.2.7 Review of Academic Achievement Hypotheses

From the results given above, it appears that the achievement levels of student academic writing of this nature are influenced by relations of derivation and the density of lexical errors. The correlations between these variables and achievement levels are, however, at best moderate. Grobe (1981) who found that "good" narrative writing is associated with a diversity of vocabulary (a finding which is not directly confirmed by the present study), claims that the ability of the writer to compose narrative text is based on the interaction of sets of multivariate effects. The way these multivariate effects interrelate is not known at present but could if studied provide an insight into the process of composing. It may be that the interaction of various text features may provide an insight into the composition processes of a particular type of text within a particular context.

4.2 RELATIONSHIP OF RESULTS TO AIMS

The aims of the study were drawn up along three dimensions: theoretical, descriptive and applied. Only the theoretical and descriptive aims can be related directly to the statistical findings of the study. The results may be applied to the instruction of writing and the implications of the findings for this specific situation will be dealt with in Chapter 5.

The first theoretical aim was:

To determine whether and to what extent lexical cohesion contributes to the perceived coherence of a text of student academic writing.

It appears as if lexical cohesion generally contributes to the perception of coherence

of a text of student academic writing. The correlation between lexical cohesion and coherence in these texts is not as high as might be expected and the extent to which lexical cohesion contributes to coherence in this case may not be very high. In terms of specific lexical relations, only derivational ties appear to contribute to the coherence rating of these student texts. This does not seem surprising in the light of the important part affixed words and nominalisation play in academic writing style in general. The result correlating ties in the closing paragraph with the question seems to indicate that a sense of closure in a text is an important factor in the assessment of student texts as coherent.

The second theoretical aim was:

To determine whether lexical errors affect the perceived coherence of a text of student academic writing.

It appears that lexical errors do influence the perception of coherence in a text of student academic writing. The correlation coefficient indicates that this influence is only moderate to low. However, other studies on cohesion errors such as those undertaken by Hubbard (1989) and Maringa (1995) have revealed no significant relation between errors and coherence in student writing at various levels. It is thus interesting that lexical errors have some influence on coherence ratings at this level.

At the descriptive level the first aim was:

To determine whether there is a relationship between lexical cohesion in a text of student academic writing and the academic achievement level of that text.

There does not appear to be a relationship of any significance between the lexical cohesion of a text and the academic achievement level of that text for this sample. However, in terms of individual cohesion types it seems that for this particular sample the achievement level of the text was affected by inclusion and derivational relations.

The second descriptive aim was:

To determine whether there is a relationship between the lexical errors in a text of student academic writing and the academic achievement level of that text.

It appears that the achievement level of a text is moderately influenced in a negative way by the lexical errors in that text. There is a stronger correlation between the errors and achievement level than between errors and perceived coherence. There is a body of literature that supports this finding (cf. Santos, 1988, as cited in Engber, 1995; Politzer, 1978, as cited in Sonaiya, 1988), claiming that readers of student texts are intolerant of lexical errors as they can distort the meaning of a text. In this case the students were writing in response to an examination of the subject *Practical English* which focusses on improving students' language skills. The examiner was, therefore, probably more intolerant of errors than were the raters.

The third descriptive aim was:

To determine whether there is a relationship between the coherence of a text of student academic writing and the academic achievement level of that text.

This was one of the more important findings of the study. There appears to be a highly significant relationship, which strongly correlates the coherence of a text of student academic writing and the achievement level of that text. This confirms a body of previous research (cf. Hubbard, 1989; McCulley, 1985; Witte and Faigley, 1981) which defines coherence as an important aspect of writing ability. Witte and Faigley (1981) state that coherence provides the "pragmatic unity" between the text and the reader's world (1981:201) which allows the reader to understand the text (Enkvist, 1990).

4.3 CONCLUSION

The results of the tests on the hypotheses and interpretations of each were given in this chapter. The most noteworthy results were the correlations between derivational ties and error density with the perceived coherence of the texts as well as with their achievement levels and the relationship between these achievement levels and coherence. Although correlational tests can never determine a causal relationship between variables, the results of these tests have helped to increase our understanding, with regard to one set of student academic writing data, of the complex relationship between cohesion, coherence and academic achievement.

TABLE 4.1

Matrix of The Spearman Correlation Coefficients between each variable and Holistic Coherence Rating respectively, showing the correlation coefficient and the level of significance for each correlation

Significance of relationship	Variable	Holistic Coherence Rating
Highly significant positive	Lexical cohesion ties	0,3241 (0,006)**
	Derivation	0,3767 (0,001)**
	Lexical ties in closing paragraph	0,3843 (0,002)**
Highly significant negative	Lexical error	-0,3669 (0,002)**
Non-significant relationship	Repetition	0,2144 (0,075)
	Synonymy or near-synonymy	0,2122 (0,078)
	Opposition or contrast	0,0595 (0,624)
	Inclusion	0,0752 (0,536)
	Lexical ties in opening paragraph	0,1067 (0,383)

** indicates a highly significant correlation ($p \leq 0,01$)

TABLE 4.2

Matrix of The Spearman Correlation Coefficients between each variable and Academic Score respectively, showing the correlation coefficient and the level of significance for each correlation

Significance of relationship	Variable	Academic Score
Highly significant positive	Holistic coherence rating	0,6790 (0,000)**
	Derivation	0,3476 (0,003)**
Near significant positive	Inclusion	0,2328 (0,053)*
Highly significant negative	Lexical error	-0,4715 (0,000)**
Non-significant relationship	Lexical cohesion ties	0,1692 (0,161)
	Repetition	0,0507 (0,677)
	Synonymy or near-synonymy	0,1642 (0,174)
	Opposition or contrast	-0,0374 (0,759)
	Lexical ties in opening paragraph	-0,0720 (0,557)
	Lexical ties in closing paragraph	0,1325 (0,293)

* indicates a significant correlation ($p \leq 0,05$)¹

** indicates a highly significant correlation ($p \leq 0,01$)

¹ Here indicates a near significant relationship.

CHAPTER 5

CONCLUSION

5.0 INTRODUCTION

This chapter reviews the study and examines the implications of the findings for the writing classroom. It also lists the limitations of the study and in the light of these provides suggestions for future research.

5.1 REVIEW

Drawn up against a background of a large corpus of literature that confirms student difficulties in producing coherent, acceptable academic writing (Johns, 1986; Lieber, 1981; Stotsky, 1983), the purpose of the study was to establish whether there is a relationship between the lexical cohesion devices used in texts of student academic writing and both the perceived coherence and the achievement levels of these texts. The study further sought to determine the effect of the lexical errors in the texts on their coherence and achievement levels. Lastly, the study set out to determine whether a correlation between the coherence of the texts and their academic achievement levels could be established. The corpus under analysis consisted of a sample of 70 examination scripts of expository essays written in response to a question on a literary text. A small body of literature exists which suggests a relationship between textual coherence and the writing quality of the text in terms of which achievement levels are normally assessed (cf. Hubbard, 1989; McCulley, 1985). There is also a significant body of research to suggest a positive relationship between the effective use of lexical cohesion devices (usually delimited in terms of Halliday and Hasan's 1976 model) and the coherence or writing quality of a text of student writing (cf. McCulley, 1985; Neuner, 1987; Patricio, 1993; Stotsky, 1986; Wessels, 1993; Witte and Faigley, 1981). Much of this research, however, used composition scripts as data with little research being done on the kind of expository writing which forms the basis of student academic writing. For this reason, the present study tested the effect of lexical cohesion on the perceived coherence of expository scripts. Nine hypotheses were drawn up to determine the possible

contribution of lexical cohesion and lexical errors to both perceived coherence and academic achievement levels and to determine the relationship between coherence and academic achievement.

Chapter 2 provided a literature survey on the basis of which the hypotheses had been drawn up which included a brief review of the quantitative research undertaken to further examine the concepts of lexical cohesion and errors and in each case their relationship to writing quality or coherence. The quantitative research examined revealed that in general there appears to be a relationship between certain aspects of lexical cohesion such as synonymy, hyponymy and the effective use of chains of lexical cohesion devices on the one hand and measures of writing quality and coherence of a text of student writing on the other. Quantitative research into lexical errors revealed a preponderance of lexical errors in student writing as compared to errors of syntax and morphology. It also revealed a negative relation between density of errors and the perceived writing quality of texts. Findings revealed that in general readers react negatively to lexical errors in texts as they are inclined to distort or obscure meaning.

In Chapter 3 the analytical framework was outlined and the research methods used in the study delineated. The research design was a quantitative one in which the densities of lexical cohesion and lexical errors were each correlated with the holistic coherence ratings of the corpus of texts of student academic writing. The coherence ratings of the texts were then correlated with their academic scores to determine whether there is a relationship between the perceived coherence of student academic texts and their academic achievement levels. The lexical cohesion devices were categorised following Stotsky's (1983) model, itself an adaptation of the Halliday and Hasan (1976) model. Lexical errors were determined as including both word choice and form and were categorised according to Engber's (1995) model. Lexical cohesion ties were counted and tabulated (sample analyses were presented at the end of the chapter). Densities of cohesive devices and errors were determined according to the number of F-units in each text.

Chapter 4 presented the findings of the study. Because the achievement scores of the scripts showed an abnormal or skewed distribution, the Spearman-rank order correlation coefficient (a non-parametric statistical measure) was used to correlate the variables of the study. The study found that there was a highly significant relationship between the coherence ratings of the scripts and their achievement levels. This result also revealed a high rate of correlation between the two variables. Other results revealed highly significant positive relationships between ties of derivation and both coherence ratings and achievement levels. Highly significant negative relationships were found between densities of lexical error and both coherence ratings and achievement levels. Despite the significance of the two previously mentioned results, the correlation coefficient in each case revealed nothing more than a moderate correlation. However, these findings may be regarded as interesting in these instances as a high degree of correlation had not been anticipated by a large body of previous research. Of particular interest to this study was the highly significant relationship between lexical cohesion ties of the closing paragraph and the question prompt, and the coherence ratings of the texts. Examining this relationship provided an additional dimension to the study in that it involved analysing intertextual cohesion between the corpus of student writing and the examination question prompt. No previous body of research was found to anticipate the outcome of this aspect of the research. The result suggests that readers take into consideration whether a text has a sense of closure when assessing its coherence value. This might provide impetus for further research into the role played by lexical cohesion in providing textual organisation.

