Aston University

Some parts of this thesis may have been removed for copyright restrictions.

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

THE DISCOURSE OF BIOLOGY LECTURES:
ASPECTS OF ITS MODE AND TEXT STRUCTURE

FARIDA BAKA

Doctor of Philosophy

THE UNIVERSITY OF ASTON IN BIRMINGHAM

April 1989

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

BEST COPY

AVAILABLE

Variable print quality

THESIS CONTAINS VIDEO

THE UNIVERSITY OF ASTON IN BIRMINGHAM

THE DISCOURSE OF BIOLOGY LECTURES: ASPECTS OF ITS MODE AND TEXT STRUCTURE

SUMMARY

The present thesis investigates mode related aspects in biology lecture discourse and attempts to identify the position of this variety along the spontaneous spoken versus planned written language continuum. Nine lectures (of 43,000 words) consisting of three sets of three lectures each, given by three lecturers at Aston University, make up the corpus.

The indeterminacy of the results obtained from the investigation of grammatical complexity as measured in subordination motivates the need to take the analysis beyond sentence level to the study of mode related aspects in the use of sentence-initial connectives, sub-topic shifting and paraphrase.

It is found that biology lecture discourse combines features typical of speech and writing at sentence as well as discourse level: thus, subordination is more used than co-ordination, but one degree complexity sentence is favoured; some sentence-initial connectives are only found in uses typical of spoken language but sub-topic shift signalling (generally introduced by a connective) typical of planned written language is a major feature of the lectures; syntactic and lexical revision and repetition, interrupted structures are found in the sub-topic shift signalling utterance and paraphrase, but the text is also amenable to analysis into sentence-like units . On the other hand, it is also found that: (1) while there are some differences in the use of a given feature, inter-speaker variation is on the whole not significant; (2) mode related aspects are often motivated by the didactic function of the variety; and (3) the structuring of the text follows a sequencing whose boundaries are marked by sub-topic shifting and the summary paraphrase.

This study enables us to draw four theoretical conclusions: (1) mode related aspects cannot be approached as a simple dichotomy since a combination of aspects of both speech and writing are found in a given feature. It is necessary to go to the level of textual features to identify mode related aspects; (2) homogeneity is dominant in this sample of lectures which suggests that there is a high level of standardization in this variety; (3) the didactic function of the variety is manifested in some mode related aspects; (4) the features studied play a role in the structuring of the text.

Key Words

Biology lecture discourse
Speech and writing
Grammatical complexity
Textual features

Farida Baka
Ph.D.
1989

To my husband Omar ,
with love

ACKNOWLEDGEMENTS

I would like to thank Mrs Catherine Johns Lewis, my previous supervisor, for helpful academic discussions and comments during the primary stages of this work. I would like to acknowledge my indebtedness to Mr Tom Bloor who undertook the supervision of the crucial stages of this thesis. My thanks are also due to Professor Dennis Ager and Professor Francis Knowles for ensuring the continuity of the present research.

I would also like to thank the Ministry of Higher Education in Algeria for offering me a grant to carry out this research in the United Kingdom.

To my husband Omar, I owe insightful academic discussions about different parts of this work, moral support and love.

My thanks to Helen Higson who typed early drafts of the thesis.

Finally, any shortcomings or defects in the work remain, of course, my own responsibility.

LIST OF CONTENTS

		Page
CHAPTER ONE:	AN INTRODUCTION TO THE PRESENT STUDY	17
1.0		
1.1	The present study	17
1.2	Corpus	18
1.2.1	Types and sources of data used in the study	18
1.2.2	Rationale and limitations of the present corpus	19
1.2.2.1	Rationale for the corpus	19
1.2.2.2	Limitations of the corpus	20
1.2.3	Methodology of data collection	20
1.2.4	Transcription conventions	21
CHAPTER TWO:	LECTURE DISCOURSE : A VARIETY OF SPOKEN LANGUAGE, THEORETICAL BACKGROUND AND THE PRESENT APPROACH	23
2.0		
2.1	Non-linguistic interest in lectures	23
2.1.1	Anthropological approach to lectures	23
2.1.2.	Educational approach to lectures	24
2.2	Linguistic interest in lectures	25
2.3	Approaches to language variety	31
2.3.1	The rhetorical approach to variety	32
2.3.1.1	The grammatical-rhetorical approach to variety	32
2.3.1.2	The rhetorical approach to variety	33
2.3.2	The non-rhetorical approach to language variety	34
2.4	Speech versus writing	39
2.4.1	Spontaneity versus forethought	41
2.4.2	Fragmentariness versus integration and involvement versus detachment	42
2.4.3	Lakoff's grid	44
2.4.4	Major features of unplanned spoken versus planned written discourse	45
2.4.5	Grammatical complexity	48
2.5	Approaches to discourse structure	52
2.5.1	"Discourse" and "text"	52
2.5.2	Pragmatic approach to discourse structure	53
2.5.3	Schema as cognitive approach to discourse structure	54
2.6.	The present approach	58
2.6.1	Biology lectures : Planned or unplanned discourse	58
2.6.2	Unplanned spoken versus planned written discourse	59
2.6.3	Approach to features of unplanned spoken as opposed to planned written discourse in the present study	59
2.6.3.1	Grammatical complexity	60
2.6.3.2	Textual features	60

		Page
2.6.3.2.0		
2.6.3.2.1	Textual features investigated	61
2.6.4	Theoretical implications of the present approach to the study of features of speech/writing in biology lecture discourse	63
2.6.5	Variation within lectures	63
	Footnotes	65
CHAPTER THREE:	ASPECTS OF GRAMMATICAL COMPLEXITY IN BIOLOGY LECTURE DISCOURSE	67
3.0		
3.1	The sentence in spoken language	67
3.1.0		
3.1.1	Identifying sentence boundaries	69
3.1.2	Types of sentence in this analysis	71
3.2	Depth and complexity	72
3.3	Co-ordination versus subordination/embeddedness	75
3.3.0		
3.3.1	Co-ordination	75
3.3.2	Subordination	77
3.3.3	Embeddedness	79
3.4	Finite versus non-finite clauses	82
3.5	Types of dependent clause	84
3.6	Major types of embedded clause	85
3.6.1	Restrictive relative clause	85
3.6.2	That-clause	90
3.7	Summary and conclusion	92
	Footnotes	94
CHAPTER FOUR:	SENTENCE-INITIAL CONNECTIVES	95
4.0.		
4.1	Preliminaries	95
4.1.1	Terminology	95
4.1.2	Sentence-initial connectives in the corpus	96
4.2	Kinds of sentence-initial connectives: frequency of occurrence in the corpus	98

		Page
4.2.2	Frequency of occurrence of one-word sentence-initial connectives in the corpus	99
4.3	Relevance and sentence-initial connectives	100
4.3.1	The concept of relevance	100
4.3.2	The concept of relevance and biology lecture discourse	102
4.4	Semantic relevance	104
4.4.1	Semantic progression within one sub-topic	104
4.4.2	Reformulation within one sub-topic	106
4.4.3	Shifting of sub-topic focus	107
4.4.4	Features of semantic relevance in the use of sentence-initial connectives and the spontaneous spoken versus planned written discourse distinction	108
4.5	Pragmatic relevance	110
4.5.1	Features of speaker/audience interaction	110
4.5.2	Features of the speaker's monitoring of his speech	112
4.5.2.1	"Well" as answer initiator following a rhetorical question	112
4.5.2.2	"Well" as editing marker of a reformulation	113
4.5.3	Pragmatic relevance in the use of sentence-initial connectives and the spontaneous spoken versus planned written distinction	114
4.6	Connectives in sentences used in relation to the visual element	116
4.7	Sentence-initial connectives in relation to the structure of the text	117
4.7.1	Sentence-initial connectives in relation to the macrostructure of the text	117
4.7.2	Sentence-initial connectives as signals of the internal structure of a chunk of text	120
4.7.2.1	Connectives as sentence initiator following a sub-topic shift signalling utterance	120
4.7.2.2	Connectives as initiators of a sentence following a digression	121
4.7.2.3	Sentence-initial connectives as initiator of a feature of the internal structuring of a chunk of text	122
4.7.3	Features of spontaneous spoken versus planned written discourse in the use of sentence-initial connectives in relation to text structure	122
4.7.3.1	Features of spontaneous spoken discourse in the use of connectives in relation to text structure	122
4.7.3.2	Features of planned written discourse in the use of sentence-initial connectives in relation to text structure	124
4.8	Summary	124
4.9	Conclusion	125
	Footnotes	127

		Page
CHAPTER FIVE:	SUB-TOPIC SHIFT SIGNALLING IN BIOLOGY LECTURES	128
5.0		
5.1	Sub-topic shift signalling in the lectures	128
5.1.0		
5.1.1	Topic and sub-topic : scope of the sub-topic shift signalling (STSS) utterance	129
5.1.2	Grammatical status of the sub-topic shift signalling (STSS) utterance	130
5.2	Identifying the sub-topic shift signalling (STSS) utterance	130
5.2.1	Criteria: nature of criteria for identifying the sub-topic shift signalling (STSS) utterance	131
5.2.2	Types of criteria	132
5.2.2.1	Metatextual criterion	132
5.2.2.2	Semantic criterion	133
5.2.2.2.1	Identifying new-information (NI)	133
5.2.2.2.2	Textual criteria for identifying new information (NI)	134
5.2.2.3	Phonological criterion for identifying the sub-topic shift signalling (STSS) utterance	136
5.2.2.4	Syntatic criterion	136
5.2.2.5	Textual criteria	137
5.2.2.6	Non-verbal criterion	138
5.3	Types and distribution of the sub-topic shift signalling (STSS) utterance. Features of spontaneous spoken versus planned written language in the use of types of the signalling utterance	139
5.3.1	The metatextual comment	139
5.3.2	The statement	140
5.3.3	The change of structure	141
5.3.3.1	Interrogative clause	142
5.3.3.2	Imperative clause	143
5.3.4	Marginal types of the signalling utterance	144
5.3.5	Summary table of the frequency of occurrence of types of sub-topic shift signalling (STSS) in the corpus	145
5.4	Textual features of sub-topic shift signalling (STSS)	146
5.4.1.	Sentence-initial connectives in the signalling utterance	146
5.4.1.1	Distribution of sentence-initial connectives in the three types of the sub-topic shift signalling (STSS) utterance	146
5.4.1.2	Features of spoken language in the use of connectives in the STSS utterance	147
5.4.2	The STSS utterance in relation to the macrostructure of the text	148
5.4.2.1	General and partial sub-topic shift signalling	148
5.4.2.2	The sub-topic shift signalling (STSS) utterance and the summary paraphrase	151
5.4.2.2.1	Types of sub-topic shift signalling preceded by a summary paraphrase	152

		Page
5.5	The sub-topic shift signalling (STSS) and visual aids	152
5.5.0		
5.5.1	Reference to the visual in the sub-topic shift signalling (STSS) utterance	154
5.5.2	Features of oral delivery in the interaction between the text and the visual	155
5.6	Grammatical features of the sub-topic shift signalling (STSS) utterance	156
5.6.0		
5.6.1	New-information carrier : Types and functions	156
5.6.1.1	Types of new information (NI) carrier	156
5.6.1.2	Functions of new information carrier	158
5.6.2	Personal pronouns in the sub-topic shift signalling utterance	160
5.6.3	Aspects of grammatical complexity in sub-topic shift signalling in relation to spoken versus written language	162
5.6.3.1	Subordination versus co-ordination in the sub-topic shift signalling (STSS) utterance	163
5.6.3.2	Degree of subordination in sub-topic shift signalling (STSS)	164
5.6.3.3	Finite versus non-finite clauses in the STSS utterance	165
5.6.3.4	Major types of dependent clause	165
5.7	Features of spontaneity in the sub-topic shift signalling(STSS) utterance	168
5.7.1	Syntactic revision	168
5.7.2	Lexical revision	169
5.7.3	Reiteration of part of the signalling utterance	169
5.7.4	Digression in the STSS utterance: use of a summary paraphrase interrupting the STSS utterance	171
5.8	Sub-topic shift signalling : a manifestation of the organising of the information structure of lectures	172
5.9	Summary and conclusion	173
	Footnotes	176
CHAPTER SIX:	PARAPHRASE. ITS TYPES AND TEXTUAL STRUCTURING	177
6.0		
6.1	Paraphrase in the lectures	177
6.1.0		
6.1.1	Scope of paraphrase in this study	178
6.1.2	Defining paraphrase in this study	183
6.2	The summary paraphrase (SP)	185
6.2.0		
6.2.1	Criteria for identifying the summary paraphrase (SP)	185
6.2.1.1	Semantic criterion	185

		Page
6.2.1.2	Textual criteria	186
6.2.2	Pseudo summary paraphrase	187
6.2.3	Textual features of the summary paraphrase (SP)	187
6.2.3.1	The summary paraphrase (SP) in relation to the macrostructure of the text	188
6.2.3.1.1	General versus partial summary paraphrase in the lectures	188
6.2.3.1.2	The summary paraphrase as a marker of a boundary in the macrostructure of the text	190
6.2.3.2	Use of initial "so", a feature of spoken language in the summary paraphrase (SP)	191
6.2.4	Ellipsis in the summary paraphrase	191
6.2.4.1	Cases of ellipted constituents in the summary paraphrase (SP)	191
6.2.4.2	Ellipsis as a feature of spoken language in the summary paraphrase (SP)	192
6.3	Equivalence paraphrase (EP)	193
6.3.0		
6.3.1	Criteria for identifying the equivalence paraphrase	194
6.3.1.1	Criteria	194
6.3.1.2	Typical examples	198
6.3.2	Types of equivalence paraphrase (EP)	200
6.3.3	Motivations for lexical and syntactic differences between the utterance and its reformulation in the equivalence paraphrase (EP)	201
6.3.3.0		
6.3.3.1	Motivations for lexical differences in the equivalence paraphrase (EP)	201
6.3.3.2	Motivations for syntactic differences in the equivalence paraphrase (EP)	203
6.3.4	Textual features of the equivalence paraphrase (EP)	204
6.3.4.0		
6.3.4.1	The equivalence paraphrase (EP) is a feature of the sequence	204
6.3.4.2	Distribution of the equivalence paraphrase (EP) in relation to the original utterance	205
6.3.4.3	Dependence on co-text	205
6.4	The inference paraphrase	207
6.4.1	Criteria for identifying the inference paraphrase	207
6.4.2	Textual features of the inference paraphrase (IP)	210
6.4.2.0		
6.4.2.1	The inference paraphrase (IP) is a feature of the sequence	210
6.4.2.2	Dependence on co-text	211
6.5	Textual continuity and paraphrase	213
6.6	Distribution of the three types of paraphrase in the corpus	215
6.7	Features of unplanned spoken versus planned written discourse in paraphrase	216
6.7.0		
6.7.1	Situational features as markers of unplanned spoken discourse in paraphrase	216

		Page
6.7.1.1	Features of the interaction between the textual realization of paraphrase and the visual element	216
6.7.1.1.1	Reference to the visual element in paraphrase in the lectures	216
6.7.1.1.2	Features of oral delivery in the reference to the visual in paraphrase in the lectures	218
6.7.1.2	Audience oriented features in paraphrase	219
6.7.2	Non-situational features in paraphrase in the lectures in relation to unplanned spoken versus planned written discourse	220
6.7.2.1	Aspects of grammatical complexity in paraphrase	220
6.7.2.1.1	Subordination versus co-ordination in paraphrase	221
6.7.2.1.2	Depth or degree of subordination in paraphrase	221
6.7.2.1.3	Finite versus non-finite clauses in paraphrase	222
6.7.2.1.4	Favoured types of dependent clause in paraphrase	223
6.7.2.2	Deictic "so" as paraphrase initiator	225
6.7.2.3	Ellipsis, a feature of spoken language in paraphrase	226
6.7.2.4	Repetition as a feature of oral delivery in the lectures	227
6.7.2.5	Interrupted structures in paraphrase	229
6.8	Conclusion	230
	Footnotes	232
CHAPTER SEVEN:	SOME ASPECTS OF INTER-SPEAKER VARIATION IN THE USE OF GRAMMATICAL COMPLEXITY, CONNECTIVES, SUB-TOPIC SHIFTING AND PARAPHRASE	233
7.0		
7.1	Comparability between speakers	233
7.2	Variation between speakers	235
7.2.1	Variation in the use of zero degree grammatical complexity	235
7.2.2	Variation in the frequency of occurrence of one-word connectives "but", "now", "so", and "well"	236
7.2.3	Variation in the use of the textual structure : sub-topic shift signalling following a summary paraphrase (SP & STSS)	237
7.2.4	Variation in the use of personal pronoun "I" in the sub-topic shift signalling utterance	237
7.2.5	Variation in the frequency of occurrence of paraphrase and the summary paraphrase	237
7.2.6	Variation in the use of the visual in paraphrase	238
7.2.7	Non-significant difference in the use of types of change of structure in sub-topic shift signalling (STSS) and in the use of the visual element in the signalling utterance	238
7.3	Accounting for variation between speakers	239
7.3.1	Personnal style	239
7.3.2	Discourse topic	241
7.3.3	Structuring strategies	242

		Page
7.4	Conclusion	243
	Footnotes	245
CHAPTER EIGHT:	CONCLUSIONS AND THEORETICAL IMPLICATIONS	246
8.0		
8.1	Summary of research	246
8.2	Summary of findings	251
8.2.1	Features of unplanned spoken language	251
8.2.2	Features of planned written language	252
8.2.3	Other features	252
8.3	Theoretical implications	253
8.3.1	Variety and mode	253
8.3.2	Variety and inter-speaker variation	257
8.3.3	Variety and purpose	258
8.3.4	Text structure and the present study	260
8.4	Potential applications	262
8.5	Suggestions for further research	265
	Footnote	267
BIBLIOGRAPHY		268
Appendix 1:	(Corpus)	283
1. - Transcript of the lectures		283
Set (1)	Lecture (1) Tape A1 Side A, and B 4.5.1984	283
Set (1)	Lecture (2) Tape A2 Side A, and B 11.5.1984	290
Set (1)	Lecture (3) Tape A3 Side A, and B 18.5.1984	296
Set (2)	Lecture (1) Tape B1 Side A 16.5.1984	302
Set (2)	Lecture (2) Tape B1 Side B 16.5.1984	307
Set (2)	Lecture (3) Tape B2 Side A, and B 23.5.1984	314
Set (3)	Lecture (1) Tape B2 Side B 30.4.1985	321
Set (3)	Lecture (2) Tape C1 Side A 22.5.1984	328
Set (3)	Lecture (3) Tape C1 Side B 29.5.1984	334

2. - Copy of lecturers' notes 341

A - Lecturer's notes (Set 1) 341

B - Lecturer's notes (Set 3) 351

Appendix 2: (Appendix to chapter three) 365

1. Table (1) : Number of clauses joined by co-ordination in the same sentence 365

Table (2) : Frequency indices for co-ordinate clauses in compound and complex sentences. 365

2. Examples and discussion of types of subordinate/embedded clause not dealt with in the text 366

Appendix 3: (Appendix to chapter six) 368

Edited transcript of a structured interview with specialist informants 368

LIST OF TABLES AND FIGURES

		Page
Table (1)	General description of the corpus used in this study	22
Table (2)	Chafe's features of spoken and written language	43
Table (3)	Lakoff's six criteria for classifying types of discourse	45
Table (4)	Frequency index for co-ordinate versus subordinate/ embedded clauses	76
Table (5)	Frequency index for sentences distinguished on the basis of their degree of complexity in the corpus	78
Table (6)	Frequency index for subordinate versus embedded clauses	79
Table (7)	Number of dependent clauses in each type of construction	81
Table (8)	Frequency index for finite versus non-finite clauses	83
Table (9)	Frequency index for types of embedded clause in the corpus	86
Table (10)	Frequency index for types of subordinate clause in the corpus	87
Table (11)	Frequency index for kinds of sentence-initial connectives	99
Table (12)	Frequency of occurrence of one word sentence-initial connectives in the corpus	101
Table (13)	Frequency of occurrence of the three types of change of structure in the three sets of lectures	144
Table (14)	Summary table of the frequency index for each type of STSS utterance	145
Table (15)	Frequency of occurrence of sentence-initial connectives in the three types of STSS utterance	147
Table (16)	Frequency of occurrence of types of STSS utterance following an SP	153
Table (17)	Frequency of occurrence of types of STSS utterance accompanied by a visual	154
Table (18)	Distribution of personal pronouns in the three types of STSS utterance	161
Table (19)	Frequency index for subordinate versus co-ordinate clauses in the STSS utterance	163

		Page
Table (20)	Frequency index for sentences distinguished on the basis of their level of complexity in the STSS utterance	164
Table (21)	Frequency index for finite versus non-finite clause in the STSS utterance	165
Table (22)	Frequency index for the major types of clause in the STSS utterance	166
Table (23)	Frequency of occurrence of general summary paraphrase and partial summary paraphrase	189
Table (24)	Frequency index for the summary paraphrase (SP), the equivalence paraphrase (EP) and the inference paraphrase (IP) in the corpus	215
Table (25)	Frequency of occurrence of each type of paraphrase accompanied by a visual in the corpus	218
Table (26)	Frequency index for co-ordinate versus subordinate clauses in paraphrase	221
Table (27)	Frequency index for sentences distinguished on the basis of their degree of complexity in paraphrase	222
Table (28)	Frequency index for finite and non-finite clauses in paraphrase	223
Table (29)	Frequency of occurrence of that-clause in paraphrase	224
Fig (1)	Subordination versus embeddedness in the sentence	80
Fig..(2)	Sub-topic expansion	149

NOTATION

Transcription conventions and symbols

// Double bars indicate sentence boundary
/ Single bar indicates clause boundary

The symbols M, Sub, Em(or E), and c indicate respectively : main clause, subordinate clause, embedded clause and co-ordinating conjunction

Key to reference to corpus

Reference to the corpus reads a number in which the first digit from the left indicates the set of lectures, the next three digits indicate page number in the corpus and the two digits from the right indicate the line in which the quoted example starts. Example (127126) reads set (1), page 271, line 26.