The following section looks at the implications of these findings for teachers of writing instruction.

5.2 APPLIED LINGUISTIC IMPLICATIONS

The applied linguistic aim of the study (see 1.1.3) was to establish whether insights gained from the research undertaken in fulfilment of the theoretical and descriptive aims might be useful in the teaching of student academic writing. In the section

below the implications of the findings of the study for the academic writing classroom will be explored and suggestions for how to teach students to write coherently for an academic audience will be given.

5.2.1 Approaches to the teaching of writing

The type of instruction which focusses on preparing students for the expository writing needed in the studying of their academic subjects may be classified as the instruction of English for Academic Purposes (EAP) which falls under the teaching of English for Specific Purposes (ESP). This is the type of teaching that might take place as part of a university course specifically designed to provide supplementary support for or development of students' academic learning, or which might be provided in the context of pre-university teaching aimed at establishing a "bridge" between previous schooling and tertiary level study.

Academic writing is of importance for students at every level and it is not surprising that there has been a range of approaches towards the teaching of it (Jordan, 1997:164). What the most effective approach to adopt might be has been an ongoing source of debate. Much writing research has centred around the needs of learners for whom English is a second language. The present study did not distinguish between first language and second language speakers of English as the aim of the study was to explore what features contributed to successful, coherent writing. However, the findings of the study are probably most relevant to second language speakers of English as those students who enrol for the subject *Practical English* are most often second language speakers who do so in order to improve their English language skills.

The various approaches to the teaching of writing have largely been influenced by trends and shifts in linguistic research particularly with regard to the study of text (Connor 1987). Raimes (1993) details the history of the four predominant approaches to writing instruction beginning with the approach which focussed on form, an approach which dominated teaching in the 1960s and 1970s. In accordance

with this approach, texts were analysed in terms of their grammatical and rhetorical structure and students were taught sentence structure and textual features such as cohesive devices amongst others. This approach is still prevalent in second language teaching and in writing research today. In the 1970s, a reaction to this approach occurred following the increasing realisation of its inadequacies in preparing students to write communicative texts. The reaction brought in the process approach to the teaching of writing, a psycholinguistic approach which focussed on the writer as "language learner and creator of text" (Raimes, 1993:239). Learners were encouraged to express their thoughts communicatively in response to personal writing topics often of their own choice. Language and form were of secondary importance and were relegated to the revision or editing stage of the writing process. Whereas the previous approach had focussed exclusively on the text as finished product, the process approach was concerned with the process undergone by the writer in producing a text. Much research accumulated around this approach, often deriving from protocol analyses (Perl, 1980; Raimes, 1985).

The process approach came to be criticised on several fronts. The fixed-stage approach to writing of "Read - Plan - Write" was found to be inadequate (Orr, 1995:192). In addition to this the choice of personalised topics and the writing of numerous drafts were criticised on the grounds that these did not prepare students for examination writing where students were assessed on their ability to produce pieces of timed essay writing on academic topics (Angelil-Carter and Eberhard Thesen, 1990; Horowitz, 1986a; Johns, 1995; Stotsky, 1995). As is argued by Stotsky, the overuse of personalised topics advocated by the expressivist approach to process writing stimulates excessive use of narrative structure (1995:765) rather than giving students experience in the writing of idea-centred writing by exposing them to a variety of genres including those of other academic disciplines (Stotsky, 1995:766).

The reaction to the perceived inadequacies of the process approach brought about two further approaches to the writing instruction of the 1980s. The content-based approach (cf. Shih, 1986) accommodated the idea that writing takes place via certain

processes but concentrated on the academic content of the writing. The reader-based approach focussed on the academic audience for whom the writing was produced. Research into this second approach has looked at the expectations of faculty members, at how readers react to writing genres and at writing across disciplines (Raimes, 1993:242).

Although the reader-based approach to the teaching of writing includes aspects of the process approach, at the same time there is a renewed emphasis on the finished product and on form. This time, however, the focus is on both linguistic and rhetorical forms of academic writing rather than predominantly on grammatical form of text in general (Raimes, 1993:242).

All of the approaches listed above are in use today and as Raimes points out, the teacher of writing now has a variety of approaches to choose from as opposed to the 1960s when only the product approach had been researched and was being advocated (1993:242). Many aspects of all the approaches mentioned above are useful in the teaching of academic writing. White points out that:

“... much EAP writing is product-oriented, since the conventions governing the organization and expression of ideas are very tight. Thus the learner has to become thoroughly familiarized with these conventions and must learn to operate within them ...”

(White, 1988a, as cited in Jordan, 1997:168)

At the same time he points out that academic writing primarily involves the manipulation of ideas, a feat best achieved through activities of the writing process (cf. Davies, 1988:130).

The present study falls within the product approach in terms of its focus on specific features of text. As the study occurs within the context of student academic writing, it seems most applicable to lodge the findings of the study within the reader-based approach, in particular within the genre-based approach.

The genre-based approach derives from many theoretical orientations but most

obviously from Halliday's register theory, a precursor to the theory of discourse analysis. There are several proponents of the genre-based approach who share theoretically similar principles but have differences in their applications of the approach at the level of classroom practice (Bloor, 1998:55). Swales (1990) and Bhatia (1993) unite in defining genre in terms of communicative purpose. Bhatia sums up each genre as "the successful achievement of a specific communicative purpose using conventionalized knowledge of linguistic and discoursal resources" (Bhatia, 1993:16). For Swales and Bhatia specific genres are the product of the discourse community which uses them. The academic discourse community has specific conventions for the genres it uses such as the research article and even student academic writing.

The Australian school has applied this approach at many academic levels (Johns, 1995). Within this school, there are several definitions of genre and several approaches to teaching it. However, the proponents of it do seem to agree on the essentials of the approach and genre is generally defined by this school as "repeated social action" (Miller, 1984:151, as cited in Johns, 1995:186). This form of the genre-based approach has led to the "Critical Language Teaching" approach which is based on Critical Discourse Analysis, the focus of which is to examine critically "... the reproduction of sexism and racism through discourse; the legitimation of power; the manufacture of consent; the role of politics, education, and the media; the discursive reproduction of dominance between groups; the imbalances in international communication and information ... " (Kress, 1991, as cited in Kaplan, 1991:202).

Johns (1995), who finds the applications of both Swales (1990) and the Australian school inappropriate for an academic writing class at tertiary level, develops her own application of the genre-based approach. She sets out to teach students how to analyse genres beginning with those with which they are familiar from their own communities and then moving on to academic genres. She maintains that students need to know the following facts about texts:

1. No text is autonomous; instead, specific, situated texts result, in part, from the many discourses and texts which precede them. The issue of intertextuality (Widdowson, 1993) is one that we need to address directly and repeatedly.
2. The features of texts and task are determined by many rhetorical factors in addition to the intertextual: task, writer's purposes and experiences, context, community values and traditions, readers' and writers' roles, as well as other factors.
3. Some text elements are repeated in a text exemplar of a genre as particular rhetorical features listed in Item 2 [above] reoccur. Thus, a genre is referred to as a "repeated social action" (Miller, 1984, p. 151).
4. Approaches to task and text processes depend on the context and community as well as the readers' and writers' roles and past experiences with texts and tasks.
5. Students have rich sociocultural backgrounds to bring to texts and tasks, including their prior experiences with genres within the communities in which they live and work.

(Johns, 1995:186)

These aims applied to the academic writing classroom allow the students to learn the sociocultural factors involved in the production of text which may be useful in enabling them to learn to write within a specific academic context.

5.2.2. Implications for writing instruction of the findings of the study

An important finding in the present study was the highly significant, strongly correlated relationship between coherence and academic achievement levels in the texts under analysis. The strength of this finding suggests that an important focus for writing instruction should be the teaching of coherence. This might be considered a difficult task given that coherence "embodies a large number of variables" (Enkvist, 1990, as cited in Connor and Johns, 1990:2). However, other findings in the present study indicated that apart from lexical cohesion as a whole, certain features of lexical relations such as derivational relations are significantly related to coherence in

student academic texts. These are, therefore, features which might profitably be taught to students of academic writing to further their ability to produce coherent texts.

Teaching coherence from a genre-based perspective requires one to work within the constraints of a genre-based syllabus, which is defined by Davies (1988:132) as one "which is founded on the identification and analysis of the specific genres the students are required to read and write in their subject-based studies". This study was primarily concerned with the investigation of student writing. However, it has been shown that reading plays an important part both in the writing process itself, for example in the revising stage where writers need to read their own work critically (Stotsky, 1984:6), and in the ability to analyse specific text types. Therefore, reading will be incorporated into the discussion of the teaching of coherence below.

In terms of the genre-based syllabus, reading is an essential skill required by students in order to analyse texts of appropriate academic genres for coherence patterns which they can then adopt as part of their own writing strategies (Davies, 1988:138). Thus appropriate genres to be analysed should be identified in terms of the types of writing students are required to produce. Orr (1995:194) maintains that the types of writing students come across in the course of their studies, such as tutorial letters, textbooks and scholarly articles, give students no indication of the type of prose expected of them in the academic essay. She argues that students should rather be exposed to prose models of texts used in their particular subject disciplines which are produced by other student writers and which can be analysed for "features such as the structure, style, register, coherence devices, academic etiquette, and vocabulary characteristic of and valorised by the discipline" (1995:193). Advantages of this procedure are firstly that it helps students improve their writing strategies by giving them access to discourse patterns which reduce "some planning burdens on short-term memory" (Daiute 1984:222, as cited in Orr, 1995:195), and secondly that it improves their ability to analyse text as they become accustomed to recognising various schemata (Reid, 1984, as cited in Orr, 1995). Johns (1986) advocates that students in an academic writing class are given

samples of student writing to analyse as an exercise to promote the recognition and acquisition of useful writing strategies.

Texts are processed and understood at different levels. Readers work from the bottom-up to establish the meanings of the words and sentences they come across in the text. They also work from the top-down using their background knowledge of the world and what they expect from the text in terms of its genre and context to predict what the next sentence will be (Brown and Yule, 1983:234; Johns, 1986:250).

The student writer needs to be aware of both of these processing strategies when writing so that they can guide the reader through the text in their attempt to produce "reader considerate" text (Armbruster and Anderson, 1984, as cited in Johns, 1986:251).

An important aspect of the top-down strategies which student writers need to learn, therefore, is that of the reader or audience. The present study showed that coherence ratings and academic achievement levels correlated in a strongly significant way which suggests that coherent texts go some way towards satisfying the expectations of the examiner who awards academic scores to the texts. This seems to confirm that coherent writing takes into account the expectations of the reader or audience. Johns (1986:250) maintains that a reader's understanding of a text and its underlying structure and, therefore, its coherence value depends on whether the reader's expectations, which are based on background knowledge of the content and structure of the type of writing presented, are consistent with the text. In the case of student academic writing, the reader usually has more knowledge of both the content and structure of the writing than the student writer does. Johns (1993:85) argues that students in academic writing classes need to understand " ... the interaction between their purposes, the interests and values of real audiences, and the genres that are appropriate for specific rhetorical contexts". Many student writers initially conceptualise the audience as a vague entity but seem better able to define their reader as they progress through the stages of composing (Ede and Lunsford, 1984, as cited in Johns, 1993:85). To facilitate student sensitivity to the expectations of their audience, Johns (1993) advocates that academic writing

teachers interview expert members of specific discourse communities about their expectations of student writing and about conventions expected of the important genres used in their communities. The teacher should bring back insights gained from this research to the classroom in order to teach students explicitly how to write for their specific audience.

Academic writing students also need to learn how to develop and structure their ideas in their writing. Flower (1979) maintains that transforming writer-based prose into reader-based prose requires students to transform their writing into "a rhetorical structure built on the logical and hierarchical relationships between ideas and organised around the purpose for writing, rather than the writer's process" (Flower, 1979:37). Davies (1988) suggests that students should be taught the top-down strategy of identifying the functions of the different sections of relevant text types such as examination essays. She argues particularly for the analysis of "end" sections (1988:139) which embody the discussion or conclusion of the essay. Students should be taught to analyse conclusions to determine whether they follow on from the introductions of the essays in terms of whether they fulfill the aims set out in the introductions.

Following the bottom-up approach, students need to learn how to provide the lexical signals required to indicate the functions of the conclusions to their essays. The present study revealed that the density of lexical ties between the closing paragraphs of the student essays and the question prompt served to distinguish the more coherent writing from the less coherent writing. This suggests that lexical relations are an important mechanism for signalling the functions of summary and closure provided by the conclusion. In this regard, students should be encouraged to analyse the question prompt before beginning to write and to refer back to the question as they develop their writing. Students who are aware of the specific demands of the question are more likely to be aware of the need to focus on the topic. Davies (1988) maintains that students' lexical choices can serve to develop topics and mark the boundaries between them. If students are encouraged to provide their audience with a sense of closure by referring lexically to the question

prompt, they are more likely to build up their arguments towards a conclusion.

The study also revealed that the density of lexical cohesion ties generally had a significant relation to perceived coherence of student expository writing. This suggests that skilled writers are better able to signal semantic relations in their texts than are less skilled writers. Stotsky (1983:440) argues that the use of lexical ties to create text-forming relationships is derived from underlying processes such as classifying and comparing and contrasting which do not need to be taught. Learners do need to be taught how to signal these relationships, however, and to do this they need to develop a wide range of vocabulary and a knowledge of how these items relate to each other.

An interesting approach to the teaching of lexical relations is encompassed in the writing instruction book *Think write* (Rodseth, Johanson and Rodseth, 1992) which, as part of teaching the writing process, focusses on how ideas are connected and how these connections are signalled in terms of lexical relations. This book can be of use to students at higher education level as it deals with the analysis and production of informative passages rather than literary or narrative texts.

In terms of the specific findings of the study, it appears that in the scripts analysed, density of repetition ties, although high, did not contribute to the perception of coherence. Previous studies such as those of Stotsky (1986) and Witte and Faigley (1981) suggest that although repetition ties are frequently used by both skilled and unskilled writers alike, these ties are often used redundantly by poorer writers. This suggests a need for students to be taught effective ways of using repetition to focus on and develop the topic and to provide signals of emphasis (Rodseth *et al.*, 1992:42 - 43). *Think write* teaches students how to construct paragraphs by linking each supporting sentence to the previous sentence and ultimately to the topic sentence using repetition of key words. At the same time each sentence develops the topic by providing new information about it. Wessels (1993;1994) links the idea of lexical repetition to that of conceptual repetition where weaker writers merely repeat ideas. A technique such as that mentioned above will prevent students from constantly

repeating old information and encourage them to develop their ideas by introducing new information into each consecutive sentence.

The study also showed that writers did not use many ties of synonymy or opposition to compare and contrast the conditions on 'Animal Farm' with those on 'Manor Farm'. Students need to be taught how effective use of synonyms and opposites can contribute to successful writing. Rodseth *et al.* (1992) alert students to the effective use of synonyms and antonyms in passages of informative writing and suggest ways of acquiring these types of words through the use of a thesaurus (cf. McCarthy, 1990).

The study found that a near significant relationship between the density of inclusion ties and academic achievement levels in the scripts. It was suggested in explanation of this that successful writers probably signalled the hierarchical relationships which exist within the text world of *Animal Farm* thus contributing to the development of their ideas. *Think write* (1992) advocates moving from general to specific information as a way of developing the topic and suggests activities to teach inclusion hierarchies of superordinate and subordinate items. These are useful techniques to teach students how to structure lexical relationships in academic discourse.

Strategies and exercises for learning and practising lexical relationships in discourse generally have been proposed by authors such as Gairns and Redman (1986) and McCarthy (1984; 1990) who advocates the use of specially constructed discourse topics and activities which require the use of specific types of lexical ties (1984:18). Crombie (1985: 91 as cited in Carter 1987: 175) also advocates encouraging learners "to develop a sensitivity to the ways in which the various relationships between lexical items may themselves contribute to semantic relational identification". An approach to lexical development based on the idea of semantic fields and collocational relationships has been developed by Rudzka, Channell, Putseys and Ostyn (cf. *The Words You Need* by Rudzka *et al.*, 1981). It seems that there is general acknowledgement amongst vocabulary instructors of the need to teach lexical items in terms of the relationships they form in discourse.

Developing a knowledge of how to use derivational devices in student academic writing seems of particular importance given the highly significant relationship between density of derivational ties and coherence ratings for the corpus. The use of affixed words in academic writing is well documented (Bhatia, 1993; Stotsky, 1983). Students need to learn how to use derivational ties to develop and extend the academic concepts that form the themes of their writing. Stotsky (1979) advocates teaching the functions of derivations to students, taking as a starting point the examination of students' own writing for potential derivational elements. She suggests that students should be taught the meaning and use of prefixes and suffixes but cautions that this should be done accurately.

The findings of the present study revealed that lexical errors had a low to moderate influence on the perceived coherence of texts but a moderate one on the achievement level of the texts. This suggests that accuracy of word choice and form is of at least some importance in text production. Operating from within the paradigm of the lexical syllabus which places a stronger emphasis on lexis than did traditional language learning syllabi, Willis (1993) suggests that accuracy should relate to the relationship between form and meaning of lexical items. In order to encode meanings accurately, learners need to understand that language is a resource for meaning which entails understanding how words combine to create meaning in relations of various language units. It appears that learning how lexical items form relationships in discourse is part of learning how to use them accurately.

5.3 LIMITATIONS AND SUGGESTIONS

A notable limitation of the present study was that the data analysed proved to be negatively skewed, possibly a result of limiting the texts to be analysed in terms of minimum length. The fact that the sample was skewed meant that it could only be tested using a non-parametric statistical equivalent to Pearson's product-moment correlation, in this case the Spearman-rank order correlation coefficient, which is less powerful than the parametric test would be.

Correlational tests cannot be used to show relationships of causality between variables. They can only show the degree of relationship between variables and correlational results are usually interpreted cautiously. In addition to this, skewed distributions usually have the effect of depressing correlational coefficients obtained (see 4.1 above). The results realized by the study were therefore possibly less significant than if the sample had had a normal distribution of scores.

In the process of rating the scripts for holistic coherence, the raters were expected to assign ratings according to the categories of Bamberg's scale without allowing for the option of assigning half marks to the essays. This was a limitation in the data collection process which became a weakness of the study in terms of assigning scripts to a particular band on the scale. There were many scripts assigned to bands 3 and 4. Allowing raters to assign half marks to the scripts may have enabled raters to distinguish more precisely between scripts and this could have resulted in a wider range of ratings being awarded. However, as the achievement scores awarded to the scripts were generally high, it may be that the coherence ratings accorded the scripts reflected the perception on the part of the raters that the scripts tended towards being coherent.

In terms of delimiting the scope of the study, the lexical cohesion subcategories used in the study were restricted to the subcategories from Stotsky's (1983) model of lexical cohesion, adapted from Halliday and Hasan's (1976) model in order to create a classificatory scheme better suited to the study of expository writing. The results justified using a research design based on the counting of specific types of lexical cohesion in that specific categories of lexical cohesion as defined by Stotsky (1983) were revealed to be important features of coherent student academic writing in this instance. A possibility for further research might be to explore the use of chains of lexical cohesion in student expository writing using Stotsky's (1983) lexical cohesion classifications. Another research possibility might be to explore the function of cohesive ties by examining Hartnett's (1986; see 2.2 above) subcategories of static and dynamic ties and establishing the role these subcategories play in accounting for perceived coherence in student academic writing.

Only Stotsky's category of reiteration was chosen for examination in the study. The category of collocation, delimited by Stotsky to include only those lexical items which randomly co-occur in similar texts, was excluded from the study. This was on the grounds that it would be difficult to prevent inter-rater subjectivity in identifying collocation (Hasan, 1984) in an empirical study such as this one. In some research collocation has been established as an important predictor of coherence and writing quality in student writing (McCulley, 1985; Witte and Faigley, 1981). The study of collocation presents further research possibilities, although it might benefit from a qualitative approach such as a research design incorporating surveys and interviews to assess the reactions of academic readers to the use of collocational relations in student academic texts. A corollary to this might be to determine student ability to comprehend academic texts in terms of the collocational relationships expressed in these texts. Research in this direction includes the study of the contextual support required by students in terms of academic and technical vocabulary present in their academic study texts (Hubbard, 1996).