CHAPTER ONE

AN INTRODUCTION TO THE PRESENT STUDY

1.0. Educational interest in the study of lectures has a long standing tradition. No such tradition is known within the frame of linguistics.

The present work aims at studying the speech versus writing dimension in biology lectures as a variety of English. This will be done through: firstly analysing some aspects of grammatical complexity in the lectures and secondly, investigating three textual features, namely sentence-initial connectives, topic shifting and paraphrase. It is hoped that this investigation will give us some insight into the nature of the speech versus writing dichotomy and the notions of planned versus unplanned discourse.

1.1. The present study

The first step taken towards achieving the above mentioned aims, is to look at some aspects of grammatical complexity in the present corpus. As it will be seen later, the study of subordination versus coordination has been an acknowledged technique in investigating spoken versus written language (see 4.4). However, the speech versus writing distinction, known as mode, remains but one dimension suggested in the study of text typologies (see Halliday et al, 1964; Gregory 1967; Ure & Ellis 1977; Dell Hymes 1972). By focussing on this dimension of language, it is hoped that we shall be able to enhance our knowledge of the discourse of biology lectures. It is also hoped that we shall ultimately be able to see the extent to which this dimension is in its own right a viable linguistic notion, if it is viable at all, to identify these lectures as belonging to the spoken or to the written form of language.

The three features studied in this work, namely sentence-initial connectives, sub-topic shifting and paraphrase, have one thing in common; they are found at a boundary of a chunk of text. Hence, we hope that in addition to their being used to reveal the speech versus writing properties of lectures, they will tell us something about the macrostructure of the text. The main question is, again, whether, assuming a continuum or "cline" between the extremes of typical spoken and typical written mode, biology lecture discourse can be placed nearer to the spoken or to the written end of the "cline" (see Berry 1975: 26). This question comes into focus when it is asked in relation to the realisation of the specific textual features being studied, i.e. sentence-initial connectives, sub-topic shifting and paraphrase. When it comes to these specific features, the questions are the following: why does each feature function in the way it does in these lectures? What are the speaker's strategies and/or reasons for using a given feature? To what extent would a particular feature be common to all lectures in the corpus? In other words, to what extent could individual variation be identified with the use of a feature?

1.2 Corpus

1.2.1 Types and sources of data used in the study.

Two types of data make up the present corpus:

1. Textual data:

   (i) It includes nine recorded biology lectures for first year undergraduate students at Aston University (academic year 1983-84; 1984-85). This data amount to about 43,000 words (see Table 1).

   (ii) It also includes notes used by the lecturers as main points for the relevant lectures.

MISSING

PRINT

(iii) It finally includes a small scale structured interview aimed at getting feedback from the lecturers about certain parts of the analysis (see 6.3). The use of the structured interview enables us, firstly, to elicit the interpretation provided by a specialist informant (see Selinker 1979) in the analysis of paraphrase, and secondly, to verify the description adopted in regard to some types of paraphrase.

2. Non-verbal data:

It includes 80 slides and transparencies used in the lectures recorded.

The observation to be made here, is that these slides and transparencies are only used by 2 speakers. The third speaker draws his diagrams on the board, as the lecture proceeds.

1.2.2 Rationale and limitations of the present corpus

1.2.2.1 Rationale for the corpus

Three points can be made about the decisions underlying the selection of this corpus. They are:

Firstly, in order to control text type, two parameters are held constant. They are: (i) the field of discourse (see Halliday et al 19640; the nine lectures are all biology lectures; (ii) the situational setting: all lectures take place in a comparable academic setting; i.e. in a lecture room, with an audience of learners of a similar academic level etc...The control of these parameters provides solid ground to maximise comparability in the analysis of a feature, since variation can be explained as resulting from factors other than change in the field of discourse or the setting.

Secondly, biology and not any other discipline, has been selected as the topic of the lectures to be investigated for three reasons; the first reason is that as far as

the knowledge of the author of this work is concerned, biology lectures have not been the aim of a full treatment in an academic work (see 2.2). The second reason is that lectures in biology rely heavily on the use of non-verbal aids, which would provide a valuable source of evidence in the identification of properties of spoken language in the lectures. The third reason is the analyst's familiarity with the subject due to academic training up to baccalaureat level.

Thirdly, the present corpus being produced by male speakers will guard us against any ambiguity that may result from male versus female variability in speech production.

1.2.2.2 Limitations of the corpus

The size of the corpus had to be relatively small (3 sets of lectures of 43,000 words) for the following reasons. Firstly, analysing the data at the grammatical level imposes a limit. on the size of the material that can be handled manually, since every sentence, clause and phrase has to be analysed. Secondly, the analysis of low level features will involve an amount of details which would make it unmanageable if a larger corpus were to be used. Thirdly, the analysis has to be done manually, because no computer programme has as yet been developed to handle textual features such as those I propose to study.

1.2.3 Methodology of data collection

Three sets of lectures, each set including three lectures from the same lecturer have been recorded. The equipment used was a cassette tape recorder with an ordinary microphone. There have been unsuccessful attempts to use the radio microphone for good quality recording. Indeed, one of the problems encountered later because an ordinary microphone was not so suitable was that a tiny percentage

of the text was difficult or even impossible to make out because of the distance between the speaker and the microphone at certain points in the lecture. The slides and the transparencies used with the recorded lectures were borrowed, and copies of them were made. As mentioned above, this type of data would provide essential evidence for identifying features of the interaction between the text and the visual element of the lectures, in the textual realisation of two of the features to be studied,namely sub-topic shifting and paraphrase.

1.2.4 Transcription conventions

In transcribing the data, the following decisions have been made:

1. Pauses are indicated by dots. One dot stands for a brief pause, two dots stands for a long pause, three dots for a very long pause. (No instrumental measuring of the duration of pauses has been taken).
2. Features of speech such as false starts, repetition of a word, syntactic and lexical revision, interrupting structures are enclosed between braces { }.
3. A dash following a word indicates a broken word.
4. A blank between brackets indicates that a word or a short phrase is unintelligible (as in Montgomery, 1977). This does not usually exceed three to four words. However, there is one case where the text is interrupted for five seconds, (Appendix (1) page 289), because of a technical failure of the recording equipment.

Texts	Number of Lectures	Number of Words	Number of Non-Verbal	Name of Lecturer	General Topic	Date of Recording	Cassette Mark	Additional Data
Set 1	3	14,200	28	Dr Perris	Endocrino -logy	4.5.84 11.5.84 18.5.84	A1 A2 A3	- Notes from lecturer structured interview
Set 2	3	14,200	0	Dr Hughes	Plant biology	16.5.84 16.5.84 23.5.84	B1 B2 B2	
Set 3	3	14,600	52	Dr Arm-strong	Plant biology	30.4.84 22.5.84 29.5.85	B2 C1 C1	- Notes from lecturer structured interview
Total	9	43,000	80	All at Aston Univer-sity				

Table (1) : General description of the corpus used in this study

CHAPTER TWO

LECTURE DISCOURSE: A VARIETY OF SPOKEN LANGUAGE. THEORETICAL BACKGROUND AND THE PRESENT APPROACH

2.0 Interest in lectures has various origins. This fact is manifested in the different disciplines which have concerned themselves with lectures. They range from anthropology (Goffman 1981), to education (Dudley-Evans & Johns 1979; Chaudron and Richards 1986; Sawa 1985, among others) to applied linguistics (Cook 1975; Lilley 1975; Montgomery 1977; Murphy and Candlin 1979).

None of the applied linguistic studies offers an investigation of the discourse of lectures in relation to spoken versus written language.

This chapter will first review existing approaches to lectures. It will then undertake a discussion of the theoretical issues of variety of language, and the spoken versus written language dichotomy. In the context of this theoretical background, the present approach will be presented.

2.1 Non-Linguistic interest in lectures

2.1.1 Anthropological approach to lectures

As pointed out above, lectures have been looked at from different points of view. Although the main thrust of the research into lectures is educational, anthropology has also concerned itself with lectures. Goffman (1981) studies the lecture as a "form of talk" in terms of the physical, social and cultural environment in which it occurs. From this point of view a taxonomy of lectures as types of social events is provided (Luncheon 'talk', closing/opening college lecture, information transfer lecture). Goffman relates different types of participants (i.e. audience/speaker) with each type of lecture. He also provides insightful

observations about stylistic characteristics of lectures distinguished according to their mode of delivery, i.e. whether the lecture is memorized, read aloud or delivered as a fresh talk (ibid: 172) thus he notes the presence of features of spoken language such as restart, repetition, ...etc in the "fresh talk" type of lecture (ibid: 188). However, his classification and observations are not supported by any linguistic data.

2.1.2 Educational approach to lectures

There are two main trends within the educational approach to lectures. Firstly, there is the trend of studies concerned with lectures as a teaching method. Thus Bligh (1972) provides a summary of findings for improving lecturing. Brown (1972), on the other hand, offers a broad description and classification of lectures in terms of "the classical, the problem centred, the sequential etc. Within this classification, the science lecture is said to generally belong to the first category, i.e. the classical method. However, the classification is not based on a linguistic analysis of lectures, but on a general parsing of their content structuring.

Secondly, there is the trend of studies whose main concern is pedagogical. Lecture discourse is investigated with the view of identifying areas of difficulty in lecture comprehension, mainly with L2 learners, and devising appropriate teaching materials and curriculum (Dudley-Evans and Johns 1979; Chaudron 1983, Long 1985, Kelch 1985, Chaudron and Richards 1986, Sawa 1985). Most of these studies identify a given textual feature of lectures but they do not provide a textual analysis of the feature being studied. For example, Chaudron (1983) and Sawa (1985) have examined features of discourse organisation, namely signalling of topic (Chaudron) and repetition and paraphrasing of information (Sawa) and their effect

on recall and free recall. But their study is not based on a textual analysis of these features.

Chaudron & Richards (1986) have investigated the effect of discourse "markers" on facilitating lecture comprehension. These markers range from "logical connectors" (because, no) to "temporal links" (they, at that time), intersentential connectives and utterances such as "let's go back to the beginning" referred to as "macro-markers" (ibid: 115). However, their paper does not go into a linguistic analysis of these discourse markers.

2.2 Linguistic interest in lectures

Almost all of the linguistic studies of lectures are based on a communicative approach to discourse (Lilley 1975; Cook 1975; Montgomery 1977; Murphy and Candlin 1979). Lilley, Cook, and Murphy and Candlin, following Cook, explicitly endorse Widdowson's view that discourse structure must be dealt with in terms of the communicative function assigned to the utterance (Widdowson 1973). They also apply an adapted version of the descriptive apparatus developed by Sinclair and Coulthard in the analysis of classroom discourse (Sinclair and Coulthard 1975). The descriptive apparatus of Sinclair and Coulthard is itself based on Halliday's "scale and category" proposals developed in his highly influential paper on grammatical structure (Halliday 1961). The Sinclair and Coulthard model is a strict descriptive model which accounts for all the data at different levels of "delicacy". In applying it to the discourse of lecture monologue, these studies have retained the use of a rank scale. Thus, lecture monologue structure is seen as being of a hierachical nature, consisting of a finite number of ranks, each rank being made up of units from the rank immediately below.

The descriptive apparatus in Montgomery , and Coulthard & Montgomery (1977, 1981) consists of four ranks; they are (compared with Sinclair & Coulthard's):

Montgomery	Sinclair & Coulthard
lecture	lesson
episode	transaction
period	exchange
member	move
	act

Starting from the highest rank of lecture, the next rank below is episode, which is composed of periods. A period is in turn composed of members. The lowest rank of member conflates the ranks of both move and act found in Sinclair and Coulthard's model. The "transaction" and "exchange" ranks in Sinclair & Coulthard emphasize the interactive nature of classroom discourse; while the terms "episode" and "period", in Montgomery, appear instead to indicate that this type of interaction is not assumed to be a feature of lectures. The interactive aspect of lectures, being different in nature, is captured by Montgomery in a different way.

Montgomery distinguishes between "main" and "subsidiary" discourse, a distinction which is derived from Sinclair's theoretical model of "planes of discourse" (1981), whereby "language in use" is described in terms of two planes of discourse, called the "autonomous" and the "interactive plane" of discourse. The autonomous plane has to do with the content of the message; the interactive plane, with the way in which the message is conveyed (see Sinclair 1981:)[1]. In Montgomery "main discourse" is used to refer to Sinclair's "autonomous plane", while "subsidiary discourse" refers to the "interactive plane" of discourse.

Thus, in his model of analysis, the lowest unit in the hierarchy, i.e. the member, is seen by Montgomery as realising a function which mainly belongs to either main or subsidiary discourse. As Coulthard & Montgomery put it: "some discourse members are mainly oriented towards the subject matter of the discourse, others towards the reception of this subject matter" (Coulthard & Montgomery 1981: chap 1 : 35). Within this distinction, members are further classified in terms of the function they are described to realise on each plane of discourse[2]. The point to retain in relation to this distinction of functions on the two planes of discourse is that it ultimately captures certain interesting features of lecture discourse, relevant to the present study. Namely, what Montogmery has called prospective focus, retrospective focus and restatement. The following two excerpts from Montogmery, are instances of utterances whose "function:" is that of prospective and retrospective focus, respectively:

> "I'm going to talk about the electronic circuits":
>
> "and so our general amplifier will have two properties a voltage gain and a current gain"
>
> (1977: 8-9)

the analysis, however, does not go beyond a mere labelling of acts. It does not, for instance, provide any systematic description of the nature and textual realization of each "function".

If we move to Cook's study of lectures (1975), we find that his descriptive apparatus consists of five ranks. These ranks are lecture, exposition, episode, move and act. The choice of the terms "exposition" and "episode" in the hierarchy,where in Sinclair & Coulthard's model of classroom interaction "transaction" and "exchange" appear, point to the non-interactive nature of lecture monologue in the sense of speech exchange as found in classroom interaction (see Murphy & Candlin 1979: 16). As seen above, this aspect of the descriptive

apparatus when applied to lectures, appears later in Montgomery (and in Coulthard & Montgomery).

The particular interactive nature of lectures is captured, by Cook, by a category of functions which either relates to the "organisation of information" (ibid: 55), or which under the label of "manipulative functions" relate to "the exploitation of levels of formality" (ibid).

The interesting aspect of Cook's analysis, followed later by Murphy & Candlin (1979), is its concern to account for the way in which the rhetorical organisation of the lectures is signalled. Both Cook, and Murphy & Candlin, are concerned with how the lecturer carries out the signalling of the organisation of the message he is transmitting, by using signposts to shape and direct audience (i.e. student) expectations. However, it must be pointed out here, that the purpose of both analyses is ultimately pedagogical as they are to be used as a basis for setting up ESP (English for specific purposes) teaching materials for L2 Science students. Their particular concern with the signalling of the organisation of information in lectures stems from the assumption that its aim is to facilitate students comprehension of the lectures.

Cook examines, among other things, the functions of connectives and other devices which serve as indicators of topic continuation. Murphy & Candlin identify a number of markers of the rhetorical organisation of lecture discourse, including what they refer to as "metastatements" (such as "I want to mention two types of generators"); "starters" (such as "well now, let's go on with") and "markers" (such as "well now") (1979: 12).

Another aspect of Cook's analysis which deserves mentioning in relation to the present work is his general discussion of lectures as a "category" of spoken

language (1975: 45). He makes a few interesting suggestions such as that lectures should be "arranged on a scale of relative formality ... with at one extreme the lecture read entirely verbatim from a written script, and at the other extreme, the lecture given spontaneously in a very informal setting" (ibid: 45-46). He further notes the occurrence of "repetition, reformulation, parenthetics.." (ibid: 51), as features of spontaneity in categories of lectures other than the read aloud lecture. However, his discussion does not go beyond these observations. Rather, to meet the general pedagogical purpose of his study which is to provide a basis for designing teaching materials, one lecture of "standard level of formality" seen as most representative of the style of lectures at large is selected as data base for his analysis. Finally, we must point out that Cook's descriptive model has been criticised on the basis that it does not satisfy the very pedagogical purpose it has set out to fulfill. As Murphy and Candlin have pointed out, Cook's sub-classification of moves has resulted in a proliferation of acts which makes the model "over-elaborate to serve as a material designer" (Murphy and Candlin 1979: 19).

Finally, in his paper on a science lecture, Lilley (1976) attempts to deal with the structure of lecture discourse. However, although his analysis fully espouses the theoretical frame underlying Montgomery, Cook and Murphy and Candlin studies, it does not provide a model of description of the kind developed in these studies. Starting from the assumption that the information patterning of science lectures "reflects the structure of scientific-thought and methods" (1976: 7), he describes the structure of lectures as consisting only of two parts: "problem identification" and "proposed solution". Then each part would consist of one or more macro-acts (or moves) and each macro-act is realised by one or more micro-acts. However, no explicit criterion is provided for identifying the stages or for allocating rhetorical categories to utterances. Further, Lilley's analysis is based on

a "mini-discourse" examination of a short video recorded lecture, assumed to be representative of expository discourse. The applicability of this classification to science lectures at large remains to be proved.

The major criticism addressed to all these studies, is that they apply speech act theory which has been strongly undermined for a number of reasons (see Levinson 1983). Among the more important are: that the interpretation of the illocutionary force of an utterance is generally based on intuition. The theory does not provide an explicit and objective basis for the allocating of rhetorical categories; nor does it provide any principled basis for evaluating alternative schemes (Johns-Lewis & Skelton 1987: 6). The model imposes limits on the number of speech acts that can occur in a given text type (Levinson 1983: 280).

In relation to the present work, the application of speech act theory to analyse our data would yield a number of problems which have been identified in the use of speech acts to deal with language data. Thus, if we consider a feature such as paraphrase, which is textually realised in this data either by a short utterance such as:

(i) so they will freely. diffuse. into the cell...

(130049)

or a long stretch of text:

(ii) so what I'm trying to say to. you is that. in the control of something. apparently as trivial. as blood glucose levels. we have insulin. we have glucagon. we have adrenalin. we have ACTH. we have growth hormone. and indeed some of the releasing hormones from the hypothalomus. all acting. together. in. concert. in sequence. in a cascade type response. to control. this simple process. of er. this simple. parameter. blood glucose.

concentration... so although they may be secreted independently of each other. often they act in collaborative fashion.