The results of the present study showed that while the relationship between cohesive devices generally and holistic coherence was highly significant, in terms of statistical probabilities, the degree of correlation between the two variables was low. This suggests that there are other determinants of coherence unaccounted for by the study. An interesting feature related to coherence which was not explored in the present study is that of reader inference-making. A possible area for future research might include the ability of student writers to make and interpret inferences where necessary in academic discourse in their interaction with the text both as producers and receivers. Empirical research on the role of inference in interpreting naturally occurring texts should be explored still further.

Another feature of coherence not given central place in the present study is that of textual organisation in terms of topic development. Wessels (1993), on the basis of her research findings, suggested that further research might benefit from exploring the role of lexical cohesion in topic-opening and topic-closing sentences. The present study went some way toward examining this in establishing a significant

relationship between the cohesion between lexical items in the closing paragraph and the question prompt, and the perception of coherence in texts of student academic writing. This suggests that a sense of closure has a role to play in determining holistic coherence in the genre of student academic writing. Further possibilities for research might include determining student ability to provide textual structure and organisation in their writing in terms of lexical signalling. Establishing specifically how student writers structure text to develop topics (cf. McCutcheon and Perfetti, 1982; Stotsky, 1986; Tierney and Mosenthal, 1983) may be a useful direction for this type of research to take. Tierney and Mosenthal (1983) maintain that the presence of cohesive ties in a text is a result of a coherent presentation of the argument of the text (1983:228). It would be interesting to examine how good writers signal the development of their arguments using lexical ties. Further assessment might concentrate on how students cope with structuring their texts in accordance with the prescriptions presented by the other writing modes they come across in their fields of study. The data in this study consisted of texts incorporating narrative style within an expository mode which required them to provide comparison and contrast between two states. Other common modes expected of students in producing assignment and examination essays include those of synthesis and analysis (Orr, 1995:193). Further research possibilities might explore student ability to provide appropriate lexical signals of text cohesion and topic development while operating within the constraints provided by these discourse modes.

The present study looked at scripts of students at university level. Other literature (for example, McCutcheon and Perfetti, 1982; Yde and Spoelders, 1985) has compared groups of children of different ages to determine whether a developmental trend exists in children's ability to use cohesive ties. In both the studies mentioned above, results showed that older children had a greater ability to structure their writing cohesively. Stotsky (1986:289) speculates that poor writers and younger writers write substantially less than do good writers and older writers respectively partly because they have smaller vocabularies and because they are unable to control their expository writing using different types of lexical cohesion. She suggests that longitudinal case studies of developing writers be carried out to test the

hypotheses that:

"... (1) the more concepts a writer can form to make assertions about, the more interconnections among these concepts there should be; (2) the larger the number of different words used in an essay, the more lexical cohesive ties there should be; and (3) the more this network of cohesive ties develops, both within and across sentences, with increasingly longer semantic units entering into many different types of ties spanning all portions of a text, the better organized and developed the essay should be."

(Stotsky, 1986:289)

It appears that the study of lexical cohesion in student writing might benefit from developmental studies comparing different groups of students such as those at intermediate and advanced levels of tertiary study or from longitudinal studies such as that suggested by Stotsky (1986).

Apart from carrying out developmental studies to determine growth of ability of student writers to produce semantic networks of lexical cohesion, comparative studies could be done to assess the development of lexical proficiency in terms of accuracy of word choice and form. Engber (1995) suggests that the intermediate level is a useful level to study in that students at this stage produce enough material to make the testing of hypotheses possible. She cites evidence from a study undertaken by Laufer (1991, as cited in Engber, 1995:151) which suggests that advanced learners stop acquiring vocabulary when they reach the average proficiency level of their group. At this stage they begin to hone their writing to make it more precise rather than actively accumulating more vocabulary. A comparison of two groups at different levels might provide insights into how the writing process develops.

An area which has not been explored in this study but which appears to hold promise for future research is that of the effect of the distance between ties on coherence in student academic texts. Halliday and Hasan (1976) propose two taxonomies of cohesive ties. The first taxonomy relates to the function of the types of cohesive tie and the second relates to the distance between them (Witte and Faigley, 1981:194). Halliday and Hasan (1976:330 - 332) posit four classes in terms of which the

distance between elements of a cohesive tie can be measured. Elements of a tie can be 'immediate', to each other. They can also be 'mediated', 'remote' or at a 'mediated-remote' distance from one another. As part of their 1981 study, Witte and Faigley used these measures to test whether the distance between ties of student writing had any effect on the writing quality produced. Their findings revealed that better writers used a high percentage of immediate ties, an indication that they were better able to establish stronger bonds between their T-units and extend the concepts they used. The weaker writers, on the other hand, seemed to interrupt many of their ties which showed that they did not elaborate their ideas through successive T-units. An interesting proposition for further research might be to extend Witte and Faigley's (1981) research to look at the performance of students writing academic texts in this regard.

5.4 CONCLUSION

This study has primarily attempted to explore the relationship between lexical cohesion and coherence in student academic writing. Although there are several limitations to the study and many areas which still need to be researched, it is felt that the study was in some way able to fulfill its central aim of beginning to explicate the concept of coherence within the context of student academic writing. As Brandt (1986:106) states:

"Techniques of discourse analysis which relate text to context and parts of text to each other can aid ... teachers in reading students' papers more fully for the interpretative processes which underlie them."

APPENDICES

APPENDIX A: EXAM QUESTION

2. Animal Farm

Read the following extract from *Animal Farm* and then answer the two questions that follow:

The farm was even more prosperous now, and better organized: it had even been enlarged by two fields which had been bought from Mr Pilkington. The windmill had been successfully completed at last, and the farm possessed a threshing machine and a hay elevator of its own, and various new buildings had been added to it. Whymper had bought himself a dogcart. The windmill, however, had not after all been used for generating electrical power. It was used for milling corn, and brought in a handsome profit. The animals were hard at work building yet another windmill; when that one was finished, so it was said, the dynamos would be installed. But the luxuries of which Snowball had once taught the animals to dream, the stalls with electric light and hot and cold water, and the three day week, were no longer talked about. Napoleon had denounced such ideas as contrary to the spirit of Animalism. The truest happiness, he said, lay in working hard and living frugally.

Somehow it seemed as though the farm had grown richer without making the animals themselves any richer - except, of course, for the pigs and the dogs. Perhaps this was partly because there were so many pigs and so many dogs. It was not that these creatures did not work, after their fashion. There was, as Squealer was never tired of explaining, endless work in the supervision and organization of the farm. Much of this work was of a kind that the other animals were too ignorant to understand. For example, Squealer told them that the pigs had to expend enormous labours every day upon mysterious things called "files", "reports", "minutes", and "memoranda". These were large sheets of paper which had to be closely covered with writing, and as soon as

they were so covered, they were burnt in the furnace. This was of the highest importance for the welfare of the farm, Squealer said. But still, neither pigs nor dogs produced any food by their own labour; and there were many of them, and their appetites were always good.

- (a) In what ways do conditions on Animal Farm differ from those on Manor Farm?
(15)

APPENDIX B: BAMBERG'S (1984) HOLISTIC COHERENCE SCALE

4 = Fully Coherent

Writer clearly identifies the topic

Writer does not shift topics or digress

Writer orients the reader by creating a context or situation

Writer organizes details according to a discernible plan that is sustained throughout the essay

Writer skilfully uses cohesive ties such as lexical cohesion, conjunction, reference, etc. to link sentences and/or paragraphs together

Writer often concludes with a statement that gives the reader a definite sense of closure

Discourse flows smoothly - few or no grammatical and/or mechanical errors interrupt the reading process

3 = Partially Coherent

If writer does not explicitly identify the topic, s/he provides enough details so that reader can probably identify the specific subject

Writer has one main topic but there may be minor digressions

Writer provides some reader orientation, either by briefly suggesting the context or by directly announcing the topic

Writer organizes details according to a plan, but may not sustain it throughout or may list details in parts of the essay

Writer uses some cohesive ties such as lexical cohesion, conjunction, reference, etc., to link sentences and/ or paragraphs together

Writer does not usually conclude with a statement that creates a sense of closure

Discourse generally flows smoothly although occasional grammatical and/or mechanical errors may interrupt the reading process

2 = Incoherent

Some of the following prevent the reader from integrating the text into a coherent whole:

Writer does not identify the topic and the reader would be unlikely to infer or guess the topic from the details provided

Writer shifts topics or digresses frequently from the topic

Writer assumes the reader shares his/her context and provides little or no orientation

Writer has no organizational plan in most of the text and frequently relies on listing

Writer uses few cohesive ties such as lexical cohesion, conjunction, reference, etc., to link sentences and/or paragraphs together

Writer creates no sense of closure

Discourse flow is irregular or rough because mechanical and/or grammatical errors frequently interrupt the reading process

1 = Incomprehensible

Many of the following prevent the reader from making sense of the text:

Topic cannot be identified

Writer moves from topic to topic by association or digresses frequently

Writer assumes the reader shares his/her context and provides no orientation

Writer has no organizational plan and either lists or follows an associative order

Writer uses very few cohesive ties such as lexical cohesion, conjunction, reference, etc. and sentences do not seem connected or linked together

Discourse flow is very rough or irregular because writer omits structure words, inflectional endings and/or makes numerous grammatical and mechanical errors that continuously interrupt the reading process

0 = Unscorable / Miscellaneous

Essay consists of only one T-Unit

Writer writes only to reject the task

APPENDIX C: LIEBER'S (1979) CATEGORIES OF F-UNIT

Categories of F-unit (cf. Lieber, 1979: 93 - 95) used in the study with examples from the scripts analysed in the present study.

Coordinating Structures:

I "Full clauses joined by coordinating conjunctions or marks of punctuation constitute separate F-units... ."

[1] On the whole the working hours were longer on Animal Farm / and the rations were less than they were on Manor Farm.

[2] There will be a Rebellion / and the animals will thereafter rule themselves.

[3] The farm is more prosperous / and it is better organized.

II "Clauses exhibiting gapping in a non-initial member constitute separate F-units... ."

[4] Their rations are no more than before / and their living quarters no better.

III "Clauses containing conjoined verbal structures will be segmented into more than one F-unit"

[5] The pigs and dogs now eat different kind of food / and get even more than other animals.

[6] Pigs also used the farm house for them to sleep / and actually instruct the others to do work.

IV “Conjoined nonverbal elements within a clause will be segmented into separate F-units when an overt marker indicating a change in rhetorical function is present (i.e. *but*, *except*, or an adverbial marker or prepositional phrase)”

[7] The windmill was not used for generating electrical power, / but for milling corn.

[8] Actually, it was not the animals who changed the things, / but the pigs.

Subordinate Structures:

1. “Adverbial subordinate clauses and clause equivalents”

(Here locative and temporal clauses introduced by *where* and *when* respectively were included - cf. Hubbard, 1989: 120.)

[9] Orwell uses this contrast / to expose and ridicule the vices and follies inherent in revolutionary rule.

[10] There are meetings on Sundays / where the animals get orders from the leadership of the pigs.

[11] When work on the windmill starts, / it is back-breaking work.

[12] Whereas before they worked for Jones / now they work for the pigs.

2. “Nonrestrictive relative clauses and sentence relatives”

[13] Compared to Animal Farm, Manor Farm is neglected by Mr Jones / who spends his money on wars.

[14] The animals have successfully completed the windmill / which is used

as a threshing machine.

3. "Nonrestrictive appositives (i.e., reduced nonrestrictive relative clauses)"

[15] During the time of Snowball // the brain behind the windmill // ...

[16] Old Major // an old pig in the farm // organizes the Rebellion of animals towards Man.

4. "Nonrestrictive appositives of exemplification, identification, and renaming"

[17] According to Major the vices of man // such as: sleeping in a bed, wearing clothes and drinking alcohol // should be done away with.

5. "Absolute constructions related to adverbial clauses or nonrestrictive relatives... ."

[18] The pigs and dogs somehow always got away with not doing any work on the farm. / The excuse being that they do all the brain work.

APPENDIX D: SAMPLE SCRIPTS AND TABLES

Sample scripts detailing the analysis of cohesive ties of the whole text, the relations between the first and last paragraphs of the text and the question prompt, and errors.

Key to tables:

Heading: Type = This refers to the type of cohesive tie categorised according to Stotsky as:

- (R) = Repetition
- (S) = Synonymy
- (O) = Opposition
- (I) = Inclusion
- (D) = Derivation

A superscript w indicates within the F-unit (i.e. a structural tie).

A superscript k indicates a cataphoric tie.

The number in brackets after the presupposed item refers to the F-unit number of the item, for example: *time* (14) refers to the fact that the lexical item *time* is to be found in F-unit number 14.

Script 58 HCR = 5

Question:

In what ways do conditions on Animal Farm differ from those on Manor Farm?

(1) The farm is more prosperous (2) and it is better organized. (3) The plight of animals has not changed very much. (4) The farm has more fields (5) which were bought from Mr Pilkington. (6) The farm is also not run by Mr Jones (7) but by the

pigs, (8) the head being Napoleon. (9) The place has become more mechanised; (10) the farm has a threshing machine, hay elevator and a windmill. (11) Mr. Jones did not possess any machinery. (12) His methods of farming were less scientific. (13) He and his men were usually drunk (14) and did not feed the animals on time. (15) The animals still get the same or even **less** rations than before (16) and they work hard. (17) The pigs and dogs live better than when the farm was Manor farm. (18) The farm is expanding all the time (19) whereas when it was Manor farm (20) it was being neglected. (21) The pigs are more educated and intelligent **compared to** Mr Jones. (22) The animals seem to be working harder (23) and maybe are given even less food. (24) They do not eat when they are not supposed (25) so no barriers need to be put up.

Cohesive ties of whole script:

F-unit no.	No. of ties	Cohesive item	Type	Presupposed item	No. of words
4	1	(The) farm	(R)	(The) farm (1)	2
6	1	(The) farm	(R)	(The) farm (1)	2
7	1	(the) pigs	(O)	Mr Jones (6)	3
8	1	Napoleon	(I)	(the) pigs (7)	2
9	1	(The) place	(I)	(The) farm (1)	2
10	1	(the) farm	(R)	(The) farm (1)	2
11	3	Mr Jones	(R)	Mr Jones (6)	4
		possess	(S)	has (10)	2
		machinery	(I)	(a) threshing machine, hay elevator and a windmill (10)	8
12	1	farming	(D)	(the) farm (10)	2

F-unit no.	No. of ties	Cohesive item	Type	Presupposed item	No. of words
14	1	(the) animals	(R)	animals (3)	2
15	1	(the) animals	(R)	animals (3)	2
17	4	(The) pigs	(R)	(the) pigs (7)	2
		better	(R)	better (2)	2
		(the) farm	(R)	(The) farm (1)	2
		(The) pigs and dogs	(O)	(The) animals (15)	4
18	2	(The) farm	(R)	(The) farm (1)	2
		(the) time	(R)	time (14)	2
19	1	Manor farm	(R)	Manor farm (17)	4
21	2	(The) pigs	(R)	(the) pigs (7)	2
		Mr Jones	(O) ^w	(The) pigs (21)	3
22	2	(The) animals	(O)	(The) pigs (21)	2
		working harder	(R)	work hard (16)	4
23	1	food	(S)	rations (15)	2

Script 58 could not be analysed for ties relating the first and last paragraphs to the question prompt as there was nothing in the structure or lexis of the essay that referred back to the question prompt.

Script 17 HCR = 7

(1) Animal farm is an imaginary story. (2) Here the animals try to free themselves from the tyranny of the human beings. (3) They were looking to govern themselves. (4) They overthrow the human (5) being and was ruling themselves. (6) The

animals is tired of being **explored** by Jones. (7) They work hard (8) i.e. eggs, the hens. (9) Major has this dream of overthrowing the human and look for **selfgoverning**. (10) This is done so that animals will live together in comradeship (11) and will enjoy the **fruits** and profits of the farm. (12) The seven "commandments" and "Beast of England" is accepted by all the animals.

(13) At Battle of Cowshed all animals fought bravely (14) to overthrow human beings. (15) In reality condition on the two farms is more or less the same. (16) At the battle of Cowshed, Snowball is the hero (17) and fight bravely (18) but he is expelled from the farm (19) because he is regarded as a traitor and on the side of Jones.

(20) The pigs are in charge of the farm (21) because they are the most intelligent animals. (22) The pigs are taught (23) and **educated** (24) so that they can be the rulers on the farm. (25) The other comrades are regarded as slaves (26) and must do all the work.

(27) The pigs move in the farm house, (28) enjoy the luxury of sleeping on beds, (29) while the other animals work on the farm. (30) The pigs enjoy eating the apples and drinking milk (31) and when other animals question Squealer (32) he justifies that the pigs must have this (33) because they are working very hard on the farm.

(34) Animals are **rationed** (35) and do not have enough food to eat, (36) Squealer justifies (37) and says that the **ration** is very high. (38) **Trade** business is established with the neighbouring farmers. (39) The hens are asked to do sacrifices (40) and when they **receive** (41) they are killed. (42) Animals have to confess and **executed**. (43) Napoleon banned the song "Beast of England" (44) and the Seven Commandments are altered to **soothe** the pigs. (45) When the animals question Squealer (46) he uses language (47) to **manipulate** the animals. (48) Conditions on the farm **deteriorated** (49) but comrades were informed that things **were** better than before.

(50) Napoleon has 9 puppies to protect himself from the other animals. (51) Boxer is sold (52) after working hard (53) when they were building the windmills.

Cohesive ties in the whole script:

F-unit no.	No of ties	Cohesive item	Type	Presupposed Item	No. Of words
4	1	human being	(R)	human beings (2)	4
6	1	(the) animals	(R)	(the) animals (2)	2
9	2	overthrowing	(R)	overthrow (4)	2
		(the) human	(R)	human beings (2)	3
10	1	animals	(R)	(the) animals (2)	2
12	1	(all the) animals	(R)	(the) animals (2)	2
13	1	(all) animals	(R)	(the) animals (2)	2
14	2	overthrow	(R)	overthrow (4)	2
		human beings	(R)	human beings (2)	4
15	1	(the two) farm	(R)	(the) farm (11)	2
16	1	At the Battle of Cowshed	(R)	At Battle of Cowshed (13)	9
17	1	fight bravely	(R)	fought bravely (13)	4
18	1	(the) farm	(R)	(the) farm (11)	2
19	2	traitor	(O)	hero (16)	2
		Jones	(R)	Jones (16)	2
20	1	(the) farm	(R)	(the) farm (11)	2
21	1	(the) most intelligent animals	(I)	(The) pigs (20)	4
22	1	(the) pigs	(R)	(the) pigs (20)	2

F-unit no.	No of ties	Cohesive item	Type	Presupposed Item	No. Of words
24	2	rulers	(D)	ruling (5)	2
		(the) farm	(R)	(the) farm (11)	2
25	3	(the) other comrades	(O)	(The) pigs (22)	3
		comrades	(D)	comradeship (10)	2
		slaves	(O)	rulers (24)	2
26	1	(all the) work	(D)	work (hard) (7)	2
27	1	(the) pigs	(R)	(the) pigs (20)	2
29	3	(the) other animals	(O)	(The) pigs (27)	3
		work	(R)	work (7)	2
		(the) farm	(R)	(the) farm (11)	2
30	1	(the) pigs	(R)	(the) pigs (20)	2
31	1	other animals	(O)	(The) pigs (30)	3
32	1	(the) pigs	(R)	(the) pigs (20)	2
33	2	working very hard	(R)	work hard (7)	5
		(the) farm	(R)	(the) farm (11)	2
34	1	Animals	(R)	(the) animals (2)	2
35	1	food	(I)	apples (30)	2
36	2	Squealer	(R)	Squealer (31)	2
		justify	(R)	justifies (30)	2
38	1	farmers	(D)	(the) farm (33)	2
39	1	(the) hens	(R)	(the) hens (8)	2
42	1	Animals	(R)	(the) animals (2)	2