(129419)

The immediate difficulty we face is how to handle extracts such as (ii). How many acts does the utterance realise? On what basis shall we label these acts? Would the description of this instance be applicable to the rest of the data? Indeed, one of the criticisms Cook's analysis, for example, has provoked is aimed at the criteria he uses to identify categories of moves (see Montgomery 1977: Murphy and Candlin 1979).

We suggest that lectures discourse must be approached in a way different from this theoretical orientation which draws on speech act theory. In the next sections, the theoretical foundation of the present work is laid down. For this purpose, existing approaches to language variety are first reviewed then the theoretical issues of speech versus writing and discourse overall structure are discussed.

2.3 Approaches to language variety

Two major approaches to the description of language variety can be identified: they are the rhetorical approach and the non-rhetorical approach, although, in a few cases features from the two approaches are conflated.

The rhetorical approach concentrates on the communicative value of the utterance as the focus of linguistic analysis. Within this approach, one finds a tradition which has been established through the works of Selinker, Trimble and Lackstrom (1970, 1973, 1977) on scientific discourse, and which is known as the "grammatical-rhetorical approach" to discourse analysis.

Within the rhetorical approach a second and more important tradition can be identified in works which draw on a combination of theoretical frames in using Austin's (1962) work on speech act theory, Gumperz and Hymes, and Hymes' (1964, 1972) work on "ethnography of speaking". The works of Widdowson (1973-1978) Sinclair & Coulthard (1975); Sinclair (1980); Cook (1975); Murphy & Candlin (1979); Candlin, Burton and Coleman (1980) can be seen to belong to this tradition which applies speech act theory and which takes the view of discourse as interaction.

In the non-rhetorical approach, a linguistic feature is correlated with a linguistic notion such as "text", as found in the works of De Beaugrande and Dressler (1981); a mental frame, as in Van Dijk (1977); or extra linguistic factors such as participants in the language event, the spoken and written medium etc..., as in Halliday et al (1964). Within this approach one can also find somewhat idiosyncratic treatments of variety in the works of linguists such as Crystal & Davy (1969) and Longacre (1983).

2.3.1 The rhetorical approach to variety

As seen above, one can distinguish two major traditions here (1): the "grammatical-rhetorical" and (2) the rhetorical approach.

2.3.1.1 The grammatical-rhetorical approach to variety

The central concern in this type of approach is, as Lackstrom, Selinker and Trimble state it: "to investigate the relationship between grammatical choice and rhetorical function in the written English for Science or Technology (EST)" (1973:1). Thus the aim is to investigate the relationships between aspects of grammar such as tense, the use of articles and their rhetorical function. The choice of simple present, for example, is found to depend on the notion of degree of generality in EST.

In the same vein, Swales (1974) has investigated the use of attributive en-participle in chemistry texts. He has found that "a given" in a phrase like "a given experiment" has two principal functions, one for clarifying the "status" of the sentence, the other for specifying the "determiner range of the NP" (1974). Which of these functions is selected is determined by whether the author is "exemplifying" or "generalising".

This concentration on low level specific features, satisfactory for Lackstrom et al or Swales pedagogical enterprise, would be of no help when discussing the overall structure of the text (as Swales found in his 1981 mimeograph).

2.3.1.2 The rhetorical approach to variety

This is found in studies which have applied speech act theory to the description of discourse such as Sinclair & Coulthard's work on the discourse of classroom interaction (1975). As seen in (2.2), their work has inspired a number of studies of lecture monologue.

The application of speech act theory is also found in Widdowson (1973 - 1978), who views discourse as describable only in terms of the communicative function of the utterance. On the other hand, conversation analysts have concentrated on the type of utterances that are used to realise certain functions (which imply a kind of speech act function), such as those of "opening up closing" (Sacks, Schegloff and Jefferson 1973) in speech exchange situations. It is acknowledged now, that the application of speech act theory to discourse is problematic. (See Section 2.2 for major criticisms).

2.3.2 The non-rhetorical approach to language variety

One approach to variety which can be described as non-rhetorical is what is known as "text-linguistics", an approach developed by De Beaugrande and Dressler (1981) and Van Dijk (1972-1977).

De Beaugrande and Dressler identify seven parameters in the study of variety or "text": (1) cohesion; (2) coherence; (3) intentionality; (4) acceptability; (5) informativity; (6) situationality and (7) intertextuality.

De Beaugrande and Dressler's discussion of "situationality" is textual rather than social, although participants are included. Studying the situation in a sixty lines text, they bring down the situation to 12 "participants' common strategies" (De Beaugrande & Dressler (1981: 173). This could be seen as the study of textual content in relation to participants. The parameter of intertextuality on the other hand, is most relevant in the comparison of different texts, i.e. how one text makes use of another text via reduction, summarising, destructuring. Important as it is, this type of text treatment falls well outside the concern of the present thesis in which the prominent question of speech versus writing and text structure are studied within the notion of text/variety.

An idiosyncratic theoretical frame for "discourse typology" is provided by Longacre (1982) who classifies "discourse types" into four broad categories which he calls: narrative discourse, procedural discourse, behavioural discourse and expository discourse. Each type of discourse is further described by a set of binary features. Thus, each type of discourse later takes on the plus or minus value in regard to the features of "projection" (which has to do with a situation or action which is contemplated, enjoined, or anticipated, but not realised") and

"tension" (which has to do with whether a discourse reflects a struggle or a polarization of some sort") (Longacre 1983: 4/6).

In this framework, the data under investigation in this thesis would be classified as expository discourse (minus projection, minus tension). However, Longacre's classification is at such a level of abstraction that it is hard to see how to proceed from it in the analysis of a type of discourse.

At the other extreme, there is the detailed ethnomethodological approach to variety expanded by Dell Hymes's work (1972) on the ethnography of speaking, where seven factors are set up to describe language variety, which are (1) speech community (2) language fields, (3) speech event, (4) speech act, (5) speech style, (6) way of speaking and (7) components of speech. The last factor, "components of speech", comprises sixteen components: (1) message form (how something is said), (2) message content (topic), (3) setting (time, place and general circumstances), (4) scope (psychological setting), (5) speaker or sender, (6) addressor, (7) hearer or receiver, or audience, (8) addressee, (9) purposes outcomes (conventionally recognised and expected outcomes), (10) purposes goals (the purpose of an event from community standpoint, (11) key (tone, manner in which the act is done), (12) channels (medium of transmission), (13) forms of speech (verbal processes of a community), (14) norms of interaction (specific behaviours and properties that attach to speaking), (15) norms of interpretation, (16) genres (categories such as poem, myth, tale etc.) (Hymes 1972: 52-65).

Firstly, we notice the inclusion of 'speech act' as one minor component in Dell Hymes' inventory. Secondly, in relation to the present work, we note the importance of the components of 'audience' (that it is, in our case,an audience of

learners), purposes and outcomes (that lectures are a variety of didactic discourse), setting (that it is a pedagogical setting) and channels (that the medium used is the spoken medium). The most important component as far as this work is concerned is Dell Hymes' 'channel'. As will be seen in (2.6) the main task of this study is to investigate features of spoken as opposed to written language in biology lecture discourse, through the realisation of three textual features.

Consideration of spoken versus written mode of transmission, is also a component in the description of language variety as a "register", which is now mainly associated with the name of Halliday (1964, 1978).

This approach to variety has its origin in the works of J R Firth (1957) following Malinowsky (1930) who has laid the foundation of the view that the linguistic system is of a social nature. Building on this view, Halliday and his associates have devleoped a model of analysis of language variation: language varies in relation to the social context in which it is used. Halliday et al (1964: 87) write: "Language varies as its function varies; it differs in different situations. The name given to a variety of a language distinguished according to use is register".

Variety characteristics are identified on the basis of three situational and contextual dimensions in terms of "field", "mode" and "style" of discourse (ibid: 90). Field of discourse "reflects what is going on" (ibid: 90); mode of discourse "refers to the medium or mode of language activity" (speaking or writing) (ibid: 91) and style of discourse "refers to relations among the participants": (ibid: 92). The last dimension has been reformulated later by Halliday. Thus "style of discourse" is reformulated as "tenor of discourse" (1978: 62).

The three dimensions set up by Halliday have been redefined and elaborated on by Ure and Ellis (1969 ; 1977) Crystal & Davy (1969); and Gregory (1967) and Gregory & Carroll (1978).

Ure and Ellis (1977 : 198) propose four "situational parameters", namely (1) subject matter, (2) medium (spoken/written) personal and social relations between speaker/writer and addressee and (4) social function of the language event, which correspond to the four linguistic features of (1) field, (2) mode, (3) tenor and (4) role of language (ibid: 200). A register is established, according to Ure and Ellis only when a "correlation" between linguistic features on one hand, and elements of situation on the other, is found.

Crystal & Davy (1969) have also redefined the parameters of variety classification. They set up their own set of five dimensions. They are: the dimensions of (1) discourse, which includes the features of medium (speech/writing) and participation (difference between monologue and dialogue), and the dimensions of (2) province (subject-matter associated with a given field such as the field of physics), (3) status (social relations), (4) modality (linguistic features correlatable to purpose of utterance) and (5) singularity (personal, occasional features) (Crystal & Davy 1969: 68-77). Two dimensions of Crystal & Davy's are of interest to us. The dimension of "discourse medium" and "singularity". Firstly, medium, "mode" in Halliday and "medium" in Ure and Ellis, refers to the speech versus writing distinction. Crystal & Davy rightly make the point that this distinction is not always clear cut: a spoken variety can make use of features of written language and vice versa. They use a qualifier with "medium" which is then either simple medium (to yield features of clear cut variety of spoken or written language), or complex medium (to describe varieties which make use of features of both speech and writing). It would seem that "complex medium"

is better suited to describe lecture discourse. Secondly, the dimension of singularity, which does not exist in Halliday, or in Ure and Ellis, can apply to account for linguistic idiosyncratic features. The inclusion of the dimension of singularity as a descriptive category of variety will capture, within the framework of variety description, common as well as idiosyncratic features of a given variety. It remains true, however, that singularity is a dominant and desirable feature only in a limited number of language varieties, primarily literature.

As far as the present data are concerned, it is important to examine whether variation of linguistic features is due to different individual styles. It should be mentioned in this connection, that in Crystal & Davy's analysis of lectures as a variety, the question of singularity is not dealt with.

Gregory & Carroll (1978) are most relevant for the present concern with speech and writing because they offer the most elaborate expansion on the parameter of "mode of discourse" which they label "modes of discourse". (We may note that the plural form is used with all three parameters: "fields of discourse", "tenors of discourse"). They start by making the point that the distinction between speech and writing is not a simple one; there is a wide scope of variation between these two "polar" modes of discourse. They also draw a distinction between monologue and conversation (1978: 40). They define monologuing as:

> "the users' medium relationships in those speech situations in which the other people present (if any) do not join in or at least are not meant to..."
>
> (ibid: 40)

this is certainly characteristic of lectures where the speaker indisputably holds the floor. (see also Goffman 1981: 165).

Features of the mode of a variety will depend on whether "the text.. is written to be spoken as if not written (e.g. (recited) script), "written to be spoken" (many political speeches and lectures), "written not necessarily to be spoken" (stories: telephone directories) (ibid: 42). Interestingly, they state that "there is a cline of more to less spontaneity": (ibid: 40). Thus, using their idea of a cline, we can observe that more spontaneity places the variety nearer to the spoken end of the cline, less spontaneity places it nearer to the written. Applying this to lecture discourse one of the concern of the present study, is where on the cline of more or less spontaneity do biology lectures stand? are they nearer to the spoken or to the written end of the cline?

The next section explores the research and findings which have concentrated on comparing the two modes of language.

2.4 Speech versus writing

As seen in the preceding section, speech versus writing is referred to as "channels" in Dell Hymes, "mode of discourse" in Halliday, "modes of discourse" in Gregory & Carroll, "medium" in Ure & Ellis and Crystal & Davy.

It has has always been implicity or explicitly maintained that speech or writing can be analysed. Dell Hymes (1972) for example, writes that "for particular purposes - any one aspect (i.e. any component among the components of speech) can be segregated for analysis" (ibid: 23).

Speech versus writing has been the object of interest and study in linguistics as well as in other disciplines such as education, psycholinguistics, anthropology. Each discipline has approached the dichotomy with its own concerns and questions. Thus psycholinguistics is interested in the acquisition of reading/writing skills

(Brown and Bellugi 1964); Sociolinguistics, in some aspects of speech, reading and writing focusing on their context and the influence of these contexts on participants or learners (Cook-Gumperz & Gumperz 1981); on the other hand, linguistics has been concerned with speech/writing in various ways. The notion of "primacy of speech" has ever been quoted in the linguistic definition of language (Strang1962; Robins 1964; Hockett 1958). Within the concern of distinguishing different language varieties, lexical and grammatical features are associated with either speech or writing. (Ure and Ellis 1977; Halliday 1979).

Recently, discourse analysis has brought to the forefront the question of the relationship between speech and writing in its investigation of language use. Extending the domain of linguistic analysis beyond sentence level, discourse analysis studies samples of spoken as well as written language drawn from performance data. The emphasis is on differences as well as similarities between spoken and written language. The type of language data which has been mostly used is narrative discourse, which can be narrative about a film (Kroll 1977; Chafe 1979; Tannen 1982; Beaman 1984; Farag 1986); Story-telling in conversation (Tannen 1982); personal narrative (Tannen 1982). Conversation has also been used as a type of language data (Chafe 1982). Although the findings of the different studies are not always congruent as to the nature of the differences between spoken and written language and the extent to which these differences are found, it is generally argued that there are certain characteristics of spoken language as opposed to written language. Before discussing these findings, a number of distinctions which have mostly emerged out of the study of spoken versus written language and which are more delicate than the primary distinction of spoken/written language, are reviewed.

2.4.1 Spontaneity versus forethought

One of the major distinctions which have appeared in the literature on discourse is the notion of spontaneity as opposed to non-spontaneity or forethought. This distinction ultimately cuts across the spoken versus written distinction.

Spontaneity versus non-spontaneity was first used in Gregory & Carroll (1965; 1978: 38), to account for variation in modes of discourse, although its application is restricted to the spoken mode of discourse.

A parallel distinction which expresses the same idea and which has now gained a lot of currency, is that suggested by Ochs (1977; 1979), namely planned versus unplanned discourse. Ochs (who incidentally does not focus on the spoken versus written language distinction) defines unplanned discourse as "discourse that lacks forethought and organisational preparation", and planned discourse as "discourse that has been thought out and organised (designed) prior to its expression" (1979: 55). Thus, we understand that discourse differs according to the amount of planning time available for the language user. Speech or writing can be planned or unplanned. If planned they share features which would be lacking in unplanned speech or writing.

Lecture monologue seems to be describable by a complex combination of the two notions of planned and unplanned discourse. Ochs refers to lectures as planned unplanned discourse (Ochs 1979: 77).

Darian (1983) classifies lectures as semi-planned, as they generally undergo a certain degree of planning. (This does not apply to the lecture which is read aloud). Characteristics of planned and unplanned discourse are said to be

manifestations in lectures of the switching of levels of formality (ibid: 33), as observed in the use of lexical items, asides and digressions (see also Cook 1975).

2.4.2 Fragmentariness versus integration and involvement versus detachment

Chafe (1979; 1982) working within his cognitive approach to linguistics, identifies two essential differences between the processes of speaking and writing:

> "That speaking is faster than writing". "That speakers interact with their audiences directly, whereas writers do not".
>
> (1982: 39)

Consequently, Chafe argues speaking shows a "fragmented" nature as opposed to the "integrated" characteristic of writing where the writer has the time to "integrate" his or her ideas into a complex, coherent whole; speaking shows an "involved" quality as opposed to a "detached" quality found in writing.

Drawing on an experimental investigation of four "styles" of language (dinner table conversation, lectures, letters and academic papers), Chafe works out a set of devices identified in the different "styles", which reveal these features of spoken as opposed to written language. These devices are presented in Table (2) overleaf:

Integration	Involvement
Nominalizations Past/present participles Attributive adjectives Conjoined phrases Series (of noun/verb phrases) Sequences of prepositional phrases Complement clauses Relative clauses Introductory phrases (e.g. "in view of this") Adverbial conjunction	First person references Speakers' mental processes (e.g. "I think") Monitoring information flow Emphatic particles (e.g. "just", "really") Fuzziness, vagueness, Hedges Direct quotes
Attachment	**Fragmentation**
Passives Non active subjects Precision Indirect quotations	Co-ordinating conjunction Afterthoughts Repetitions False starts

Table (2) : Chafe's features of spoken and written language

The four features of integration, fragmentation, detachment and involvement form the poles of two intersecting continua on which various types of discourse can be placed.

In connection with the present work, it would be interesting to see which devices are exploited in biology lectures and consequently which features would describe biology lecture discourse. It must be pointed out that although Chafe uses lectures as part of his spoken data, the results of his analysis apply only to the "two maximally differentiated samples: spontaneous conversational language on the one hand and formal academic papers on the other" (Chafe 1982: 49). His data have been later criticised (see 2.4.5).

2.4.3 Lakoff's grid

Lakoff (1973) proposes a set of six criteria to be used as binary features in investigating samples of discourse. According to her, any type of discourse (spoken or written) is to be classified along the criteria of: "visibility", "reciprocity", "informality", "spontaneity", "empathy" and "inconsequentiality".The theoretical basis underlying these criteria is that, in Lakoff's terms, there is "a continuum of discourse, arranged as to the purpose of the discourse and the environment in which it occurs" (Lakoff 1979: 23). The continuum over-rides the spoken versus written dichotomy, which Lakoff argues, is a false dichotomy.

If we apply Lakoff's criteria to the present data, biology lectures would have the characteristics of: plus visibility (speaker/audience are both present in the situation); plus/minus informality (assuming that there is generally a shift of levels of formality); plus/minus spontaneity (features of spontaneous/non-spontaneous discourse can be found), minus empathy, and minus inconsequentiality (lecture discourse, as didactic discourse, is primarily interested in the message). The following table summarises Lakoff's criteria and shows how they would apply to the present corpus :

Criteria	Prototype of speech (spoken oral dyad)	Biology lecture discourse	Prototype of writing (written expository prose)
visibility	+	+	-
reciprocity	+	-	-
informality	+	+/-	-
spontaneity	+	+/-	-
empathy	+	-	-
inconsequentiality	+	-	-

Table (3) : Lakoff's six criteria for classifying types of discourse

2.4.4 Major features of unplanned spoken versus planned written discourse

A number of characteristics which researchers generally argue belong more to one mode of discourse or the other have been established. These characteristics can be seen to be either syntactic, or organisational.

(i) Syntactic characteristics of unplanned spoken as opposed to planned written discourse

Unplanned spoken language is found to make use of the following features:-

1 - Repetition (lexical or syntactic) (Chafe 1982)

2 - Ellipsis (Darian 1983)

3 - Subject omission (Kroll 1977)

4 - False starts (Chafe 1982; Darian 1983)

5 - Left/right dislocation (Brown & Yule1983; Farag 1986)

6 - Deictic usage (Kroll 1977)

7 - Fuzzy expression (Chafe 1982)

8 - Zero clause connection (Chafe 1982)

Planned written language, on the other hand, is characterised by the use of:

1 - Subordinate rather than co-ordinate constructions (O Donnell 1974, Kroll 1977, Ochs 1979; Chafe 1982).

2 - Complex syntactic and semantic structures (noun groups; nominalisation -) (Huddleston 1971; Poole and Field 1976; Akinnaso 1982).

3 - Declarative and subjunctive rather than imperatives interrogatives and exclamations (Ochs 1979; Akinnaso 1982).

4 - Higher frequency of certain grammatical features such as gerunds, participles, attributive adjectives, modal and perfective auxiliary (O Donnell 1974, Ochs 1979; Chafe 1982; Akinnaso 1982).

(ii) Organisational characteristics of planned written as opposed to unplanned spoken discourse

1 - Planned written language shows a 'reliance on the explicit organisation of information structure manifested in the use of "thesis", "topic sentence", "supporting evidence" (Olson 1977; Rubin 1980; Akinnaso 1982).

2 - Planned written language shows the tendency of being explicit, of producing complete information whereas unplanned spoken language would rely on immediate context in other extra textual parameters to complete the verbally expressed message. (Chafe 1982; Akinnaso 1982; Lakoff 1979; Darian 1983).