F-unit no.	No of ties	Cohesive item	Type	Presupposed Item	No. Of words
43	1	Beast of England	(R)	Beast of England (12)	6
44	2	Seven Commandments	(R)	(the) seven "Commandments" (12)	4
		(the) pigs	(R)	(the) pigs (20)	2
45	3	(the) animals question Squealer	(R)	(other) animals question Squealer (31)	6
		animals	(R)	(the) animals (2)	2
		Squealer	(R)	Squealer (31)	2
47	1	(the) animals	(R)	(the) animals (2)	2
48	2	Conditions	(R)	condition (15)	2
		(the) farm	(R)	(the) farm (11)	2
49	2	comrades	(R)	(the other) comrades (25)	2
		were better	(O)	deteriorated (48)	3
50	2	Napoleon	(R)	Napoleon (43)	2
		(the) other animals	(O) ^w	Napoleon (50)	3
52	1	working hard	(R)	work hard (7)	4

Relation of first and last paragraphs to the question:

Paragraph	F- units	Repetition	Synonymy	Opposites	Inclusion	Deriva tion	words
Opening	12	1	0	0	0	0	4
Closing	4	0	0	0	0	0	0

Script 6 HCR =11

(1) In many respects conditions on Animal Farm greatly differs from that of the old Manor Farm of Mr Jones. (2) This short exposition will try to reveal the conditions of the Animal Farm. (3) After the departure of Snowball, Napoleon became the undisputed, supreme leader or head of the Animal Farm. (4) In the task of administering the Farm, (5) he was assisted only by the pigs. (6) Decision making was in their hands. (7) Nobody had the right to challenge. Napoleon's decisions.

(8) The Animal Farm was well organized (9) and as a result it became very very prosperous. (10) We **realized** this (11) because new lands were bought (12) and added to it (11) from Mr Pilkington. (13) The windmill was also completed (14) and it was used as a threshing machine for **cornmill**. (15) It also had a hay elevator of its own. (16) Other new buildings had been added.

(17) There was trade with the outside world (18) and we are aware that it brought in very handsome dividends.

(19) During the time of Snowball (20) the brain behind the windmill, (19) the animals were told that the windmill was to provide things like electricity, and hot and cold water! (21) After its completion, we **realized** that it was unable to provide these items. (22) In fact Napoleon **demande**d these items as contrary to the spirit of Animalism.

(23) Despite the fact that outwardly the farm looked prosperous, (24) the bulk of the animals never enjoyed anything! (35) Life seemed to be worse for all the animals (26) apart from the pigs and the dogs! (27) There was not much food to eat! (28) It seemed even worse when compared with Jones time - under the Manor Farm

management! (29) All the lofty ideas of Old Major was not fulfilled. (30) no animal was free! (31) Everything good was for Napoleon and his family the pigs, (32) assisted or well guarded by the dogs!

(33) The ruling class, (34) that is the pigs led by Napoleon, (33) were basically on equal footing with the human beings. (35) The master servant relationship between animals and men was **erased!** (36) Ironically this same master/servant relationship (37) which other fought against (36) was ruthlessly enforced on the animals in the Animal Farm by Napoleon and his clansmen (38) assisted by the Security Police (39) - that is the dogs! (40) There was no justice on the Animal Farm. (41) The **massive** execution of the dissidents testifies to this.

(42) This brief essay clearly shows that conditions on the Animal Farm completely differ from those on Manor Farm. (42)

Cohesive ties in the whole script:

F-unit no.	No. of ties	Cohesive item	Type	Presupposed item	No. of words
2	2	conditions	(R)	conditions (1)	2
		Animal Farm	(R)	Animal Farm (1)	4
3	2	head	(S)*	leader (3)	2
		Animal Farm	(R)	Animal Farm (1)	4
7	2	Napoleon's	(R)	Napoleon (3)	2
		decisions	(R)	decision making (6)	3
8	1	(the) Animal Farm	(R)	Animal Farm (1)	4
16	2	buildings	(I)	hay elevator (15)	3
		added	(R)	added (12)	2
17	1	trade	(I)	bought (11)	2
19	1	Snowball	(R)	Snowball (3)	2
20	1	(the) windmill	(R)	windmill (13)	2
19	2	(the) windmill	(R)	windmill (13)	2

F-unit no.	No. of ties	Cohesive item	Type	Presupposed item	No. of words
		things	(I) ^{wk}	electricity and hot and cold water (19)	7
21	3	completion	(D)	completed (13)	2
		unable to provide	(O)	provide (19)	4
		items	(I)	electricity and hot and cold water (19)	7
22	3	Napoleon	(R)	Napoleon (3)	2
		items	(R)	items (21)	2
		Animalism	(D)	(the) animals (19)	2
23	2	(the) farm	(R)	(the) farm (4)	2
		prosperous	(R)	prosperous (9)	2
24	1	(the bulk of the) animals	(R)	(the) animals (19)	2
25	1	(all the) animals	(R)	(the) animals (19)	2
26	1	(the) pigs and the dogs	(O)	all the animals (25)	7
28	4	worse	(R)	worse (25)	2
		Jones	(R)	Mr. Jones (1)	3
		Manor Farm	(R)	Manor Farm (1)	4
		management	(S)	administering (4)	2
30	1	(No) animal	(R)	(the) animals (19)	2
31	2	Napoleon	(R)	Napoleon (3)	2
		(the) pigs	(R)	(the) pigs (5)	2
32	3	assisted	(R)	assisted (5)	2
		well guarded	(S) ^w	assisted (32)	3
		(the) dogs	(R)	(the) dogs (26)	2

F-unit no.	No. of ties	Cohesive item	Type	Presupposed item	No. of words
34	2	(the) pigs	(R)	(the) pigs (5)	2
		Napoleon	(R)	Napoleon (3)	2
35	2	animals	(R)	(the) animals (19)	2
		men	(S)	(the) human beings (33)	3
36	8	Ironically	(D)	basically (33)	2
		master/servant relationship	(R)	master servant relationship (35)	6
		ruthlessly	(D)	basically (33)	2
		(the) animals	(R)	(the) animals (19)	2
		(the) Animal Farm	(R)	Animal Farm (1)	4
		Napoleon and his clansmen	(O) ^w	(the) animals (36)	5
		Napoleon	(R)	Napoleon (3)	2
		clansmen	(S)	(his) family (31)	2
38	1	assisted	(R)	assisted (5)	2
39	1	(the) dogs	(S)	(the) Security Police (38)	3
40	1	(the) Animal Farm	(R)	Animal Farm (1)	4
42	6	brief essay	(S)	short exposition (1)	4
		shows	(S)	(will try to) reveal (2)	2
		conditions	(R)	conditions (1)	2
		(the) Animal Farm	(R)	Animal Farm (1)	4
		differ	(R)	differs (1)	2
		Manor Farm	(R)	Manor Farm (1)	4

Relation of the first and last paragraphs to the question:

Paragraph	F- units	Repetition	Synonymy	Opposition	Inclusion	Deriv ation	Word
Opening	13	7	0	0	0	0	24
Closing	1	3	0	0	0	0	10

Script 49 HCR = 12

(1) Major describes the conditions on Manor Farm as miserable, laborious and unacceptable to the animals. (2) He says that the humans have no regard for the animals, (3) the animals are used only for as long as they are useful. (4) The animals will not see out their natural life, (5) rather end up being slaughtered. (6) He says that the humans are the only animals that consume and never produce. (7) He tells the animals that they will labour (8) until every **atom** of strength has been sapped. (9) The animals on Manor Farm do not see the fruits of their produce. (10) The hens' eggs are not allowed to hatch, (11) the milk is sold (12) and Clover's foals are sent away before their first birthday. (13) The animals, in the beginning, are underfed, not cared for and taken for granted. (14) This very fact leads to the uprising on Manor Farm (15) when Jones and the others are forced to leave.

(16) Initially, after the Rebellion, things go well on Animal Farm. (17) There is plentiful grain, meat and milk. (18) All animals are classified as equal (19) and new fervour abounds. (20) The animals set about establishing themselves on Animal Farm; (21) even think about building a windmill. (22) as time progresses, (23) the conditions they left (24) and said goodbye to, (25) re-emerge. (26) The initial idealistic view propogated by Major, (27) and then Snowball and Napoleon (26) was one of adherence to Seven Commandments and later to the Principles of Animalism. (28) Time changes all this. (29) Every Commandment is bent to suit the pigs and especially Napoleon, (30) Class distinction is introduced. (31) Although Animal Farm seems to prosper in terms of the trappings of material wealth, the animals are no

better off. (35) The luxuries they were promised do not materialize (36) and the spirit of unity as mentioned by Major has faded. (37) The pigs become the aristocracy and privileged class. (38) In actual fact, the conditions on Manor Farm do not nearly differ from those on Animal Farm in respect of the animals. (39) The only change that takes place is the removal of Jones and the humans (40) to be replaced by Napoleon and his entourage. (41) The values propagated by Napoleon are not really different to those of Jones. (42) One sees the analogy of Animal Farm in many countries of the world to-day. (43) The masses rebel to overthrow an unjust government or dictator (44) only to see him replaced by someone even worse. (45) True freedom comes from education (46) when people do not follow like the sheep in Animal Farm.

Cohesive ties in the whole script:

F-unit no.	No. of ties	Cohesive item	Type	Presupposed item	No. of words
1	1	unacceptable	(D) ^w	Miserable (1)	2
2	1	(the) animals	(R)	(the) animals (1)	2
3	2	(the) animals	(R)	(the) animals (1)	2
		useful	(D) ^w	used (3)	2
4	1	(the) animals	(R)	(the) animals (1)	2
5	1	slaughtered	(O)	see out their natural life (4)	6
6	3	humans	(I) ^{kw}	(the) animals (1)	2

F-unit no.	No. of ties	Cohesive item	Type	Presupposed item	No. of words
		(the only) animals	(R)	(the only) animals ()	2
		consume	(O) ^{kw}	produce (6)	2
7	2	(the) animals	(R)	(the) animals (1)	2
		labour	(D)	laborious (1)	2
9	2	(The) animals	(R)	(the) animals (1)	2
		Manor Farm	(R)	Manor Farm (1)	4
10	1	eggs	(I)	(the) fruits of their produce (9)	5
11	1	(the) milk	(I)	(the) fruits of their produce (9)	5
12	1	foals	(I)	(the) fruits of their produce (9)	5
13	1	(The) animals	(R)	(the) animals (1)	2
14	1	Manor farm	(R)	Manor farm (1)	4
16	1	(the) Rebellion	(S)	(the) uprising	2

F-unit no.	No. of ties	Cohesive item	Type	Presupposed item	No. of words
17	1	grain, meat and milk	(I)	(the) fruits of their produce (9)	7
18	1	(All) animals	(R)	(the) animals (1)	2
20	2	(The) animals	(R)	(the) animals (1)	2
		Animal Farm	(R)	Animal Farm (16)	4
23	2	(the) conditions	(R)	(the) conditions (1)	2
		left	(R)	leave (15)	2
25	1	re-emerge	(O)	left (23)	2
26	2	Major	(R)	Major (1)	2
		Animalism	(D)	(All) animals (18)	2
28	2	Time	(R)	time (22)	2
		Changes	(O)	Adherence (26)	2
29	2	(Every) Commandment	(R)	(Seven) Commandments (26)	2
		Napoleon	(I)*	(the) pigs (29)	2
31	2	(All) animals	(R)	(the) animals (1)	2

F-unit no.	No. of ties	Cohesive item	Type	Presupposed item	No. of words
		equal	(R)	equal (18)	2
32	1	equal	(R)	equal (18)	2
33	1	Animal Farm	(R)	Animal Farm (16)	4
34	1	(the) animals	(O)	(the) pigs (29)	2
35	1	materialize	(D)	material (33)	2
36	1	Major	(R)	Major (1)	2
37	2	(The) pigs	(R)	(the) pigs (29)	2
		privileged class	(S) ^w	aristocracy (37)	3
38	4	(the) conditions	(R)	(the) conditions (1)	2
		Manor Farm	(R)	Manor Farm (1)	4
		Animal Farm	(R)	Animal Farm (16)	4
		(the) animals	(R)	(the) animals (1)	2
39	3	change (n)	(D)	changes (v) (28)	2
		Jones	(R)	Jones (15)	2
		(the) humans	(R)	(the) humans (2)	2

F-unit no.	No. of ties	Cohesive item	Type	Presupposed item	No. of words
40	2	replaced	(O)	removal (39)	2
		Napoleon	(R)	Napoleon (27)	2
41	5	(The) values	(S)	(the) Principles (26)	2
		propogated	(R)	propogated (26)	2
		Napoleon	(R)	Napoleon (27)	2
		different	(D)	differ (27)	2
		Jones	(R)	Jones (15)	2
42	1	Animal Farm	(R)	Animal Farm (16)	4
43	3	masses	(O)	(the) aristocracy (37)	2
		rebel	(D)	(the) Rebellion (16)	2
		dictator	(I)*	Unjust government (40)	3
44	2	replaced	(R)	Replaced (40)	2
		worse	(O)	better (34)	2

F-unit no.	No. of ties	Cohesive item	Type	Presupposed item	No. of words
46	2	people	(S)	humans (39)	2
		Animal Farm	(R)	Animal Farm (16)	4

Relation between the first and last paragraphs and the question:

Paragraph	F-units	Repetition	Synonymy	Opposition	Inclusion	Derivation	Word
Opening	5	2	0	0	0	0	6
Closing	9	6	0	0	0	1	24

BIBLIOGRAPHY

- Angelil-Carter, S. and Eberhard Thesen, L. 1990. A process approach to academic writing: does it work? In Chick, K. (ed.) 1990: 585 - 594.7
- Bamberg, B. 1983. What makes a text coherent? *College Composition and Communication*, 34(4): 417 - 429.
- Bamberg, B. 1984. Assessing coherence: a reanalysis of essays written for the national assessment of education progress, 1969 - 1979. *Research in the Teaching of English*, 18(3), October: 305 - 319.
- Bardovi-Harlig, K. 1990. Pragmatic word order in English composition. In Connor, U. and Johns, A.m. (eds): 43 - 65.
- Bartholomae, D. 1994. The study of error. In Tate, G, Corbett, E.P.J, and N. Myers (eds) 1994: 338 - 352.
- Bizzell, P. 1986. What happens when basic writers come to college? *College Composition and Communication*, 37(3): 294 - 301.
- Bhatia, V.K. 1993. Analysing genre: language use in professional settings. London: Longman.
- Bloor, M. 1998. English for specific purposes: the preservation of the species (some notes on a recently evolved species and the contribution of John Swales to its preservation and protection). *English for Specific Purposes*, 17(1): 47 - 66.
- Brostoff, A. 1981. Coherence: "next to" is not "connected to". *College Composition and Communication*, 32: 278 - 293.
- Brown, J.D. 1988. *Understanding research in second language learning. A teacher's guide to statistics and research design*. Cambridge: Cambridge University Press.
- Brown, G. and Yule, G. 1983. *Discourse analysis*. Cambridge: Cambridge University Press.
- Brandt, D. 1986. Text and context: how writers come to mean. In Couture, B. (ed) 1986: 93 - 107.
- Carrell, P.L. 1982. Cohesion is not coherence. *TESOL Quarterly*, 16 (4), December: 479 - 488.
- Carstens, W.A.M. 1987. De Beaugrande en Dressler (1981) se "standaarde van tekstualiteit". *South African Journal of Linguistics*, 5 (2): 17 - 32.
- Carstens, W.A.M. 1997. *Afrikaanse tekslinguistiek*. Pretoria: Van Schaik.

- Carter, R. 1987. *Vocabulary: applied linguistic perspectives*. London: Allen and Unwin.
- Carter, R. and McCarthy, M. 1988. *Vocabulary and language teaching*. London: Longman.
- Chapman, J. 1983. *Reading development and cohesion*. London: Heinemann.
- Charolles, M. 1985. Text connexity, text coherence and text interpretation processing. In Sözer, E. (ed.) 1985:1 - 15.
- Chick, K. 1990. *Searching for relevance. Contextual issues in applied linguistics in Southern Africa*. Durban: SAALA.
- Clark, H.H. and Haviland, S.E. 1977. Comprehension and the given - new contract. In Freedle, R.O. (ed.) 1977: 1 - 39.
- Connor, U. 1987. Research frontiers in writing analysis. *TESOL Quarterly*, 21(4), December: 677 - 695.
- Connor, U. and Johns, A.M. (eds) 1990. *Coherence in writing*. Virginia: TESOL.
- Cook, G. 1989. *Discourse*. Oxford: Oxford University Press.
- Coulthard, M. (ed.) 1994. *Advances in written text analysis*. London: Routledge.
- Couture, B. (ed.) 1986. *Functional approaches to writing: research perspectives*. London: Frances Pinter.
- Creswell, J.W. 1994. *Research design: qualitative and quantitative approaches*. Thousand Oaks: Sage Publications.
- Crothers, E.J. 1978. Inference and coherence. *Discourse Processes*, 1: 59 - 71.
- Cruse, D.A. 1986. *Lexical semantics*. Cambridge: Cambridge University Press.
- Davies, F. 1988. Designing a writing syllabus in English for academic purposes: process and product. *Academic Writing: Process and Product, ELT Documents*, 129: 130 - 142.
- De Beaugrande, R. and Dressler, W. 1981. *Introduction to text linguistics*. London: Longman.
- Eiler, M.A. 1983. Meaning and choice in writing about literature. In Fine, J. and Freedle, R.O. (eds) 1983: 169 - 222.
- Ellis, R. 1994. *The Study of second language acquisition*. Oxford: Oxford University Press.

- Engber, C. A. 1995. The relationship of lexical proficiency to the quality of ESL compositions. *Journal of second language writing*, 4 (2):139 - 155.
- Enkvist, N.I. 1990. Seven problems in the study of coherence and interpretability. In Connor, U. and Johns, A.M. (eds) 1990: 11 - 28.
- Fahnestock, J. 1983. Semantic and lexical coherence. *College Composition and Communication*, 34(4), December: 400 - 416.
- Fine, J. and Freedle, R.O. 1983. *Developmental Issues in discourse*. New Jersey: Ablex.
- Flood, J. (ed.) 1984. *Understanding reading comprehension: cognition, language, and the structure of prose*. Delaware: International Reading Association.
- Flower, L. 1979. Writer-based prose: a cognitive basis for problems in writing. *College English*, 14(1), September: 19 - 37.
- Freedle, R.O. (ed.) 1977. *Discourse production and comprehension*. New Jersey: Ablex.
- Gairns, R. and Redman, S. 1986. *Working with words*. Cambridge: Cambridge University Press.
- Grobe, C. 1981. Syntactic maturity, mechanics, and vocabulary as predictors of quality ratings. *Research in the Teaching of English*, 15 (1), February: 75 - 85.
- Hatch, E. and Farhady, H. 1982. *Research design and statistics for applied linguistics*. Rowley: Newbury House.
- Halliday, M.A.K. 1989. *Spoken and written language*. (2nd ed.) Oxford: Oxford University Press.
- Halliday, M.A.K. and Hasan, R. 1976. *Cohesion in English*. London: Longman Group.
- Hartnett, C.G. 1986. Static and dynamic cohesion: signals of thinking in writing. In Couture, B (ed) 1986: 142 - 153.
- Hasan, R. 1984. Coherence and cohesive harmony. In Flood, James (ed.) 1984: 181 - 218.
- Hoey, M. 1983. *On the surface of discourse*. London: George Allen & Unwin.
- Hoey, M. 1991. *Patterns of lexis in text*. Oxford: Oxford University Press.
- Horowitz, D. 1986a. Process, not product: less than meets the eye. *TESOL*