The characteristics outlined above are not exclusive to spoken or written language. They rather qualify the prototype of either mode.

In light of the above characteristics, the following sample from our data can be looked at:

(i) now the function of the haemoglobin. {or let's put it another way} the haemoglobin has two functions.

(230339)

(ii) erm. now let's see what's next... {so you can see that. er. both aporsity. and superfluity of. hormones can have. quite. dramatic effects.} now what I. want to do now. is tell you a little something. of the. chemical nature of hormones.

(129111)

1 - In both excerpts, the speaker announces what is coming next in the lecture (as text) i.e. the sub-topic of the coming fragment of text is announced (see Chapter Five). This may be looked at as the "topic sentence" of the new stretch of text. It can also be looked at as a manifestation of the organisation characteristic of planned written language whereby the planning of the information structure of the text is made explicit.

2 - On the other hand, the "topic sentence" is interrupted in both (i) and (ii). In (i) it is interrupted by a "metatextual comment", in (ii) by a recapitulation (underlined in the excerpt) (see Chapter Six). The recapitulation applies to the preceding stretch of text.

3 - In (i) the metatextual comment which interrupts the "topic sentence" is realized by the utterance. "let's put it another way". This can be seen as a manifestation of the switching of level of formality (from formal to informal

style). This can also be seen in (ii) in the use of the comment "let's see what's next" which introduces the whole utterance.

4 - The use of the personal pronouns "You" and "I" in excerpt (ii) refers to the participants in the lecture (as language event), i.e. audience and speaker.

The feature described in (1) is a normal feature of planned written discourse, whereas features in (2), (3) and (4) are found in unplanned spoken discourse. These excerpts (i and ii) exemplify the type of situation presented in the lectures which form the data base of this study. In one utterance features of both planned written and unplanned spoken discourse are made use of. If both types of features are found in the lectures, the question which remains to be answered is where on the continuum these lectures can be placed? Are they nearer to the planned written end of the continuum or the unplanned spoken end? What type of features is mostly favoured? Is it possible to come to a clear identification of the discourse of biology lectures along the planned written/unplanned spoken language continuum?

One way of tackling this question is to look at complexity (a characteristic of planned written language) as an index for identifying where lecture discourse belongs.

2.4.5 Grammatical complexity

We have seen earlier that it is generally agreed that planned written language uses more complex grammatical structures (such as subordination) than unplanned spoken language - (Kroll 1977, Chafe 1982; Milroy & Milroy 1985).

A number of experimental studies concerned with differences between spoken and written language, have used grammatical complexity as measured in the use of

subordinate as opposed to co-ordinate constructions in selected samples of spoken and written discourse, as a test for higher complexity in the written as opposed to the spoken samples investigated(3).

Thus Kroll (1977) provides evidence supporting the general view that written language uses more subordination whereas spoken language uses more co-ordination. Her findings are that written narratives contain 35 per cent of subordinate/dependent structures, whereas spoken narratives contain only 14 per cent (ibid:98).

Chafe (1982) comparing the structure of "idea unit" (a unit which exhibits one of a small set of syntactic structures, characterised by a coherent unit of intonation contour and bounded by pauses), in the speech of fourteen subjects (Faculty and graduate students), has found that one of the devices which contribute to the "integrated" nature of writing is the favoured use of subordination as opposed to co-ordination which is more used in speaking.

O' Donnell 's (1974) findings go along the same lines. A significantly greater number of dependent clauses are found in written discourse.

However, conflicting results have also been the outcome of other experimental studies, also concerned with spoken versus written language differences. Thus, Poole and Field (1976) argue that spoken discourse shows a significantly greater degree of subordination than written discourse. With these findings, they provide support for Halliday's view that speech rather than writing is the more complex mode of discourse: "written language also tends to be simpler than spoken language in its grammatical organisation; speech, especially informal speech such as casual conversation, displays complexities of sentence structure that

would be intolerable (because they would be unintelligible) in writing" (Halliday 1978: 224).

A study which partly supports Halliday's view is Beaman's (1984). Comparing spoken and written narratives of a short movie shown to forty subjects, she argues that "spoken narrative is on the whole just as complex as, if not more in some respects, than written narrative" (Beaman 1984: 75). She found that spoken narratives, on the whole, used more subordinate/dependent clauses than written narratives (ibid: table 19 p.75). However, an analysis of the types of subordinate clauses used in each medium enables Beaman to show that some types of subordinate clauses are more frequent in one "modality" or the other. Thus written narratives favour the use of non-finite nominal clauses, which are complex syntactic constructions, spoken narratives use more restrictive relative clauses, which according to Beaman are less complex constructions than the participle or ing-clause for example (1984: 68). Beaman concludes that, if subordination implies complexity, neither speech nor writing is more complex. Rather, the two "modalities" differ in the type of complexity which is manifested since each "modality" favours certain types of subordinate clause. However, a closer look at her analysis reveals that subordination is not carefully defined.

These conflicting results, have led linguists to reflect upon these various studies and question aspects of the methodology which has been used, such as:

1 - the data base used: some studies have not controlled certain parameters which can influence language performance; thus, Akinnaso (1982) writes in relation to Chafe's study:

"differences in setting,context and purpose between formal (published) academic papers and informal conversation are in themselves variables that may affect lexical and syntactic choices:

(Akinnaso 1982 : 108)

Beaman, on the other hand, observes that: "the differences between spoken and written language that many linguists in the past have found may be more a result of differences in the formality and purpose of discourse than in spoken versus written distinction (1984: 46).

2 - There is no general concensus on the nature of the unit of analysis used. O' Donnell used the "T-unit" (one independent clause and the dependent clauses, if any, syntactically related to it). (1974: 103), Chafe, Kroll, the "idea unit" (a cognitive category identifiable on syntactic and phonological grounds); Horowitz and Newman, also used the "idea unit" (but with them it is a purely cognitive category).

3 - "Subordination" as an indicator of complexity is approached in different ways in the various works. (Farag 1986; Johns Lewis 1987).

4 - "These studies being all quantitative "each researcher decides on what to count and how" (Akinnaso 1982: 109).

The point which can be made in relation to work on complexity here, is that the discrepancies found between the findings of these studies have not been attributed to the original hypothesis as such, namely that complexity is typically a feature of writing. Its validity is not questioned. What has been questioned is the methodology used in verifying this hypothesis.

The present study adopts the assumption that complexity is typically a feature of written language and that subordination as opposed to co-ordination, implies

complexity. However a different approach to complexity is used. (see 3.2). It is hoped that the application of the analysis of grammatical complexity in biology lectures will help us to identify biology lecture discourse along the unplanned spoken versus planned written language continuum.

2.5 Approaches to discourse structure

Three broad approaches to the description of the overall structure of discourse can be identified. Before we review them, it is desirable to say a few words about the use of the terms "discourse" and "text" as they are used in the literature and as they will be used in this work.

2.5.1 "Discourse" and "text"

While both "discourse" and "text" are used to refer to the level of analysis beyond the boundaries of the sentence,the two notions have been used differently by different scholars. Thus, "text" is used to refer to a theoretical notion distinguished from instances of text by European linguists such as Dressler (1978), De Beaugrande & Dressler (1981); Dressler (1978) states that textlinguistics views "spoken and written texts as the minimal free unit of language" (1978: 2). This use of text as an abstract notion is also found in Winter (1978) who writes: "we call any report, article, memo, letter extracts, etc. text "(1978: 4). On the other hand, "text" has been used to refer to non-interactive monologue, while discourse is used to refer to speech exchange situations (i.e. spoken discourse) (see Stubbs 1983 for a review). A third type of distinction between "text" and "discourse" is to be found in Widdowson (1973) who argues that text is made up of sentences and has the property of cohesion, while discourse is made up of utterances and has the property of coherence (this distinction is dropped in his later writings as Hoey has pointed out 1983: 8). Fourthly, and

finally, discourse is used to refer to an abstract notion, actualised in text. (Halliday 1964)[4], (Sinclair 1981; Sinclair & Coulthard 1975). Sinclair and Coulthard (1975) use "classroom discourse" to refer to a variety of language and text as an instance of "classroom discourse". This distinction is adopted in this work: thus "biology lecture discourse" refers to the language of biology lectures, while each lecture is an instantiation of the discourse of biology lectures; each lecture is a text[5].

2.5.2 Pragmatic approach to discourse structure

The major work which applies speech act theory to the analysis of discourse is that of Sinclair and Coulthard (1975). As discussed earlier (see 2.2), basing their model on Halliday's strict interpretation of structure as a rank scale, they set up a hierachical apparatus in which the lowest rank is the act, to describe the structure of classroom discourse.

There are three aspects in Sinclair and Coulthard's model which make it unsuitable for our purpose. Firstly, as mentioned earlier, it is a strict mode of description in that all the data have to be accounted for. The present study is not primarily concerned with the overall structure of lectures; the description it provides is loose in that it is not concerned with accounting for all the data, but rather aims to examine the role of the features under study in the overall structure. Secondly, their model applies speech act theory which has exposed their minimal unit. i.e. the act, to severe criticisms. Thirdly, even assuming that their model of analysis is successful in dealing with monologuic discourse in lectures (as claimed to be by Cook, Montgomery and others), still, for the purpose of this work, it would be unsuitable because it will force us to go to a low level of delicacy, taking "act" as the minimal unit of analysis.

Widdowson (1973; 1978) also applies speech act theory, but in a different descriptive framework. As mentioned in 2.3, he argues that coherence[6] which underlines textual patterning is carried by the illocutionary force of the utterance. Three points can be made in relation to Widdowson's interpretation of discourse structure. Firstly, all the criticisms which have been addressed to speech act theory, remain valid in relation to Widdowson's view of coherence. Secondly, the data he uses are usually conversational and do not exceed a few short utterances such as the following:

A - What are the police doing
B - I have just arrived

(1978: 28)

Thirdly, Cook's application of the speech act on the structure of a monologue has run into considerable difficulties (see 2.2). Moreover, in the context of the present work, it would seem that if speech acts labels are attached to a stretch of text such as a summary paraphrase, no insights into the overall structure are gained by the exercise.

2.5.3 Schema as cognitive approach to discourse structure

Schema theory is "basically a theory about knowledge, a theory about how knowledge is represented and how that representation facilitates the use of knowledge in particular ways" (Rumelhart 1984 : 2).

Three broad Schema approaches to discourse structure can be identified : Schema as social structure, Schema as cognitive structure and Schema as text structure. Schema as social structure highlights the social context of discourse interaction (Cicourel 1980; Chafe 1977). The knowledge of a particular social

Schema i.e. a particular context of interaction is brought into play in interpreting and anticipating on-going discourse.

Schema as cognitive structure postulates a cognitive representation of situations, actions, sequences of actions, events, which is activated in processing discourse. (Rumelhart 1977; Van Dijk 1977). Van Dijk, for instance, uses "office frame" (frame is his term for Schema) as an example of cognitive representation which is activated in the reading of an extract from a "crime story". He explains that what makes the passage coherent is that its content conforms to the expectations brought about by the representation in the reader's mind of what an office is like, the kind of activities which take place in an office, what their sequence is and the like.

The approach to Schema as cognitive structure has been elaborated on in later studies which have introduced the theoretical distinction of culture specific Schema and formal versus content Schema (Carrell 1983). In light of the formal versus content Schema distinction, the type of Schema activated in the crime story quoted above would be in Carrell's terms "content" Schema.

Schema as text structure views discourse structure in terms of information patterning i.e. in terms of a set of categories which represent the information blocks of a given text.

The type of text structure schema which has received a good deal of attention in the last two decades is the problem-solution schema. It has been mostly applied to the narrative by a number of linguists who have either used it as it is or developed a related structure. Thus Labov (1972) identifies the structure abstract orientation - complication - action - evaluation-result or resolution-coda. Longacre (1974) similarily proposes the following pattern: aperture-setting, inciting moment -

developing conflict, climax, denouement, final suspense, closure. Van Dijk (1977) and Grimes (1975) refer both to the structure of the narrative and of scientific discourse in acknowledging the existence of the pattern problem - solution.

Hutchins (1977) building upon two patterns developed by Bremond (1970) working on French fairy tales, and Gopnik (1972) analysing three types of scientific paper, puts forward the following structure fo scientific papers.

the problem	("current' hypothesis/paradigm (demonstration of inadequacies (statement of "problem"
the solution	(statement of "new" hypothesis or (of alternative hypothesis (testing of hypothesis or of one (of alternative hypotheses
	Implications of solution

(Hutchins 1977: 31)

Winter (1978) has applied the schema of problem solution to text structure. This has been further discussed and elaborated on by Hoey (1979-1983). The application of the problem solution pattern is first discussed in Winter (1978) in a published paper. Winter uses the following constructed example:

> I was on sentry duty. I saw the enemy approaching. I opened fire. I beat off the enemy attack.
>
> (Winter 1978: 7)

which, according to him, illustrates the pattern, situation - problem - solution - evaluation. These elements are claimed to constitute the fundamental categories of information structure of science texts. (Winter's study was applied on different types of articles selected from the New Scientist magazine).

It can be argued that: Firstly, the sample text used by Winter (and Hoey, later) (sentry duty story) is constructed and thus it may be seen as designed to fit the pattern rather than the other way around. Secondly, if we move to our data, we can observe that the problem-solution schema will not be adequate to handle the structure of a biology lecture as approached in this study. It could apply to a longer text such as a series of lectures (or a thesis), but not to a single lecture. This can be seen in the fact that a series of biology lectures can conceivably be framed as starting a major question(s) and then devoting a lecture to deal with a given aspect of the question. In one lecture, however, as exemplified in set (3) Lecture (2) (Appendix (1)) we cannot find any traces of problem - solution schema. This particular lecture starts with the introduction: "okay. today. we begin our look ..... and we're going to be looking at the most primitive of them. that are. generally described as the gymnosperms", and then moves on to provide an expository description of aspects of the gymnosperms. Furthermore the narrow scope of analysis in this study motivated by the purpose of identifying speech/writing characteristics in biology lectures, would not yield a full problem/solution analysis. Finally, we can add that Hoey (1983) acknowledges different structural analysis of scientific discourse).

For the purpose of the present study, an approach which has a much narrower scope of focus is needed.

Swales (1981), investigating introductions to science papers identifies a schema consisting of four moves: (1) establishing the field; (2) summarising previous research; (3) preparing for present research; (4) introducing present research. Each move is realised by a number of categories. Thus, by way of example, the first move, the field of research is established by:

a) asserting centrality

b) stating current knowledge

c) describing key characteristics

Swales also discusses the prominent lexical and grammatical features which appear in each move of the pattern. The interest of Swales' approach to the present work is essentially in the narrow scope of focus of his analysis as it concentrates on specific low level textual features, which he calls "functions", to reveal the underlying structural patterning of the text (or in his data, part of the text as he only deals with the introduction to the science paper). As stated earlier, this narrow focussing is appropriate to the general aim of the present study to identify characteristics of speech versus writing in lectures. The following section presents the details of the present approach.

2.6 The present approach

2.6.0 The main concern of this study is the relationship between speech and writing in biology lecture discourse. For this purpose, the analysis although a textual analysis, concentrates on the low level of grammatical as well as textual features in the attempt to capture characteristics of speech as opposed to writing in biology lecture discourse, and to identify where, on the unplanned spoken versus planned written discourse continuum, these lectures can be plotted.

2.6.1 Biology lectures: planned or unplanned discourse

In the nine lectures of (43,000 words) recorded and used as a corpus in this study, the speakers use notes in lecturing, although not all of the three speakers depend on notes in the same way. With one speaker (the 3rd set of lectures) notes are almost used continuously, while with the other two speakers (1st and 2nd sets of lectures) notes are referred to only occasionally as aide memoire. If we apply

Dudley Evans & Johns classification of lectures, then lectures in set (3) would be nearest to style A (reading style):; lectures in set (2), nearest to style B (conversational style); and lectures in set (1) would combine style B and C (conversational and rhetorical style) (Dudley Evans & Johns 1979: 34). This use of notes means that all the lectures which make up the data have received some degree of planning (which varies from one set to the other).

If biology lectures have undergone a certain degree of planning, what type of characteristics, of spoken or written language, are going to be mostly exploited in the three features under analysis.

2.6.2 Unplanned spoken versus planned written discourse

Although the distinction of planned versus unplanned discourse cuts across the spoken versus written distinction (as observed earlier), it remains the case that the prototype of unplanned discourse is spoken, and the prototype of planned, discourse is written. The two form the poles of a continuum on which varieties of language would be placed according to their degree of planning. Further, in this study, when discussing a given feature of biology lectures in relation to the polarity planned written discourse is used to refer to written academic textbooks.

2.6.3 Approach to features of unplanned spoken as opposed to written discourse in the present study

2.6.3.0 Two different linguistic dimensions are investigated in this study to identify characteristics of unplanned spoken as opposed to planned written discourse in biology lectures (1) grammatical complexity; (2) study of textual features.

2.6.3.1 Grammatical complexity

This study adopts the assumption that planned written language because it has "undergone forethought and organisation prior to its expression" as it is not produced under the constraints of "real" time, shows more complex grammatical constructions than unplanned spoken language which is produced in "real" time and is not prepared prior to its production. (Chafe 1982). It also adopts the view that subordination implies complexity. Thus, grammatical complexity as measured in the use of subordination as opposed to co-ordination is used as an indicator of "spokenness" as opposed to "writtenness" in biology lecture discourse.

The first point which can be made about the application of grammatical complexity analysis to biology lecture discourse has to do with the nature of complexity. Because complexity is a feature of the sentence, sentence grammar is used to describe the data. The second point is that the analysis of grammatical complexity is used here as a methodology to determine features of unplanned spoken as opposed to planned written discourse in the data under study.

2.6.3.2 Textual features

2.6.3.2.0 The main point to be made about the use of grammatical complexity in this study, is that it applies to the sentence. In order to capture more features of biology lecture discourse in relation to unplanned spoken versus planned written discourse, certain utterance(7) features are looked at. These features can be realised in an element in an utterance (such as a connective), or in a sentence or more than one sentence (such as one topic shift signallling and paraphrase, (see 2.6.3.2.1 below). It is hoped that the investigation of these textual features will capture certain aspects of unplanned spoken as opposed to planned written discourse, which grammatical complexity cannot capture. Further, unlike grammatical complexity which looks at language internally from within the linguistic system,

the investigation of textual features places a given feature in its textual context as well as situational context of participants (audience of learners), purpose of the activity (transfer of information), setting (institutionalised academic setting). It is to be noticed here, that one of the situational factors in the lectures is the oral mode of delivery which is expected to entail some performance features such as repetition, interruption and the like.

2.6.3.2.1 Textual features investigated

As mentioned in Chapter One, three textual features are selected for investigation: sentence-initial connectives; sub-topic shift signalling and paraphrase. The choice of these features has been made on the basis of their prominence in these biology lectures.

In investigating these features, the primary source used has been the core data (i.e. transcript of lectures). Then as a second step, lecturers' notes have been used, in set (1) and set (3) specifically in the analysis of sub-topic shift signalling to verify that places where a change of sub-topic is located, is correctly identified. As a third and final resource, lecturers have been consulted about some aspects and types of paraphrase to verify that instances of certain types of paraphrase defined on the basis of the criteria set up in this study are approved by the author of the text as a specialised informant.

The following excerpt from the data illustrate the three features under analysis. (The initial utterance in this excerpt has already been quoted in 2.4.4 above).

> now the function of the haemoglobin, or let's put it another way, the haemoglobin has two functions, one is to protect. the. bacterium from oxygen. by. combining. with. any. oxygen. that is. diffused. into. the root nodule.. and since respiration is very active in these nodules. erm it's very necessary that oxygen

should. readily diffuse. into it... but in addition. to protecting the bacterium. the haemoglobin is a source. of oxygen... at low partial pressures... no oxygen is available. both to the bacterium. and to the host plant. at low partial pressures. so that the. terminal oxydes. in the respiratory chain of electron carriers. {associated with a terminal train} associated with a train of electron. carriers. can. accept. oxygen. at low partial pressures. from the haemoglobin.. is that clear... {so now I'm going to look at the} having looked at the root nodule. its structure. and one aspect of its structure. in relation to its functions. at er. what might be called. a molecular level. I now want to look at the. structure and function of. the root nodule. {at a more}. at a cellular. and anatomical level. (230339)

In two instances, the content of the coming fragment of text is announced (underlined in the excerpt), which signals a shift to a different aspect of the topic under discussion. In the second instance, the utterance which announces the new sub-topic is interrupted by a recapitulation or a summary paraphrase of the main points of the preceding stretch of text (which extends beyond the quoted excerpt). There is also an instance of another type of paraphrase, an inference paraphrase (i.e. interpreted as an utterance whose information content is implied by that of the original utterance. See Chapter Six). The first instance of sub-topic shift signalling utterance and the inference paraphrase are initiated by a sentence-initial connective. In addition, if we look at the surface realisation of the two sub-topic shift signalling utterances, we can see that in the first instance, the utterance is interrupted by a "metatextual" comment and reformulated; in the second instance, it is interrupted by a summary paraphrase. Reformulation as a type of repetition and interruptions are known to be characteristic of unplanned spoken discourse.

In the context of what is now known about the characteristic features of unplanned spoken as opposed to planned written discourse, as discussed in (2.4.4), the main task of this work is to study features of unplanned spoken as opposed to planned written discourse in the three textual features selected for investigation. Although each textual feature is first investigated in its own right, the

focus of this work is to identify features of spoken/written language. It is hoped that, ultimately, the study of these features will enable us to place the variety of biology lecture discourse along the unplanned spoken versus planned written language continuum.

2.6.4 Theoretical implications of the present approach to the study of features of speech/writing in biology lecture discourse

The investigation of features of unplanned spoken as opposed to planned written language in the discourse of these biology lectures via the analysis of grammatical complexity and the analysis of the textual features of connectives, signalling of sub-topic shift and paraphrase has two main implications. Firstly, the study of grammatical complexity implies that the analysis has to go to minute grammatical details as it has to consider the structure of the clause. Every single sentence in the data is assigned a structural clausal description (see Chapter Three). Secondly, the investigation of each textual feature is ultimately a study of the relation between parts of the text, thus, paraphrase, by its nature, depends on the preceding stretch of text; the signalling of sub-topic shift depends on the following stretch of text and connectives express relations within the text. It is hoped that these features will enable us to reveal some aspects of the textual patterning of lecture discourse.

2.6.5 Variation within lectures

Assuming that the description of biology lecture discourse as a variety is established on the basis of similarities between lectures in grammatical complexity and in the use of the three textual features, it is worthwile to ask whether there still are differences between the three sets of lectures, in regard to the features being studied, and to what extent there are differences within this limited corpus.. It is hoped that it will be possible to identify the exact area of differences between the

three sets of lectures, i.e. whether the differences are due to speakers, topic or the use of one of the textual features. Investigating variation between different sets of lectures would provide further insight into the characteristics of biology lecture discourse as a variety.

Footnotes

(1) - Originally, a distinction between language functions as ideational textual and interpersonal was expounded in Halliday's various works. Sinclair's planes of discourse uses two instead of three Hallidayan functions of language, and he proposes that these two functions are operational at discourse level.

(2) - Montgomery has studied "asides" as a proponent of subsidiary discourse. Lwaitama (1983) has studied the phonological demarcation of asides in lectures in linguistics.("A Linguistic Investigation of so-called Asides in Lecture Discourse, with Implications for E.A.P." MSc Dissertation. Aston University).

(3) - The analysis of subordination versus co-ordination has been applied to academic lectures/textbooks in a recent study (Bilton 1987). However, the size of the corpus used (14,500 words of spoken data and 7,000 words of written data) and the methodology applied to approach the data (parts of the transcript of the lectures were left out in order to - in the author's judgement - maximize comparability between the spoken and the written data) shed doubts on the validity of the results obtained from the analysis. ("Co-ordination and subordination in academic lectures and textbooks" MSc dissertation. Aston University).

(4) - We should note that sometimes Halliday uses "text" also to refer to the unit beyond the sentence. However, he mostly uses "text" in reference to a particular instance of text (see Halliday 1976 : 294).

(5) In the current work, when the word lecture is used in the plural as in "biology lectures", it refers to the variety or genre. But the word "lecture" is sometimes used to refer to a given speech event (we notice that Goffman points to the use of the word "lecture:" to refer to the "spoken transcript" or to the "whole situation". Goffman 1981).

(6) - Widdowson distinguishes between "cohesion" which has to do with the way propositions are linked together; and "coherence" which has to do with the illocutionary force of these propositions (Widdowson 1978 : 54). His approach to "coherence" is unique in that he uses this term to mean overall discourse structure. Sinclair & Coulthard (1975) use the act as the smallest unit of discourse structure but they talk about discourse structure which is conceived of as a rank scale. Halliday & Hassan use the term "cohesion" to cover different types of inter-sentential relations but they refer to discourse structure as macrostructure (Halliday & Hassan 1976 : 322). Winter & Hoey use problem - solution schema (Hoey 1983). So the use of the term "coherence" to mean overall structure is only found in Widdowson.

(7) - The difficulty of defining the term "utterance" on the one hand, and of drawing the distinction between "sentence" and "utterance" on the other is well recognised. Different approaches to "utterance" are suggested. Bar-Hillel (1970) for example, advocates that the term "utterance" (when contrasted with sentence) is used in the sense of a sentence (or sometimes study of sentences) paired with a context (Bar-Hillel in Levinson 1983; 19).

The general trend however, is to use utterance as a pre-theoretical term to label "any stretch of talk, by one person, before and after which there is silence on the part of that person" (Harris 1951 : 14). This is also found in Lyons (1977 : 26)' Sinclair & Coulthard (1975 : 21) ; Levinson (1983 : 19) (who also uses Bar-Hillel's approach to "utterance" quoted above. Levinson : ibid).

In this study, where the whole lecture (as text) is a monologue, i.e. produced by one speaker, such an approach to "utterance" cannot apply because it would just identify the whole text as one utterance and it would link a series of utterances together without providing any basis for identifying their boundaries. Utterance is used here (in particular in talking about sub-topic shift signalling and paraphrase) to refer to a stretch of text and thus avoid using the term sentence, since utterances can be either smaller than or larger than one sentence. But the significant distinction between the two is seen in that utterance functions at the level of discourse while sentence, on the other hand, evokes all the constituent relations in different units in sentence grammar. Using the term utterance will enable us to by-pass problems related to sentence grammar.

CHAPTER THREE

ASPECTS OF GRAMMATICAL COMPLEXITY IN BIOLOGY LECTURE DISCOURSE

3.0 This Chapter studies some aspects of grammatical complexity in biology lectures discourse. The aim is to use these aspects as a measure for identifying biology lecture discourse along the spontaneous spoken versus non-spontaneous written language continuum.

This analysis is to be placed in the context of the background research which has concentrated on comparing samples of spoken and written language using as a parameter grammatical complexity as it is manifested in the use of subordination as opposed to co-ordination and types of dependent clause.

3.1 The sentence in spoken language

3.1.0 The first task which has to be faced in carrying out this analysis is to establish sentence boundaries.

The difficulty of parsing a spoken text into sentence like units is well recognised and widely discussed in the literature (Crystal & Davy 1969; Kroll 1977; Chafe 1979; Stubbs 1983). The analysis of a sample of spontaneous spoken language shows that people do not speak in sentence-like units. Spontaneous speech is marked by features like unfinished structures, repetitions, false starts and so on. (see 2.4.4). The sentence is a theoretical construct primarily set up in grammar to describe the grammatical structure of written language. To apply it to spontaneous spoken language can only meet with difficulty. This difficulty arises with a few instances in the present data. It is illustrated in the following two examples:

ASTON UNIVERSITY LIBRARY

(1) this goes. now I'll have to. so this is the bundle. here is the bundle. I've drawn it in a way {just just just to make it} erm just to exaggerate it

(230938)

(2) but if you are asked to ask the library. to get it for you on loan. I'm sure they do so. and er now some people think that when. journal. articles in learned journals.....

(230710)

Example (1) can be analysed, on grammatical grounds (see 3.1.1) into two unfinished structures: "this goes": "now I'll have to"; and three complete sentence like structures: "so this is the bundle", "here is the bundle" and "I've drawn it in a way just just just to make it erm just to exaggerate it". Part of the last sentence is reformulated (just to make it).

In (2) on the other hand, it is difficult to decide on structural grounds where the fragment 'and er' (underlined in the extract) belongs: does it end the first sentence or begin the second one. In other words, is the first sentence unfinished or does the second sentence begin with a false start? In such a case prosodic cues have to be used to decide on where the sentence boundary should be.

Examples (1) and (2) illustrate situations which are rather rare in the data as a whole. Unfinished structures only represent 3.22% of the total number of sentences in the corpus (2014). What is generally found are complete sentence-like structures, as the following extract shows:

(3) //so it is produced in one place exerting its effects at some distance away// it does not have to be a long way away// there are cells particularly in the central nervous system which may likewise produce factors or secretions which simply pass into a very tiny gap between two cells// so into the fluid surrounding this cell may emerge or emanate this chemical and it may have an effect on a cell which is only a few microns away// so the distance that the chemical travels is in a sense unimportant//

(128431)

We can say that; (1) since the lectures under study have undergone a certain degree of planning (see 2.6.1); (2) since they belong to academic discourse whose purpose is to impart information; (3) and since it is educated speech addressed to an educated audience, it is admissible to expect that the ideas expressed are formulated in recognisable structures.

3.1.1. Identifying sentence boundaries

The general working principle to identify sentence boundaries used in this study is based on Quirk et al's definition of the sentence (1985: 48)

> "The limits of the English sentence are defined in practice, wherever grammatical relations (such as those of subordination and co-ordination cannot be established between clauses."
>
> (Quirk et al 1985 : 48)

This is illustrated in extract (3) quoted above.

As pointed out above, there are a few instances in the data, where the grammatical features of a fragment of text are not indicative of the location of boundaries between sentences. In such cases we have to resort to prosodic cues to identify sentence boundaries .

There is an on-going debate about determining prosodic criteria for sentence like unit boundaries in speech. Studies in this area have shown that there are discrepancies between prosodic units and grammatical units in speech. (see Johns Lewis 1986 ; Farag 1986). However some prosodic features have gained currency as potential markers of sentence boundaries in spoken language. Pitch marking of the initial boundary of the sentence which takes the form of sentence declination (Johns-Lewis 1986 : xxi) has been investigated in a number of studies in different languages which have established that "pitch height is a marker of utterance initiality" (Johns-Lewis 1986: xxi).

This study uses pitch marking as a demarcator of sentence initiality. In extract (2) quoted above, there is a fall of pitch on "do so" which rises again on "now some people think" and goes down with the rest of the sentence. Here the cut between the 2 sentences precedes 'now'.

> //but if you are asked to ask the library to get it for you on loan
> I'm sure they do so and er// now some people think that...../

(230710)

On the other hand, there are a few cases in the data where the difficulty of deciding on sentence boundaries does not stem from the use of unfinished structures in the data, but arises with verbal structures beginning with "and". The question then is whether "and" initiates a sentence or links co-ordinate clauses. In the following example:

(4) //okay, today, we begin our look. at the third and last of the three major groups of land. plant //{I know it's a sad occasion. and I can see that you are looking glum already}. and we're going to be looking at the third group. the. division spermatophyte. or seed plants. and we're going to be looking at the most primitive of them. that are. generally described as the gymnosperms.

(33281)

The question is whether (1) the first instance of underlined "and" introduces a sentence or co-ordinates the clause it introduces with the main clause "we begin our look...."; and (2) the second instance of underlined and introduces a sentence or co-ordinates two successive clauses. In the first instance and is followed by the verbal group "are going to be looking" which is more complex than the verbal group of the main clause "begin our look". It is also repeating part of the object "the third group". The two clauses can therefore be analysed as two sentences. In the second instance, however, and initiates a clause which uses the same verbal group as the preceding clause, namely "are going to be looking". It could be said that repetition at the level of the whole verbal group is typical of speech. Therefore

to decide about sentence boundaries we have to go to the level of prosody. The change in pitch level establishes that and here introduces a new sentence; the change in pitch level also confirms that the first instance of and initiates a new sentence.

3.1.2 Types of sentence in this analysis

The sentence is analysed into:

(i) simple

(ii) compound

(iii) complex (Huddleston 1984: 378)

(i) the simple sentence is one main clause. It may be:

(a) composed of a subject and a predicate:

(5) "so the spores contain n chromosomes"

(332823)

(b) a verbless clause:

(6) "casperian thickening"

(230733)

Unfinished structures which contain a subject and a predicate are analysed as main clauses:

(7) "and there are cytoplasmic connections....."

(231129)

This structure is reformulated in the next sentence.

(ii) the compound sentence is two or more main clauses, as shown in excerpt

(8) below:

(8) //it's produced a root/ it's produced a stem/ and it's produced its first little frond from the surface//

(33256)

(iii) the complex sentence is a main clause and one or more subordinate clauses:

(9) //but should it be necessary/ or should you wish to discuss it further/ or if you have any difficulties er with any aspect of this subject / I'm to be found in room 308//

(12834)

In this excerpt the main clause comes last: "I'm to be found in room 308". It is preceded by the three co-ordinate subordinate clauses:

/but should it be necessary/ or should you wish to discuss it further/ or if you have any difficulties er with any aspects of this subject/

3.2 Depth and complexity

The present analysis of grammatical complexity makes use of two aspects of sentence grammar. Firstly, a distinction is made between subordination and embeddedness. The distinguishing feature is the relationship between the dependent or subordinate clause and the main or superordinate clause. In embeddedness, the dependent clause is a constituent of the clause in which it is embedded; whether that clause is a main clause or whether it is itself a dependent clause. In subordination on the other hand, the dependent clause is not a constituent of the clause it is subordinate to (see Huddleston 1984: 379). Excerpts (10) and (11) below exemplify each case:

(10) //we mentioned earlier that they only modify existing metabolic processes//

(129723)

(11) //bleeding can be stopped/ if you make the osmotic potential of the bathing solution too high//

(230524)

Excerpt (10) exemplifies a case of embeddedness. The dependent clause "that they only modify existing metabolic processes" is embedded as object into the main clause "we mentioned earlier". Excerpt (11) on the other hand instantiates a case of subordination. The dependent clause is a conditional clause. It is introduced by the subordinator "if". Thus this distinction between subordination and embeddedness highlights the nature of the grammatical relation that exists between the dependent clause and the main clause. We must point out that subordination is a cover term for both subordination and embeddedness. A subordinate clause is either subordinate or subordinate embedded. Thus the term "embedded" is used in this work to refer to the subordinate embedded clause.

With respect to the analysis of grammatical complexity of a sentence, the distinction between subordination and embeddedness does not have a bearing on the evaluation of the structural complexity of sentence, i.e. both the subordinate and the embedded clause score in the same way.

Secondly, following Berry who suggests that: "depth can be used as a measure of the complexity of a stretch of language" (Berry 1975: 86), grammatical complexity in this study, is primarily dealt with in terms of depth. Previous studies which have concerned themselves with the investigation of grammatical complexity in samples of spoken and written language (see 2.4.5), have not sufficiently highlighted the role of depth of structure as a marker of complexity in language. In Beaman (1984 : 59) for example, it is merely presented in tabular form.

In this study, depth in structure is used as a basic parameter. Every sentence in the data has a score which indicates its level of complexity representative of the number of recursive embeddedness and/or subordination, as shown in the following excerpts:

(12) //^{0}now this little gland [^{1}laying at the base of the brain] is able to produce chemicals/^{1} which can pass in the blood stream/^{1} and affect other glands or other tissues//

(12845)

(13) //^{0}well I'll just say something now about the structure and function of sieve tube elements simply as a means of /^{1}aiding your understanding of /^{2} what I've already said about the root nodule//

(231224)

(14) //^{0}you take a plant/ ^{0}and you turn it over/ ^{0}and you will see on the underside a whole series of little kidney shape scales//

(332324)

The sentence in (12) is a main clause, and three embedded clauses. The main clause is at (0) level of complexity; the ing-clause which interrupts the main clause (as indicated by the square brackets) is at (1) level of complexity; the next embedded clause "which can pass...." is also at (1) level of complexity since it is embedded into the main clause. The co-ordinated clause is at the same level of complexity as the clause to which it is co-ordinated. Sentence (12) is described to be at (1) level of complexity. The point to be made here is that a co-ordinated clause, whether main or subordinate/embedded, is always at the same level of complexity as the clause with which it is co-ordinated.

The sentence in excerpt (13) is comparatively more complex. Its clausal constituents are a main clause, and two embedded clauses. The first embedded clause is embedded into the main clause, and is at (1) level of complexity; the

second, into the embedded clause, and is at (2) level of complexity. The sentence is at (2) level of complexity.

Sentence (14) is three separate compounded main clauses and it is at (o) level of complexity.

3.3 Co-ordination versus subordination/embeddedness

3.3.0 Quirk et al (1985: 918) define co-ordination as opposed to subordination as a symmetrical versus non-symmetrical relation between units of the same rank. In co-ordination the clauses are of equal syntactic status. In subordination, they are not. They form a hierarchy: one clause is dependent on the other (See examples 12-14 above).

Semantically the two constructions differ in that in co-ordination, the information is asserted whereas in subordination it is (most often) presupposed as given in the subordinate clause. (Quirk et al 1985: 918).

3.3.1 Co-ordination

Co-ordination of clauses may either be signalled by a conjunction of co-ordination or it may not as shown in the following extracts:

(15) //and so for every nitrogen atom/ that's exported or in erm organic form/ two carbon atoms must come into the nodule in the form of co carbohydrate ( ) the carbohydrate be broken down to organic acid/ and combined with the anoniae to form an amno-acid or an amid//

(230430)

The conjunction of co-ordination is underlined. A blank between brackets indicates where it has been omitted. The co-ordinators which are generally used in this corpus are 'and', 'but' and 'or'. They initiate clauses. In sentence initial position, 'and', and 'but' are not analysed as co-ordinators. (see Chapter Four).

A primary analysis of the frequency of occurrence(1) of co-ordinate as opposed to subordinate/embedded clauses in the data shows the following distribution:

Texts	No of Words	F.I For co-ordinate clauses	F.I For sub/emb clauses
Set 1 (3 lectures)	14,200	(134) 9.43	(761) 53.59
Set 2 (3 lectures)	14,200	(102) 7.18	(834) 58.73
Set 3 (3 lectures)	14,600	(138) 9.45	(728) 49.86
Total (9 lectures)	43,000	(274) 8.69	(2323) 54.02

Table (4) : Frequency index for co-ordinate versus subordinate/ embedded clauses .

The results of Table (4) show a clear tendency of biology lectures discourse to use more subordinate/embedded than co-ordinate constructions.

The dominant use of subordination is generally considered a feature of planned written language. (O'Donnell 1974 ; Kroll 1977; Ure & Ellis 1977; Ochs 1979 ; Chafe 1979 ; 1982 ; Brown & Yule 1983). In this light, the interpretation of the results obtained here in this first step analysis is that biology lecture discourse shows a feature of planned written language in its preference for subordination as opposed to co-ordination.

3.3.2 Subordination

As pointed out in section (3.2 above) grammatical complexity is looked at primarily in terms of depth of structure or degree of subordination or embeddedness in the sentence.

As explained earlier (3.2) the levels or degree of complexity of every sentence in the data is identified. It is representative of the number of embedded or subordinate clauses in a given sentence. The analysis yields the results presented in table (5) overleaf. They show that lectures in this data exploit up to (2) degree complexity structure (with some variation among speakers which will be taken up in (7.2.1). The most recurrent structure is of (1) degree complexity. This means that the lectures under study tend to use more simple structures (i.e. one dependent clause per sentence). If we interpret these results in the context of the widely held view that written language is grammatically more complex than spoken language (see 2.4.4), then this analysis of depth of structure shows that biology lecture discourse can be seen as nearer to spoken language.

Thus a count of subordinate clauses as opposed to co-ordinate clauses has decisively shown the tendency of these lectures to use more subordination than co-ordination. But the in-depth analysis of the degree of subordination/embedding hasprovided the evidence that lectures use mostly simple dependent structures, i.e. one subordinate or embedded clause per complex sentence.