- Quarterly*, 29(1): 141 - 144.
- Horowitz, D.M. 1986b. What professors actually require: academic tasks for the ESL classroom. *TESOL Quarterly*, 20(3), September: 445 - 462.
- Hubbard, E.H. 1989. *Reference cohesion, conjunctive cohesion and relational coherence in student academic writing*. Unisa, PhD Thesis.
- Hubbard, E.H. 1990. Topical structure analysis as a text-linguistic explication of 'coherence': a critique. *South African Journal of Linguistics*, 8(2): 80 - 87.
- Hubbard, E.H. 1993. Some coherence correlates in expository writing. *South African Journal of Linguistics*, 15: 55 - 74.
- Hubbard, E.H. 1994. A process-oriented analysis of reference errors. *South African Journal of Linguistics*, 22, December: 63 - 79.
- Hubbard, E.H. 1996. Profiling contextual support for technical and general academic terms in writing for students. *South African Journal of Linguistics*, 32: 95 - 106.
- Jackson, H. 1988. *Words and their meaning*. London: Longman
- Johns, A.M. 1986. Coherence and academic writing: some definitions and suggestions for teaching. *TESOL Quarterly*, 20(2), June: 247 - 265.
- Johns, A.M. 1993. Written argumentation for real audiences: suggestions for teacher research and classroom practice. *TESOL Quarterly*, 27(1), Spring: 75 - 90.
- Johns, A.M. 1995. Genre and pedagogical purposes. *Journal of Second Language Writing*, 4(2): 181 - 190.
- Jordan, R.R. 1997. *English for academic purposes*. Cambridge: Cambridge University Press.
- Kaplan, R.B. 1991. Concluding essay: on applied linguistics and discourse analysis. *Annual Review of Applied Linguistics*, 11: 199 - 204.
- Kroll, B. 1990. What does time buy? ESL student performance on home versus class compositions. In Kroll, B. (ed.): 140 - 154.
- Kroll, B. (ed.) 1990. *Second language writing: research insights for the classroom*. Cambridge: Cambridge University Press.
- Laufer, B. 1990. 'Sequence' and 'order' in the development of L2 lexis: some evidence from lexical confusions. *Applied Linguistics*, 11(3): 281 - 291.

- Laufer, B. 1991. Some properties of the foreign language learner's lexicon as evidenced by lexical confusions. *IRAL*, XXIX(4), November: 317 - 330.
- Lehrer, A. 1969. Semantic cuisine. *Journal of Linguistics*, 5: 39 - 56.
- Lennon, P. 1991. Error: some problems of definition, identification, and distinction. *Applied Linguistics*, 12(2): 180 - 196.
- Lieber, P. E. 1979. *Cohesion in ESL students' expository writing: A descriptive study*. New York University. PhD thesis.
- Lyons, J. 1968. *Introduction to theoretical linguistics*. Cambridge: Cambridge University Press.
- Lyons, J. 1981. *Language and linguistics*. Cambridge: Cambridge University Press.
- Malmkjaer, K. 1991. *The linguistics encyclopedia*. London: Routledge.
- Marchand, H. 1969. *The categories and types of present-day English word-formation*. Munchen: C.H. Beck'she Verslagsbuchhandlung.
- Maringa, E. 1995. *Reference cohesion in primary school English second language writing*. Unisa, MA thesis.
- Markels, R.B. 1983. Cohesion paradigms in paragraphs. *College English*, 45(5), September: 450 - 464.
- McCarthy, M.J. 1984. A new look at vocabulary in EFL. *Applied Linguistics*, 5(1):13 - 22.
- McCarthy, M.J. 1990. *Vocabulary*. Oxford: Oxford University Press.
- McCarthy, M.J. 1991. *Discourse analysis for language teachers*. Cambridge: Cambridge University Press.
- McCulley, G.A. 1985. Writing quality, coherence and cohesion. *Research in the Teaching of English*, 19(3), October: 269 - 282.
- McCutcheon, D. and Perfetti, C.A. 1982. Coherence and connectedness in the development of discourse production. *Text*, 2(1 - 3): 113 - 139.
- Mulder, J.C. 1986. *Statistical techniques in education*. Pretoria: HAUM Tertiary.
- Myers, M. 1980. *A procedure for writing assessment and holistic scoring*. Urbana: ERIC Clearinghouse on Reading and Communication Skills.
- Neuner, J.L. 1987. Cohesive ties and chains in good and poor freshman essays. *Research in the teaching of English*, 21(1), February: 92 - 105.

- Nunan, D. 1992. *Research methods in language learning*. Cambridge: Cambridge University Press.
- Nystrand, M. 1982. An analysis of errors in written communication. In Nystrand, M. (ed.) 1982: 57 - 74.
- Nystrand, M. (ed.) 1982. *What writers know*. New York: Academic Press.
- Onions, C.T. (ed.) 1972. *The shorter oxford English dictionary*. 3rd edition. Oxford: Oxford University Press.
- Orr, M.H. 1995. Teaching writing at university: an academic literacy programme. *South African Journal of Higher Education*, 9 (2): 189 - 197.
- Pagano, A. 1994. Negatives in written text. In Coulthard, M. (ed.) 1994: 250 -265.
- Patricio, E.M. 1993. *Cohesion and coherence in std 10 English second language compositions of black pupils: an analysis*. Potchefstroom University. MA thesis.
- Perl, S. 1980. Understanding composing. *College Composition and Communication*, 31: 363 - 369.
- Quirk, R. 1978. Focus, scope, and lyrical beginnings. *Language and Style*, 11: 30 - 39.
- Raimes, A. 1985. What unskilled ESL students do as they write: a classroom study of composing. *TESOL Quarterly*, 19(2), June: 229 - 258.
- Raimes, A. 1993. Out of the woods: emerging traditions in the teaching of writing. In Silberstein, S. (ed.) 1993: 237 - 260.
- Rodseth, V., Johanson, L. and Rodseth, W. 1992. *Think write*. Johannesburg: Hodder & Stoughton.
- Roets, N. 1992. Die onderrig van Afrikaans aan anderstaliges op tersiere vlak. *Tydskrif vir geesteswetenskappe*, 32(2), June: 143 - 154.
- Rudzka, B., Channell, J., Putseys, Y. and Ostry, P. 1981. *The words you need*. London: Macmillan.
- Rwigema, Q. 1996. *An investigation into the role of rule schema and the linguistic environment as factors that hamper language acquisition and competence in written English among students of English as a second language in the Diploma in Secondary Education at the University of Transkei*. University of Transkei. MA thesis.
- Sanford, A.J. and Garrod, S.C. 1981. *Understanding written language*. Chichester:

John Wiley and Sons.

- Seliger, H. W. & Shohamy, E. 1989. *Second language research methods*. Oxford: Oxford University Press.
- Shih, M. 1986. Content-based approaches to teaching academic writing. *TESOL Quarterly*, 20(4), December: 617 - 648.
- Shiro, M. 1994. *Inferences in discourse comprehension*. In Coulthard, M. (ed.) 1994: 167 - 178.
- Silberstein, S. 1993. *State of the art TESOL essays*. Illinois: TESOL.
- Sinclair, J.M. 1993. Written discourse structure. In Sinclair, J.M., Hoey, M. & Fox, G. (eds) 1993: 6 - 31.
- Sinclair, J.M., Hoey, M. & Fox, G. (eds) 1993. *Techniques of description: spoken and written discourse*. London: Routledge.
- Sonaiya, O.C. 1988. *The lexicon in second language acquisition: a lexical approach to error analysis*. Cornell University. Phd thesis.
- Sözer, E. (ed.) 1985. *Text connexity, text coherence*. Hamburg: Helmut Buske Verlag.
- Starfield, S. 1990. Contextualising language and study skills. *South African Journal of Higher Education: Special edition*: 142 - 148.
- Stotsky, S. 1979. Teaching the vocabulary of academic discourse. *Journal of Basic Writing*, 2, Fall / Winter: 15 - 39.
- Stotsky, S. 1981. The vocabulary of essay writing: can it be taught? *College Composition and Communication*, 32: 317 - 326.
- Stotsky, S. 1983. Types of lexical cohesion in expository writing: implications for developing the vocabulary of academic discourse. *College Composition and Communication*, 34: 430 - 446.
- Stotsky, S. 1984. A proposal for improving high school students' ability to read and write expository prose. *Journal of Reading*, 28: 4 - 7.
- Stotsky, S. 1986. On learning to write about ideas. *College Composition and Communication*, 37(3): 276 - 293.
- Stotsky, S. 1995. The uses and limitations of personal or personalized writing in writing theory, research, and instruction. *Reading Research Quarterly*, 30(4) October/November/December: 758 - 776.

- Swales, J.M. 1990. *Genre analysis*. Cambridge: Cambridge University Press.
- Tate, G., Corbett, E.P.J. and N. Myers. 1994. *The writing teacher's sourcebook*. (3rd ed.) New York: Oxford University Press.
- Tierney, R.J. and Mosenthal, J.H. 1983. Cohesion and textual coherence. *Research in the Teaching of English*, 17(3), October: 215 - 229.
- Van Dijk, T.A. 1977. *Text and context*. London: Longman.
- Van Zyl, M.H. 1993. The teaching of written composition at university. *Progressio*, 15(1): 43 - 58.
- Weissberg, R.C. 1984. Given and new: paragraph development models from scientific English. *TESOL Quarterly*, 18(3), September: 485 - 499.
- Wessels, E. M. 1993. *Bonding and related measures of coherence in student academic writing*. Unisa. MA thesis.
- Wessels, E.M. 1994. Lexical cohesion and the teaching of academic skills. *South African Journal of Linguistics*, 22, December: 123 - 148.
- Willis, D. 1993. Grammar and lexis: some pedagogical implications. In Sinclair, J.M., Hoey, M. and Fox, G. (eds) 1993: 83 - 93.
- Witte, S.P. 1983. Topical structure and writing quality: some possible text-based explanations of readers' judgements of student writing. *Visible Language XVII* (2): 177 - 205.
- Witte, S.P. and Faigley, L. 1981. Coherence, cohesion, and writing quality. *College Composition and Communication*, 32, May: 189 - 204.
- Yde, P. and Spoelders, M. 1985. Text cohesion: an exploratory study with beginning writers. *Applied Psycholinguistics*, 6: 407 - 415.