Texts	No of Words	O Complexity M or Mc(n..)	1 Complex M+sub or M+Emb	2 Complex $M+sub^1+sub^2$ or $M+E^1+E^2$ $M+sub^1+E^2$	3 Complex $M-sub^1+sub^2+sub^3$ or $M+E+sub^2+E^3$	4 Complex	5 Complex	6 Complex	7 Complex
Set 1 3 lectures	14,200	(208) 14.64	(247) 17.39	(182) 12.31	(93) 6.54	(56) 3.94	(20) 1.40	(6) 0.42	(0) 0
Set 2 3 lectures	14,200	(187) 13.16	(209) 14.71	(234) 16.47	(153) 10.66	(48) 3.88	(15) 1.05	(12) 0.84	(14) 0.98
Set 3 3 lectures	14,600	(354) 24.24	(299) 20.47	(220) 15.06	(81) 5.54	(36) 2.46	(25) 1.71	(0) 0	(7) 0.47
Total 3 lectures	43,000	(749) 17.41	(775) 17.55	(636) 14.79	(327) 6.60	(140) 3.26	(60) 1.39	(18) 0.41	(21) 0.48

Table (5) : Frequency index for sentences distinguished on the basis of their degree of complexity in the corpus

3.3.3. Embeddedness

As stated earlier (3.2.2) a distinction is made between subordinate and embedded clauses. The results of the analysis of the frequency of occurrence of each category of dependent clause are shown in table (6):

Texts	No of words	F.I For subordinate cl	F.I For embeddeded cl
Set 1 (3 lectures)	14,200	(221) 15.56	(540) 38.02
Set 2 (3 lectures)	14,200	(263) 18.52	(571) 40.21
Set 3 (3 lectures)	14,600	(175) 11.98	(553) 37.87

Table (6) : Frequency index for subordinate versus embedded clauses

Embedded clauses are clearly the most frequent category of subordinate clause used in all three sets of lectures (with a frequency of occurrence more than twice that of the subordinate clauses). This dominant use of embedded clauses as compared with subordinate clauses can sometimes be seen in the same sentence, as the following example from the data illustrates:

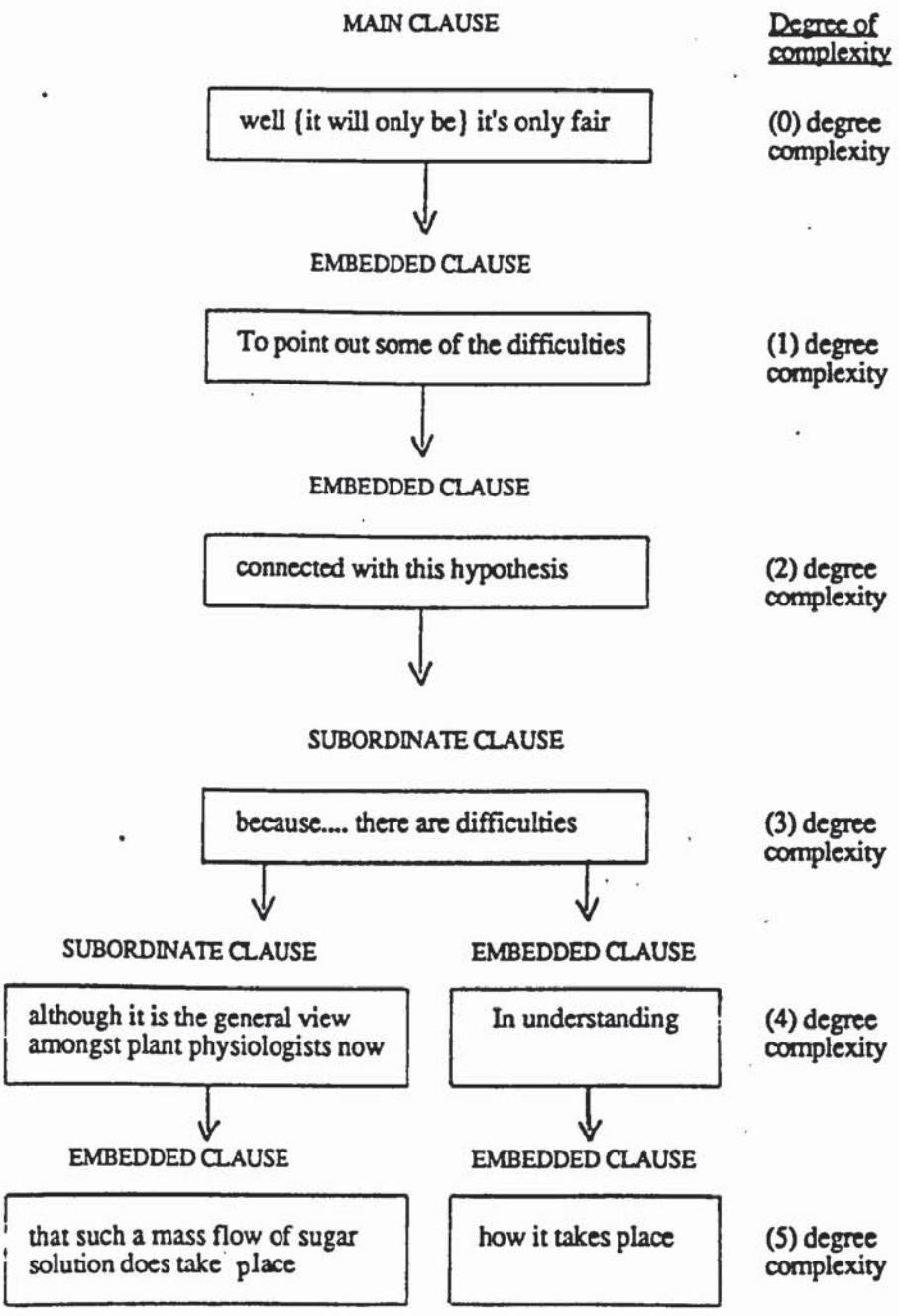
(16) //well {it will only be} it's only fair/ to point out some of the difficulties/ connected with this hypothesis/ because/ although it is the general view amongst plant physiologists now/ that such a mass flow of sugar solution does take place/ there are difficulties in/ understanding/ how it takes place//

(231426)

The structure of the sentence in this excerpt is represented by figure (1) overleaf. It shows five embedded clauses for only two subordinate clauses. It also reveals the tendency of recurrent dependency between clauses to be mostly realized

Figure (1)

SUBORDINATION VERSUS EMBEDDEDNESS IN THE SENTENCE



in a combination of embedded or embedded/ subordinate clauses, rather than a combination of two (or more) subordinate clauses. Indeed the analysis of the nature of recursiveness of dependent clauses applied to the whole data gives the following results tabulated in (7).

No of Dependent clauses per sentence	No of sentences realizing Emb/Emb	No of sentences realizing Sub/Sub	No of sentences realizing Sub/Emb or Emb/Sub
2	238	68	170
3	40	3	0
4	24	0	32
5	6	0	0
6	1	0	4
7	1	0	0
8	0	0	1
9	0	0	1
10	0	0	1

Table (7) : Number of dependent clauses in each type of construction (calculated out of the total number of sentences (2014) in the data).

The most recurrent type of combination is double embedding in the sentence which indicates that embeddedness and not subordination is the mechanism mostly exploited in complex sentences. Importantly, it is the mechanism which inderlies depth of structure in sentences in this data.

3.4 Finite versus non-finite clauses

Structurally clauses in English are described as either finite, i.e. with the verbal group "showing tense, mood, aspect and voice" or non-finite, i.e. with the verbal group "not showing tense or mood, but still capable of indicating aspect and voice" (Quirk et al 1972 : 43).

The use of finite as opposed to non-finite clauses is also one of the parameters applied in studies which have investigated subordination in contrasted samples of spoken and written language. Chafe (1982) for example, has found that present and past participles are features of integrated structures, generally typical of planned written language.

The following example is an instance of a complex sentence from the data showing a finite main clause followed by two embedded non-finite clauses:

(17) //and from the hypothalamus emanate nerves/ connecting it to a little gland at the base of the brain/ called the pituitary//

(128335)

Finite and non-finite verb forms are underlined in the excerpt. The analysis of the frequency of occurrence of finite as opposed to non-finite clauses in this data shows the following results:

Texts	No of words	F I for Finite cl	F I for non-finite cl
Set 1 (3 lectures)	14,200	(1322) 93.09	(245) 17.72
Set 2 (3 lectures)	14,200	(1328) 93.52	(236) 16.61
Set 3 (3 lectures)	14,600	(1585) 105.13	(270) 18.49
Total (9 lectures)	43,000	(4185) 97.32	(751) 17.46

Table (8) : Frequency index for finite versus non-finite clauses

Finite clauses are by far the largest category as opposed to non-finite clauses in all three sets of lectures.

Non-finite structures are generally held to be a manifestation of greater complexity in language. They have been found to contribute to the compactness of written language (Chafe 1982). They are also associated with lexical density (Halliday 1979, Beaman 1984, Johns-Lewis 1987) which is claimed to be characteristic of written language. (Halliday 1979).

The results of our analysis, that finite clauses are more found than non-finite clauses, can be interpreted in this context as further evidence that biology lecture discourse tends to use grammatically simpler structures. In this respect, it seems to be nearer to spontaneous spoken language.

3.5 Types of dependent clause

Dependent clauses are classified into the following categories, (adapted from Quirk et al 1972 ; 1985):

(i) Nominal clauses: which include the following types of embedded clauses: that-clause, WH interrogative clause; nominal relative clause; to-infinitive clause and ing-clause. The first three types of clause are finite clauses; the last two, are non-finite clauses.

Nominal clauses are so called as "they have functions that approximate to those of a noun phrase" (Quirk et al 1985 : 1047)

(ii) Relative clauses: "they generally function as restrictive or non-restrictive modifiers of noun phrases and are therefore functionally parallel to attributive adjectives" (Quirk et al 1985 : 1048). there are two other types of relative clause found in this data: (1) nominal relative clauses which belong to the above category of nominal clauses; and (2) sentential relative clauses whose relative item refers anaphorically to a phrase, clause or sentence (Quirk et al 1985). Sentential relatives function as non-restrictive relatives. The restrictive and the nominal relative clauses are analysed as embedded clauses; the non-restrictive modifiers of noun phrases or sentential relative clauses, as subordinate clauses.

Relative clauses may be finite or non-finite.

(iii) adverbial clauses: they generally function like adverb phrases. They are also often introduced by a subordinator (like when, if, although, so that, because....).

Adverbial clauses are all described as subordinate. They are generally finite clauses, but they may also be non-finite clauses(2)

3.6 Major types of embedded clause

As mentioned in the introduction to this chapter types of dependent clause have also been used in the study of grammatical complexity in samples of spoken and written language. This section looks at the most recurrent types of dependent clause in the data.

As was shown in table (6), embedded clauses are the most recurrent category of subordinate clause. The distribution of each type of embedded clause in the data is presented in table (9) overleaf.

The most frequent type of embedded clause found in these lectures is the restrictive relative clause, as table (9) shows. Its frequency of occurrence is comparable in the three sets of lectures. It is followed by that-clause which also comes second in frequency in all three sets of lectures. Further if we look at the distribution of types of subordinate clauses, as shown in table (10) overleaf, and compare the frequency of occurrence of restrictive relative clause and that-clause with that of subordinate clauses in the data, then we can say that these two types of clauses are the most frequent types of dependent clause in the whole data. The next two sections discuss each one of these two major types of clause and assess their frequency of occurrence in relation to features of spontaneous spoken versus planned written language.

3.6.1 Restrictive relative clause

The restrictive relative clause is semantically defined, as opposed to the non-restrictive relative clause, as encoding information that "forms an integral part of the

Texts	No. of Words	F.I. for RESTRICTIVE RELATIVE	NOMINAL CLAUSES				
			F.I. for that-clause	F.I. for RELATIVE clause	F.I. for ing-clause	F.I. for to-infinitive clause	F.I. for WH INTERROGATIVE clause
Set 1 (3 lectures)	14,200	(267) 18.80	(119) 8.38	(3) 0.21	(71) 5	(58) 4.08	(25) 1.76
Set 2 (3 lectures)	14,200	(262) 18.45	(137) 9.64	(7) 0.49	(65) 4.57	(66) 4.64	(41) 2.88
Set 3 (3 lectures)	14,600	(256) 17.53	(109) 7.46	(11) 0.75	(61) 4.17	(81) 5.54	(46) 3.15
Total (9 lectures)	43,000	(785) 18.25	(365) 8.48	(321) 0.48	(197) 4.58	(205) 4.76	(112) 2.60

Table (9): Frequency index for types of embedded clause in the corpus

Tests	No. of Words	F.I. for CONDIT-IONAL Clause	F.I. for REASON Clause	F.I. for RESUL-TIVE Clause	F.I. for TEMP-ORAL Clause	F.I. for WH INTERR-OGATIVE Clause	F.I. for CONC-ESSION Clause	F.I. for CONTIN-GENCY Clause	F.I. for PURP-OSE Clause	F.I. for CONT-RAST Clause	F.I. for COMPA-RATIVE Clause	F.I. for NON RESTR-ICTIVE Clause	F.I. for PLACE Clause	F.I. for MANNER Clause	F.I. for SUBOR-DINATE Clause
Set 1 (3 lectures)	14,200	(55) 3.87	(19) 1.33	(14) 0.98	(66) 4.64	(8) 0.56	(1) 0.07	(3) 0.21	(6) 0.42	(1) 0.07	(3) 0.21	(38) 2.67	(2) 0.14	(4) 0.28	(221) 15.56
Set 2 (3 lectures)	14,200	(51) 3.59	(44) 3.09	(22) 1.54	(39) 2.74	(3) 0.21	(6) 0.42	(3) 0.21	(8) 0.56	(6) 0.42	(20) 1.40	(50) 3.52	(7) 0.49	(3) 0.21	(263) 18.52
Set 3 (3 lectures)	14,600	(72) 4.93	(39) 2.67	(14) 0.95	(28) 1.91	(0)	(1) 0.06	(1) 0.06	(14) 0.95	(0)	(7) 0.47	(47) 3.21	(0)	(1) 0.06	(175) 11.98
Total (9 lectures)	43,000	(178) 4.13	(102) 2.37	(50) 1.16	(133) 3.09	(11) 0.25	(8) 0.18	(7) 0.16	(28) 0.65	(7) 0.16	(30) 0.69	(135) 3.13	(9) 0.20	(8) 0.18	(659) 15.32

Table (10) : Frequency index for types of subordinate clause in the corpus (calculated per 1,000 words)

message conveyed by the larger construction" (Huddleston 1984 : 399). Whereas the information contained in the non-restrictive relative clause is additional to that contained in the superordinate clause. Prosodically, the non-restrictive relative is marked by intonation separation; while the restrictive relative is bound to its antecedent (Huddleston : ibid). Structurally, the non-restrictive clause is dealt with in this analysis as a subordinate and not embedded clause. It is not a necessary constituent of the clause to which it is subordinate, as the restrictive relative is.

If we compare the following excerpts from the data:

(18) //the sap/ which they are tapping and the stylet do go into the sieve tube element//

(231728)

(19) //These are going to be very important in the process of dehiscence or opening of the sporangium/ which is an extremely active event/ because during dehiscence the water gland absorbs water from the cells of the sporangium//

(332345)

In sentence (18) the relative clause "which they are tapping" specifies which sap is being referred to. It identifies its antecedent "sap". It is restrictive. In (19), on the other hand, the relative clause "which is an extremely active event" gives additional information about its antecedent" the process of dehiscence or opening of the sporangium". Further, the relative clause here can be reformulated as a co-ordinate independent clause "and this is an extremely active event". It is thus non-restrictive.

It is worth mentioning that a third out of the total number of occurrences of non-restrictive relatives in this data (47 out of 135: 3.4%) is realized by a comment clause(2) (see Quirk 1985 : 1112). This is a type of parenthetical clause which in this case, is used as sentential relative as the following example shows:

(20) //Those plants/ as we'll see/ are probably derived from the ferns//

(332715)

The comment clause "as we'll see" functions here as a sentential relative (Quirk et al 1985 : 1116). "As" may be seen to function like the relative pronoun "which":

> those plants are probably derived from the fern, (which we will see).

"which" refers anaphorically to the whole sentence "those plants are probably derived from the ferns".

Furthermore, prosodically, the comment clause has, like the non-restrictive relative, a separate intonation contour which marks it off from the rest of the sentence.

Structurally, the restrictive relative clause in this data, is generally introduced by one of the relative pronouns "which", "where", "who" and "that" except when it is a non-finite restrictive relative, in which case the relative pronoun and the finite verb are omitted. This is illustrated in the following example:

(21) //the receptor part is probably the part of this protein jacket of the chromatin specificity of shape/ determining the union of the steroid with this new receptor...../

(130120)

the non-finite restrictive relative "determining the union of the....." can be reformulated as "which is determining the union of the.....", or more plausibly, "which determines....."

The findings of this study, namely that the restrictive relative is the most frequent type of dependent clause in biology lecture discourse is consistent with

Swales' observation that the relative clause (although with no specification as to restrictive or non-restrictive) is the most common type of clause in scientific English. (Swales 1971 : 54). However, in studies of grammatical complexity in spoken and written language contradictory findings and observations have been reached with regard to the use of the relative clause. Kroll (1977), Chafe (1982) and others postulate that the relative clause (to mean, mostly, the restrictive relative clause) is more used in written language. Golub (1969), Beaman (1984) have provided evidence of the higher frequency of the restrictive clause in spoken language. Graver (1971), on the other hand, states that: "when relative clauses occur in spoken language they are nearly all defining (i.e. restrictive relative) clauses" (Graver 1971 : 124). His argument is that "non-defining clauses" (i.e. non-restrictive relatives) tend to sound formal and unnatural in speech" (ibid). Consequently spoken language tends to favour the use of the "defining" rather than the "non defining" clause.

Because of this disparity in the discussion of the status of the relative clause in relation to spoken and written language, it is not clear whether the use of the restrictive relative clause is to be looked at as typical of speech or writing. Therefore, the use of restrictive relative in this study is interpreted as inconclusive with respect to qualifying biology lecture discourse in relation to spontaneous spoken as opposed to planned written language.

3.6.2. That clause

That-clause is the second major type of dependent clause in this data. It is mostly introduced by "that" which is omitted in a few cases. the following excerpts from the data are instances of this type of clause:

(22) //so as the calcium level falls/ you can see/ that parathyroid hormone levels in the blood go up//

(12957)

(23) //now of course this means/ ø the osmotic potential of the sap in the sieve tube element is high/ and ø water would be {attracted} attracted to the sieve tube element er {in} by osmosis//

In excerpt (23) "that" is omitted (indicated by ø). The verbs used in the main clause in each excerpt ((22) and (23)) are factual verbs (Quirk et al 1985 : 1180). This is generally the case in this data. However, that-clauses also include reported clauses, which following Huddleston (1984 : 150) are also analysed as embedded clauses,(3) i.e. they are necessary constituents of the superordinate clause. It must be mentioned that reported clauses are rather rare in this data. They only constitute 4.93% of the total number of occurrences of that-clause in the whole data. An example of a reported clause from the data follows:

(24) //He was a gentleman who first pointed out/ that in your body and mine the cells in fact lived in two environments//

(12876)

"that in your body and mine.... environments:" is a nominal that-clause, realizing a reported clause. It functions as direct object of the verb "pointed out" in the superordinate clause.

Nominal clauses, also called complement clauses have been found to be more used in written language (Chafe 1982). Their dominant use in writing is explained as being a manifestation of the "integrated" nature of writing, as they contribute to the structural compactness, characteristic of written language. (Chafe : ibid). It is also generally argued that the use of complex nominal constructions (such as nominal clauses) tend to be preferred in written language. (see Akinnaso 1982 for a review).

The use of nominal clauses in this data, of which that-clause is the most prominent type, can therefore be seen as a manifestation of a feature of written language in biology lecture discourse.

3.7 Summary and conclusion

This analysis of subordination versus co-ordination in the data, has shown that:

1 biology lectures make more use of subordination than co-ordination.

2 the study of depth of structure or degree of embeddedness reveals that biology lectures favour one degree complexity structure, that is, they tend to use more simple sentences.

3 biology lectures use more finite than non-finite clauses.

On the first parameter, biology lecture discourse is grammatically complex and shows a characteristic of written language. On the second and third parameters, it is less complex and seems to be nearer to spoken language. Further, a study of the major types of embedded clauses used reveals that:

- the most frequent types of clause (the restrictive relative clause) does not have a clear status with respect to the spoken versus written language dichotomy.
- but the second major type of clause (that-clause) is claimed to be a feature of written language.

Finally, the use of complete sentence-like structures in the lectures places them nearer to written language.

The aim of this analysis was to use grammatical complexity measured in the use of subordination as opposed to co-ordination, as a parameter for defining biology lecture discourse along the spontaneous spoken versus planned written

language continuum. The results of this analysis are contradictory: they suggest that these lectures seem to combine features of both spoken and written language. But they are not conclusive as to where biology lectures belong on the continuum. That is they are not conclusive as to whether biology lectures are to be placed nearer to the spontaneous spoken and or to the planned written end of the continuum.

The next three chapters will investigate each a textual feature of these lectures, namely "sentence-initial connectives", "sub-topic shift signalling" and "paraphrase". Each feature will be dealt with in terms of its grammatical and textual structuring. Characteristics of spoken versus written language will be identified in an attempt to understand better where biology lectures belong on the spontaneous spoken versus planned written language continuum.

Footnotes

(1) A frequency index per 1,000 words is used (see Beaman 1984). Figures between brackets in a table indicate absolute occurrences of a given structure in the data. This reading applies to all the tables used in this work.

(2) See appendix (2) for an exemplification and a brief discussion of the types of dependent clauses which have not been discussed in the main text.

(3) This is also the stand of Quirk et al (1985) who describe reported clauses as realized as nominal that-clauses (Quirk et al : 1985 : 1025); and Kuno (1975) who analyses reported clauses as object complement to the reporting clause.

Reported clauses have a different treatment in systemics where they are not considered to be embedded (see Halliday 1968 : 195). It must be mentioned, however, that when no "senser" or "sayer" is involved, i.e. the main clause is of the form "it is said", "it is believed", the reported clause is considered a fact clause, thus embedded as object in the main clause (see Halliday 1985 : 243).

CHAPTER FOUR

SENTENCE-INITIAL CONNECTIVES

4.0 In the previous chapter we saw that the investigation of features of spontaneous spoken versus planned written language had to be taken beyond the level of the sentence to the level of discourse. This chapter presents a study of a textual feature of these biology lectures namely the use of sentence-initial connectives. It examines some aspects of their use in sentence initial position in relation to the spontaneous spoken versus planned written discourse dichotomy. In this perspective the use of sentence-initial connectives is studied in terms of: (1) the relevance of the proposition expressed in the sentence where a connective occurs, in relation to the preceding stretch of text; (2) the relevance of the use of the sentence where a connective occurs, interpreted in the light of the audience background knowledge and assumptions about the speaker's knowledge and authority; (3) their role in signalling the structure of the text.

4.1 Preliminaries

4.1.1 Terminology

Before presenting the phenomenon it is necessary to explain why the term "connective" is used to refer to the elements under investigation. The first point to be made is that these elements are an old concern both in traditional grammar and in modern linguistic theory including text linguistics. They have been referred to by different labels according to: (1) the level of analysis where they are investigated (ie. sentence or text level, (2) the theoretical framework used to handle them. By way of example, 'and' is a coordinating conjunction in traditional grammar where it is looked at as interclausal or interphrasal; as a sentence initiator it is a conjunction (Halliday & Hassan 1976: 233) or a connective (Van Dijk 1977:210) which may

either be semantic or pragmatic. This study is concerned with the use of these elements at discourse level i.e. beyond sentence level. The term 'connective' is most appropriate to refer to them globally as its use can be interpreted at both the semantic level i.e. when it is used to connect propositions; and at the pragmatic level i.e. when its use can be interpreted in terms of the listeners assumptions about the authority of the speaker.

4.1.2 Sentence-initial connectives in the corpus

An examination of the data shows that there is a recurrent use of the connectives, 'and', 'so', 'but', 'well' and 'now' in sentence initial position(1) in the data. Consider the following excerpt:

(1) //When animals or indeed plants became multicellular and certain different parts of the animal or plant body became specialised some sort of communication system was necessary between the different parts in order to coordinate their function. //So we are saying with the development of a multicellular organism some form of internal communication became necessary.// and you are probably aware that this is accomplished by the joint activity or via the joint activities of nerves on one hand ...//

(12838)

Underlined 'so' and 'and' are instances of the occurrence of sentence-initial connectives. 'So' introduces a sentence which is a reformulation of the preceding one. The relation between the two sentences is then a paraphrase relation which is partly signalled by the use of 'so' in this position (see chapter six). And, on the other hand, introduces a sentence which stands in a different type of relationship with the preceding stretch of text: its content is additive to what has gone before. If, on the other hand, we consider the uses of 'and' in:

(2) //If we look at the next slide we see that many of these trees reach enormous age and size //And this by the way is your humble narrator in the Red Wood Forest of California.// And you really can't see the extent of this tree.....

(333312)

we can see that the first occurrence of 'and' introduces a digression or in Coulthard and Montgomery's terms, an aside, (Coulthard & Montgomery 1981 : 36). Whereas the second occurrence of 'and' initiates a sentence which resumes the main text: its content is additive not to that of the aside but to the sentence which precedes the aside.

So the three occurrences of 'and' illustrated so far realise different functions. The occurrence of 'now' and 'well' in the following examples illustrate yet a different phenomenon:

(3) //Now what is it that gives these gves this sap a high osmotic potential.// Well it is the nitrogen. The organic nitrogen has been fixed in the bacterial cells and has moved in a concentration gradient away from the infected cells towards the vascular strands //Now in considering this movement I want you to think about the symplast//

(231117)

we observe that the first occurrence of 'now' introduces a sentence which announces a 'sub-topic shift' (see Chapter Five). The second occurrence of 'now' signals a shift in topic focus. Where topic is here looked at at sentence level. The topic in the preceding sentence, is about the organic nitrogen; in the sentence initiated by 'now, it has shifted to the movement of the organic nitrogen. "Well', on the other hand, initiates the sentence which immediately follows the 'sub-topic shift signalling' utterance. In other words, 'well' marks off the boundary between the 'sub-topic shift signalling' utterance and the stretch of text which discusses the sub-topic which has been announced in the 'sub-topic shift signalling' utterance. Further, we notice that the sub-topic shift signalling utterance is realized in a question; 'well' initiates the answer to that question.

The excerpts quoted above (ie. (1) to (3)) illustrate different uses of the

connectives as they are observed in this data: 'so' and 'now' in excerpts (1) and (3) introduce a textual feature; 'well' in (2) is used as a text organising constituent; 'and' and 'now' in (1) and (3) are markers of a semantic relation.

These observations show that there are different aspects of the use of sentence-initial connectives in these biology lectures. They will be discussed in detail in the main sections of this chapter, in relation to the spoken versus written language dichotomy. But first we need to study some distributional features of these sentence-initial connectives. Namely the kinds of sentence-initial connectives found in the data; the frequency of occurrence of the connectives which make up the major type.

4.2 Kinds of sentence-initial connectives : frequency of occurrence in the corpus

4.2.0 A sentence-initial connective can be made up of one word (e.g. "and", "so", "well") or more than one word ("and now", "but nevertheless", "but in contrast"). Table (11) overleaf shows the frequency of occurrence of each kind of connective per 1,000 words, in the corpus.

Texts	No of words	F.I For all sentence-initial connectives	F.I. For one-word sentence-initial connectives	F.I. For sentence-initial connectives of more than one word
Set 1 (3 lectures)	14,200	19.85 (282)	19.08 (271)	0.77 (11)
Set 2 (3 lectures)	14,200	20.42 (290)	19.50 (277)	0.91 (13)
Set 3 (3 lectures)	14,600	16.64 (243)	16.30 (238)	0.34 (5)
Total (3 lectures)	43,000	18.95 (815)	18.27 (786)	0.67 (29)

Table (11) : Frequency index for kinds of sentence-initial connectives

We notice that one word connectives are by far the most frequent sentence-initial connectives in the corpus. Furthermore, none of the sentence-initial connectives made up of more than one word has a frequency of occurrence which exceeds four occurrences in the whole data. Accordingly this analysis will be essentially concerned with the use of one word connectives. Reference will be made to the use of sentence-initial connectives composed of more than one word, where it is seen to be relevant to the issues raised by the use of one word connectives.

4.2.2 Frequency of occurrence of one word sentence-initial connectives in the corpus

Five out of the eleven one word sentence-initial connectives found in the data are frequently used, while the remaining six have a negligible percentage of occurrences (calculated per 100 occurrences out of the total number of one word

connectives in the corpus) (see table 12 overleaf). The present analysis deals with the five most prominent one word connectives, namely "and", "but", "so", "now", and "well". If we look at the frequency of occurrence of these five one word connectives we notice that:

- "and" is the most frequent one word connective in the whole data.
- "so" is more used in set (1) than in set (2) and set (3)
- "now" and "well" are more frequent in set (2) than in set (1) and set (3).

The extensive use of 'so', 'now' and 'well' by one speaker could be seen as a manifestation of inter-speaker variability, an issue which will be taken up later in Chapter Seven.

4.3 Relevance and sentence-initial connectives

This section discusses the application of the concept of relevance to the use of sentence-initial connectives in this data, and identifies aspects of spontaneous spoken as opposed to planned written discourse as they appear in relation to the use of connectives.

4.3.1 The concept of relevance

The concept of relevance in this work is taken from Grice's philosophical essay on "Logic and conversation": (1975). Grice postulates four maxims underlying the "co-operative principle" of conversation; quality, quantity, manner and relation(2). "Be relevant" is the reformulation of the last maxim of "relation". In other words, in conversation if participants adhere to the co-operative principle every speaker's contribution should bear a relationship to what has gone before. This relationship is either explicit in what is actually said or implicit. Participants

Texts	and	but	so	now	well	right	certainly	then	hence	however	alternat-ively	Total No. of connectives
Set 1 (3 lectures)	(102) 37.65	(26) 9.59	(98) 36.16	(30) 11.07	(6) 2.21	(1) 0.36	(0) 0	(3) 1.10	(2) 0.73	(1) 0.36	(2) 0.73	271
Set 2 (3 lectures)	(124) 44.76	(12) 4.33	(42) 15.16	(70) 25.27	(27) 9.74	(0) 0	(1) 0.36	(0) 0	(0) 0	(0) 0	(1) 0.36	277
Set 3 (3 lectures)	(139) 58.40	(29) 12.18	(24) 10.08	(38) 15.96	(81) 3.36	(0) 0	(0) 0	(0) 0	(0) 0	(0) 0	(0) 0	238
Total (9 lectures)	(365) 46.43	(67) 8.52	(164) 20.56	(138) 17.55	(41) 5.21	(1) 0.12	(1) 0.12	(3) 0.38	(2) 0.25	(1) 0.12	(3) 0.38	786

Table (12): Frequency of occurrence of one-word sentence-initial connectives in the corpus (calculated per 100 occurrences of the total number of one-word connectives).

have to draw then on conversational implicatures i.e. what is implied by what is said to see this relationship.

It has been pointed out that Grice's discussion of 'relevance' offers very little by means of clarification (Werth 1981: 130; Brown and Yule 1983). Grice briefly defines the concept of relevance and observes that it involves a certain number of difficult issues which have to be dealt with such as: "what different kinds and focuses of relevance there may be..." (1975: 46).

In this work, the application of this concept in relation to the use of sentence-initial connectives in biology lectures is carried out in terms of two dimensions: the concept of relevance is applied to mean: (1) semantic relevance: (2) pragmatic relevance. 'Semantic relevance' is essentially concerned with the meaning of the propositional content of the sentence where the connective occurs. Pragmatic relevance on the other hand is interpreted in the light of the listeners' background knowledge and assumptions about the speaker's authority and academic knowledge.

4.3.2 The concept of relevance and biology lecture discourse

There are three aspects of biology lecture discourse which favour the application of the notion of relevance in our analysis of the use of connectives:

1. It is pedagogic discourse: its purpose is to successfully impart or transmit information. If this is to be achieved it is expected that the verbal representation of concepts and thoughts would show explicitness and clarity. The relation of relevance between the utterances which make up the produced text should be in most cases explicitly recoverable to the audience. In other words the articulation of ideas would have a linguistic manifestation.

2. It is expository scientific discourse: that is to say it is mostly descriptive and explanatory. These characteristics have certain grammatical manifestations such as the use of the simple present (see Lackstrom et al 1970-1973); the use of explicit reference to the phenomenon being described i.e. repeated use of technical terms in the text. Such features are part of the linguistic cues which may be used by the audience to process the text. They contribute to make the relevance of connected sentences recoverable to the listeners.
3. As seen in 2.6.1 these biology lectures are semi-planned. It can be pointed out that a text which has undergone a certain degree of planning and forethought shows more articulation and overtly marked coherence than for example, a spontaneous conversation, for two main reasons:

(i) - in semi-planned discourse the topic(s) are well defined whereas in spontaneous conversation they have to be negotiated.

(ii) - in a monologue discourse one speaker holds the floor (see Goffman 1981) while in spontaneous conversation the participants have all an equal claim to hold the floor subject to the rules of turn taking (Sacks, Schegloff & Jefferson 1973).

A conversation then might show a tentative stage where participants try to establish a common topic for discussion or change the topic of discussion. Bearing in mind the observations made above, it is reasonable to assume that the relation of relevance is more 'explicit' within one speaker's discourse than across participants' contributions.

4.4 Semantic relevance

As pointed out earlier in 4.3.1 the application of the concept of relevance in this study is seen in terms of "semantic relevance" and "pragmatic relevance". In this section features of semantic relevance in the use of sentence initial connectives are examined. They are features of:

- semantic progression within one sub-topic
- reformulation within one sub-topic
- shifting of sub-topic focus

4.4.1 Semantic Progression within one sub-topic

Consider the following example from the data, which illustrates a use of sentence-initial and:

(4) //For example insulin from the pancreas is normally only released when glucose levels in the blood go up after a meal //This provokes the release of insulin from the islets of Langerhans in the pancreas //And this insulin acts upon muscles and upon fat cells to stimulate the uptake of glucose //And that has the effect of bringing the glucose level in the blood back down to normal//

(129242)

In this excerpt both sentences initiated by 'and' express an addition of information to what has gone before. Each sentence builds upon the content of the preceding fragment of text. The first sentence initiated by 'and' expands on the concept expressed by the noun phrase (NP) 'this insulin' which has been defined in the previous sentence as 'the insulin which has been released from the islets of Langerhans in the pancreas'. Similarly, the second sentence initiated by 'and' expresses a semantic progression (c.f. Hutchins 1977: 21): the anaphoric pronoun 'that' which stands for the whole of the previous sentence ('and this insulin acts upon muscles and upon fat cells to stimulate the uptake of glucose') is expanded on.

These sentences can be analysed in terms of "given" "and" "new" information (Halliday 1967; Lyons 1977) as each sentence builds upon a 'given' item ('this insulin' in the first sentence, 'that' in the second sentence) which can be recovered from the previous fragment of text. Semantic relevance is seen here in the thematic progression realized in these sentences.

Halliday and Hassan (1976: 233) have dealt with this function of 'and' in sentence initial position. In their analysis, 'and' is referred to as 'conjunction' i.e. the realisation of a type of 'cohesive tie', and it signals an additive relation between the sentences conjoined by 'and'. Their analysis is however limited as conjunctive 'and' does not relate more than two successive sentences. In the example used in their discussion:

> " I wonder if all the things move along with us" thought poor Alice. And the Queen seemed to guess her thoughts for she cried faster, don't try to talk!
>
> (ibid)

the sentence containing 'and' is analysed as being only linked to the preceding one. As extract (4) shows 'and' in this data may relate more than two sentences: the first occurrence of 'and' relates the sentence where 'and' is initial to the two preceding sentences.

A more adequate treatment of additive 'and' is found in Van Dijk (1977: 210) where 'and' is referred to as a connective. Sentence initial 'and' is described as having a number of functions one of them is to "conjoin utterances by indicating an addition (my underlining) or a continuation of a statement" (1977: 211).

A different type of semantic progression is realised in the following excerpt from the data:

(5) //But nowadays we know that the heterocyst is the cell in which nitrogen fixation takes place. //And erm also it's a cell in which

photosynthesis does not take place.// So in other words the two functions of nitrogen fixation and photosynthesis are separated in this organism as they need to be because otherwise the nitrogenese will be oxidatively destroyed by the oxygen produced in the photosynthetic cells during photosynthesis//

(230224)

The first clause in the sentence initiated by so (underlined in the excerpt) is a result clause (Quirk et al. 1985: 1108). It is also an inference paraphrase as its information content can be inferred from the preceding stretch of text (see 6.4), which is that the blue green algae are composed of two types of cells: the vegetative photosynthetic cells and the heterocyst. The two types of cell have different properties which keep the functions of photosynthesis and nitrogen fixation separated. The rest of the sentence however expresses new information: it explains why these two functions need to be separated. This type of semantic progression which is also seen as a manifestation of semantic relevance is different from the one illustrated in (4) as it is realised by a sentence whose information structure may be described in terms of a paraphrase followed by an elaboration on that paraphrase.

4.4.2 Reformulation within one sub-topic

Reformulation within one sub-topic is observed in one type of paraphrase (see 6.2) namely the summary paraphrase. It is generally initiated by the connective 'so' or in a few instances, by the compound connective "and so":

(6) //so some extracts or emanations from the pituitary gland are clearly responsible for overall body growth and in particular the growth of the long bones in the legs//

(12889)

In this case there is no new information and therefore no semantic progression. The relevance of the paraphrase is that it states in a concise form the main point of the previous discussion leaving out details such as the description of the visual which was used in the discussion.

4.4.3 Shifting of sub-topic focus

Shifting of sub-topic focus is a feature of some of the sentences initiated by the connective 'now'. Such sentences do not signal a shift in sub-topic (see Chapter Five). Rather, they are found within one sub-topic and they are about a new point in the discussion of a given sub-topic. In the following extract:

(7) //Now during the process of fertilisation pollen grains are released from the microsporangium in enormous quantities//

(333035)

The sentence introduced by 'now' expresses one aspect of the "life cycle of the cycad", the sub-topic under discussion. This aspect has not been dealt with so far; the semantic relevance of this sentence is that is expresses a new point. We may remark that the relationship between the sentence initiated by 'now' and the preceding fragment of text is not that of semantic progression as we described it earlier in relation to excerpts (4) and (5). We saw that the semantic content of the sentences connected by 'and', or 'so' add to what has preceded by building upon 'known' information. The information structure of such sentences was analysed in terms of 'given' and 'new' where the known information is expressed by an anaphoric pronoun or the lexical repetition of a technical term. It was also analysed in terms of a paraphrase, followed by an expansion on that paraphrase. In excerpt (7) however, the sentence introduced by 'now' expresses a new aspect in the description of "the life cycle of the cycad". It does not use any term from the immediately preceding stretch of text or pro-form as old information to serve as a point of departure for further expansion. A different kind of shifting of topic focus is realised in:

(8) //Now when Bartol removed the testes from this cockerel what he noticed was that the comb shrank dramatically lost its turgid rigidity if you will and it lost its bright red coloration as did the wattle//

(128638)

where 'now' initiates a sentence which expresses the main point of the story. The preceding stretch of text sets the scene, namely that Bartol used cockerels in early experiments on hormones and that cockerels had specific secondary sexual characteristics which were"a bright red and turgid comb and wattle" and that the experiment Bartol did was to castrate the cockerel. The sentence introduced by 'now' expresses the focus of the story as it is about the effect of the experiment on the cockerel's secondary sexual characteristics. The semantic relevance of this sentence here, is that it brings into focus the main point of the story.

This section has dealt with the semantic relevance of the sentences introduced by connectives in relation to the rest of the text, as manifested in three aspects: semantic progression within a 'sub-topic', reformulation realised as a summary paraphrase and shift of focus within a sub-topic. The next section will discuss these features in relation to the spontaneous spoken versus planned written discourse dichotomy.

4.4.4 Features of semantic relevance in the use of sentence-initial connectives and the spontaneous spoken versus planned written discourse distinction

Two manifestations of semantic relevance, namely semantic progression and shift of focus within a sub-topic are evidence of the progression of ideas in these biology lectures and of the way in which the information to be transmitted is articulated.

As discussed above (4.3.2) this is a feature to be expected in pedagogic discourse (see Crystal & Davy 1969 : 227). However in relation to spontaneous

spoken versus planned written discourse semantic progression can be expected in either mode. It is a natural feature of discourse (whether spoken or written). There might, however, be differences in the surface features which realize it in one mode or the other.

In this data, the most salient feature of sentences which are additive in content to the preceding stretch of text is that they are mostly initiated by "and". As it can be seen from the results of the analysis of the frequency of occurrence of one-word connectives presented in table (12), "and" is the most frequent sentence-initial connective in the lectures (46.43% out of the total frequency of all one-word connectives in sentence initial position). Although "and" can have other functions than being additive, it generally retains this function (see 4.6.1). Additive "and" constitutes 83.56% of all occurrences of "and" in these lectures.

The examination of a random sample from two biology textbooks(3) shows that in 500 sentences looked at in this sample no occurrence of initial "and" is found. While in these lectures 97 out of 500 sentences are initiated by "and". On the other hand, the investigation of additive "and" in contrastive studies of spoken and written narrative discourse has shown that "and" is typical of spontaneous spoken discourse (Kroll 1977 ; Beaman 1984 ; Farag 1986). Therefore, it can safely be assumed that these biology lectures show a feature of spontaneous spoken language in their extensive use of additive sentence initial "and", and of initial "and" in general.

The last feature of semantic relevance as manifested in the use of reformulations introduced by the connective "so" is also primarily a feature of pedagogic discourse. The use of reformulations has essentially the pedagogic function to reinforce the transfer of information through repetition.

Repetition introduces redundancy in the text. As it will be seen in Chapter Six, reformulation, as paraphrase, is heavily exploited in these lectures. It introduces an amount of redundancy which would not be expected in a written academic text since heavy redundancy in language production is a feature of speech (see Darian 1983). Thus the pedagogical purpose of the lectures motivates the recurrent use of reformulations which manifests a characteristic of spontaneous spoken language.

It has been found that utterance initial "so" is a feature of the oral narrating of stories (Farag 1986 : 110). In the present corpus most instances of paraphrase, specifically the summary paraphrase, (see 6.2) are introduced by "so". Since it has been observed that paraphrase itself is more used in oral delivery and shows more spontaneity, it is natural to conclude that the use of "so" in paraphrase would not be much expected in academic writing.

4.5 Pragmatic relevance

Features of pragmatic relevance are found in this study in three aspects of the use of sentence-initial connectives. These features are markers of:

- speaker/audience interaction
- speaker's monitoring of his speech

4.5.1 Features of speaker/audience interaction

There are two types of features which are a manifestation of the interaction between the speaker and the audience. They are features of:

(i) Audience involvement:

This is reflected in the use of sentences such as:

(9) //And you may be familiar with some of these //Up here up here at the base of the brain is the pituitary//

(128442)

which are used to address the audience and which are almost invariably initiated by a connective. In this excerpt the speaker is explicitly inviting audience participation: the pronoun 'you' is used to address the audience. It can be seen as an overt marker of the interactive nature of this sentence.

(ii) Speaker/audience status relationship:

This is manifested in the use of 'now' and 'well' as instantiated in the following excerpts:

(10) //But on the underside it also has a vascular tissue as every leaf would have //Now it's possible to derive ovaries from such from such a structure by the fusion and folding of different numbers of these basic units or carpels//

(33361)

(11) //So can we therefore in the face of their diversity of structure say anything about their general properties// Well I think we can//

(129225)

In (10) the speaker was describing the structure of a leaf. "Now" is used to introduce a movement or shift in focus to how the derivation of ovaries can be obtained through the structure of a leaf. On the other hand "well" in (11) initiates the answer to a rhetorical question. In addition, relevance in the use of these connectives is seen here as a manifestation of the speaker/audience status relationship. The speaker has the authority and knowledge to monitor and transmit the information to an audience of learners i.e. to an audience of different academic status. Features of his control over the interaction with the audience can be seen in the structuring of his argument and in the way he monitors his speech pedagogically. In this instance, it is seen in the use of "now" to signal a shift of sub-topic focus, and in the use of a rhetorical question followed by an answer and the use of "well" to initiate that answer.

4.5.2 Features of the speaker's monitoring of his speech

They are mostly features of two uses of "well":

- as answer initiator following a rhetorical question
- to initiate the reformulation of part of the sentence where it occurs.

These two uses of "well" are discussed next.

4.5.2.1 "Well" as answer initiator following a rhetorical question:

This is realized in the following excerpt:

(12) //Where did they come from// how did they evolve// Well to find out we have to return one more time to the carboniferous period about 270 million years ago//

(33289)

The first two sentences make up a "sub-topic shift signalling" utterance and realize two rhetorical questions. The utterance which is prefaced by 'well' provides part of the answer to the rhetorical question and initiates the discussion of the new sub-topic. The use of 'well' as answer initiator has been investigated (Lakoff 1973; Svartvik 1980 ; Owen 1981). Lakoff in particular has looked at "well" in answers. She has worked out a taxonomy of pairs of question/answer based on her finding that the use of 'well' prefixing an answer distinguishes between different types of question/answer. Although she makes interesting observations, they do not apply to the use of 'well' in answers in this data for one main reason: she only looks at pairs involving real questions (Levinson 1979 : 391). As already mentioned above, the majority of questions followed by 'well' - initiated answers are rhetorical questions i.e. they are not meant to elicit a response from the audience. The use of 'well' in such a context, as can be observed in excerpt (12), is a manifestation of the way in which the speaker monitors his

speech pedagogically. It is essentially a feature of oral delivery. It is a feature of the structuring of pedagogic discourse.

The rhetorical patterning of question/answer realize the next sub-topic to be dealt with (see Chapter five). As has been observed in the previous section the relevance of the use of connective "well" in this context is seen as a manifestation of a feature of pedagogic discourse.

4.5.2.2 "Well" as editing marking of a reformulation

Consider the use of "well" in the following excerpt:

(13) //Now I asked you {well I did give you some of the evidence I didn't merely ask you} to believe that erm the movement of a sugar solution in the sieve tubes...../

(231421)

"Well" here, does not initiate the sentence but it introduces the reformulation of the beginning of the sentence. In this case, "well" is used as a checking of the flow of speech. It is drawing the attention of the audience to the fact that that something is going to be reformulated. Although this analysis is essentially concerned with sentence initial connectives, this use of "well" can be exceptionally included since it is occurring in initial position in an interrupting structure, which can be a phrase or a clause. Thus in excerpt (13) quoted above, the interrupting structure consists of two clauses "I did give you some of the evidence/ I didn't merely ask you/.....". The speaker, here, reformulates the beginning of his utterance to provide more grounds to reinforce the claim that students must be aware of one possible description of the movement of sugar solution in the sieve tubes as it has been dealt with in the previous lecture. "Well" is being used here as an editing marker (see Svartvik 1980) of the reformulation of the sentence. The audience find the reformulation to be relevant to and more specific than what has gone before it. This stems from their knowledge of the speaker as authority in the

field and his right to refine his expression and continue holding the floor unchallenged.

The use of "well" as discussed in (13) above can also be seen in:

(14) //Now if if this amount of sap was moving in a single sieve tube element that means in a tube composed of single cells placed end to end it would mean that the amount of sugar solution {well the amount of sap} which passes through each cell...../

(231828)

Here the interrupting structure is a phrase "the amount of sap". The speaker reformulates part of the sentence to use a more appropriate term namely the technical term "sap" as a substitute for the compound noun phrase "sugar solution". the reformulation is prefixed by "well". In this instance "well" can be seen as an editing marker of self-correction (Svartvik 1980 : 175). Further, in this example and in example (14) above, "well" is used as a "frame" to use Sinclair & Coulthard's term (1975:22) i.e. it is used to introduce a part of the text.

4.5.3 Pragmatic relevance in the use of sentence-initial connectives and the spontaneous spoken versus planned written language distinction

There are three aspects of the use of "and" and "well" as discussed above in 4.5.1 and 4.5.2 which can be interpreted in relation to the spontaneous spoken versus planned written discourse dichotomy.

(i) "and"

As discussed above in relation to excerpt (9), "and" is found in a sentence which is used to address the audience. It assumes authentic interaction which is marked by the use of the pronoun "you" to address the audience. It is also produced in relation to a visual (namely a slide), i.e. an element of the immediate environment which both the speaker and the audience can see. The use of this

sentence, introduced by "and", assumes real interaction i.e. taking place in real time (Lakoff's visibility criterion 1979. (See 2.4.3). The use of "and" here is a feature of oral delivery. In addition, as observed in 4.4.4, sentence initial "and" can be seen as a feature of spontaneous spoken language in the lectures as compared with a written biology textbook.

(ii) "well" as answer initiator

In this data, "well" in sentence initial position introduces the answer to a rhetorical question. As remarked above (4 5.2.1) this use of "well" is a feature of academic discourse, of the way in which the speaker/lecturer monitors his speech pedagogically. On the other hand, the use of "well" in spontaneous spoken English has been investigated (Svartvik 1980). Svartvik has found that when "well" is not used as a "degree word (e.g. you know that perfectly well) or as a manner adverb (e.g. he speaks well) "(Svartvik 1980 : 168), it is virtually restricted to spoken language (ibid : 169), whether it is found in initial position in the utterance (see also Stubbs 1983 : 69) or embedded in discourse. Therefore the use of "well" in this context in this data, can be seen as a feature of speech used in the lectures for pedagogical purposes.

(iii) "well" as editing marker of a reformulation

"Well" in this use introduces an interrupting structure in the sentence. It signals that the speaker is going to revise part of his utterance. "Well" in this use can be seen as a characteristic feature of a text in the process of being produced in real time (see Quirk 1985 : 1312). Therefore it is characteristic of a spontaneous spoken text. Moreover it has been found that the common denominator of all uses of "well" in speech is that "it signals that the speaker is going to shift grounds.... "well" signals a modification or partial change in the discourse" (Svartvik 1980 : 177). Consequently the use of "well" here is a typical feature of spontaneous spoken language in the lectures.

4.6. Connectives in sentences used in relation to the visual element

It is observed in this data, that the connective "and" often introduces a sentence which is used in relation to a visual element i.e. a slide or a diagram. In the following examples

(15) //And you can see this area in the neck just beneath the Adam's apple...../

(128832)

(16) //And this is a specimen from the Birmingham Botanical Gardens..../

(332547)

in both sentences, reference to the slide has a linguistic manifestation, "this" which in (15) refers to a part of the visual ; in (16) it refers to the whole slide as it is being made visible to the audience. "This" in both cases is a place deictic i.e. it is only interpretable with reference to elements of the immediate environment i.e. the slide. Its use is thus restricted to oral delivery. A written academic text would use a figure between brackets to refer to a non-verbal element (see 5.5.2 for further discussion of this point). On the other hand we can observe that the use of the pronoun "you" in (15) (also deictic) to address the audience is, as discussed in 4.5.3, a feature of oral delivery in this sentence as it assumes authentic interaction. The use of person and place deictics (you ; this) show that sentences used to refer to the visual element in the lectures are marked by features of oral delivery. The use of initial "and" in this context is congrunent with its characteristic of being a feature of spoken language. Finally we may observe that the use of deictic "this" is relevant to the visual element by reference to which it is interpretable. Unlike the previous cases (see 4.5) where relevance has to do with the background knowledge of the listeners and their assumptions concerning the speaker's authority, relevance here is simply established by reference to the immediate environment.

4.7 Sentence-initial connectives in relation to the structure of the text

This section discusses first the use of sentence-initial connectives in relation to the structure of the text. Then some aspects of this use are interpreted in relation to the spontaneous spoken versus planned written distinction.

Connectives are found at two different levels of the structure of the text. This can be seen in the following example:

(17) but how can we derive. the stamen the anther and the ovary. from leaf like structures.. well. it's possible. to. derive the stamen.....

(333532)

"But" in this example, initiates an utterance which announces a shift of sub-topic. It is part of an utterance which introduces a sub-topic which in this case is about the derivation of the stamen, the anther and the ovary from leaf like structures. In the structure of the text this utterance marks off the beginning of a new chunk (see 5.4.2.2). Connective "well" on the other hand, in addition to introducing an answer to a rhetorical question initiates the stretch of text which discusses the sub-topic announced in the preceding utterance. Both but and well here signal a boundary in the text. These boundaries are at different levels of the structure of the text. They are either in the internal structuring of a chunk of text (such as the boundary signalled by well in (17), or they mark off its peripheries exemplified in the boundary signalled by but in (17). This section studies the use of connectives in relation to the two levels of structuring in the text.

4.7.1 Sentence-initial connectives in relation to the macrostructure of the text

As discussed above, sentence-initial connectives in this data can be used to

initiate a sub-topic shift shift signalling utterance. "Now" in the following excerpt instantiates such a use:

(18) now let's see what's next...{so you can see that. er. both aporsity. and superfluity of. hormones can have. quite. dramatic effects.} now what I. want to do now. is tell you a little something. of the. chemical nature of hormones.

(129111)

'now' introduces a 'sub-topic shift signalling' utterance (interrupted by a "summary paraphrase" see 5.7.4) which in this instance, explicitly states that a new stage in the discussion is starting and is about "the chemical nature of hormones". Such an utterance functions as a marker of a boundary in the text (see 5.4.2.2). It opens a new stretch of text. In this use "now" is deictic, it signals a sub-topic shifting.

This use of 'now' is also dealt with in Halliday and Hasan (1976: 268), who describe it as a "marker of the opening of a new stage in communication". They rightly observe that in such a case, 'now' is reduced (produced on falling intonation and unstressed). It must be said, however, that in their analysis 'now' expresses a semantic relation as it belongs to a category of cohesive tie called 'conjunction' which has already been mentioned above in relation to additive 'and'.

Sinclair & Coulthard (1975), on the other hand, offer a more appropriate approach to the analysis of utterance initial 'now' as they focus on its function as a discourse marker.

> "occurring at the beginning of opening, focussing and framing moves and (it) indicates a boundary in discourse".
>
> (1975: 38)

"Now" is also partly identified on prosodic grounds: i.e. falling intonation and unstressed. It is observed in this data that 'now' in this use is always reduced.

The connective 'so' also assumes this function of a text boundary marker when it initiates a summary paraphrase as exemplified in (19) below:

(19) so it's a. sequence of several steps. which ultimately leads to the activation. of one or more cellular enzymes binding activated calcicalmogulin complex.

(130017)

'So" initiates an utterance which summarizes the preceding discussion about the properties of a particular protein hormone (calmogulin). This summary paraphrase is then followed by a 'sub-topic shift signalling' utterance initiated by 'but' (see appendix (1): 130019). In such a use "so" is also deictic. It signals the summary paraphrase.

The summary paraphrase also functions as a marker of a boundary in the text, as it closes a stretch of text (see 6.2.3.1.2). Connectives "now", "but", and "so" when they initiate a 'sub-topic shift signalling' utterance or a 'summary paraphrase' belong to utterances which demarcate the boundaries of a stretch of text. In this use, they are markers of the macrostructure of the text.

A third type of utterance which signals a boundary in the text and which is generally initiated by a sentence-initial connective, is the utterance which announces the end of the text. Such an utterance may be introduced by "well" or "and", examples:

(20) well I think I'll stop at that point. er. and and we'll take up the anatomical questions. in our second lecture today.

(230652)

(21) and I'll show you some examples of this.. next week

(130130)

"well" and "and" initiate utterances which terminate the text. These utterances also announce what is going to be done in the next lecture. In this sense they have a different scope compared with utterances which signal a shift in sub-topic and which apply within the same text. "Closing" utterances such as (20) and (21) cut across lectures, thus they cut across texts(4). However, as closing utterances they are also markers of the macrostructure of the text.

4.7.2 Sentence-initial connectives as signals of the internal structure of a chunk of text

4.7.2.1 Connectives as sentence initiator following a "sub-topic shift signalling utterance"

As discussed in relation to (17) sentence-initial connectives can be used to initiate a sentence which begins the discussion of the sub-topic announced in the preceding sub-topic shift signalling utterance. The connectives which are mostly found in this position are well (as exemplified in (17)) and "now", as exemplified in the following excerpt from the data.

(22) //now. er. the endodermis, is the explan necessary is the semi permeable membrane which makes root pressure possible// now to explain that. we'll have to consider the pathway that water follows...../

(230945)

Underlined now signals a boundary in the text. It is used here as a text organising element (see O.S. Al-Shabab 1986). The same applies to the use of well in (17) quoted earlier. Both "now" and "well" in these occurrences are metatextual (ibid).

4.7.2.2 Connectives as initiators of a sentence which follows a digression

As discussed in (2.2) Montgomery (1977) has proposed a descriptive model of discourse whereby he makes a distinction between "main" and 'subsidiary' discourse.

It is observed, in this data, that when main discourse is resumed it is frequently initiated by connectives 'and', in a few instances, by "so".

In the following excerpt:

(23) //today sadly. we reach the end of our. story that we've been pursuing for the last five weeks.{I can see grief stricken faces in the audience}. and we're going to talk about. the. last of the. angiospersm or. flowering plants.....

(33341)

The whole utterance introduces the topic of the lecture. It is interrupted by a digression (underlined in the excerpt) which is a metatextual remark relating to what is being said, then resumed. The sentence which resumes the introduction of the topic of the lecture is initiated by and. And signals return to main discourse. In this use it is essentially deictic: it indicates that the sentence being initiated is not to be related to the immediately preceding stretch of text, but to what has gone before (Levinson 1983 : 84). And in this context has also retained its function of being additive (see 4.4.1).

4.7.2.3 Sentence-initial connectives as initiators of a feature of the internal structuring of a chunk of text

Connectives may initiate two types of paraphrase (the equivalence and the inference paraphrase) which are found in the internal structuring of a chunk of text (see 6.3.4.1. and 6.4.2.1); example:

(24) but they do exhibit. specificity. and they interact. with. receptors. which are located. in the cytoplasm. of. the cell..

(24a) so we get.within the cytoplasm. a steroid. receptor complex..

(130052)

So initiates an inference paraphrase (24a) i.e. a paraphrase whose information content is implied by the information content expressed in (24) (see 6.4).

The inference paraphrase is found within a chunk of text demarcated on one hand by a "sub-topic shift signalling" utterance and on the other by a "summary paraphrase". Here again "so" is deictic. It signals the paraphrase.

4.7.3 Features of spontaneous spoken versus planned written discourse in the use of connectives in relation to text structure

4.7.3.1 Features of spontaneous spoken discourse in the use of connectives in relation to text structure:

1 - Resumptive "and" and "so"

We saw in 4.7.2.2 above that sentence-initial connectives "and", and "so" when they are deictic, can be found in this data to signal return to main discourse after a digression.

The occurrence of a digression in spoken discourse may have its counterpart in written discourse in the form of a parenthetical remark. However, return to main

discourse would not be expected to be initiated by a connective. It would rather be signalled by the closing of the brackets. The use of a textual connective in this resumptive function can be said to be largely restricted to spoken discourse.

2 - "So" as paraphrase initiator

As discussed in 4.4.4 sentence initial "so" is mostly related to the use of "paraphrase" (see Chapter Six) which is a prominent feature of the lectures. As pointed out then and as will be discussed again in Chapter Six, the heavy redundancy which the use of paraphrase introduces in the text is more readily tolerated in speech (see Darian 1983) than in academic writing. Therefore we can say that this use of "so" in paraphrase typifies speech rather than writing.

3 - "Now and "but" as initiators of a sub-topic shift signalling utterance

The use of the sub-topic shift signalling utterance manifests the explicit signalling of the organisation of the content of the text (see Chapter Five). It is rare to find a spontaneously produced text, i.e. a text produced under the constraints of real time, which would show an explicit planning of what is going to be said. In spontaneous speech, speakers do not have time to organise and announce what they are going to talk about. Therefore sub-topic shift signalling can be seen as typical of a planned text. However, the use of connectives to introduce the sub-topic shift signalling utterance would not be much expected in a written academic text. Academic discourse would use other devices to signal a shift of topic/sub-topic like starting a new paragraph or using a sub-title. Thus the occurrence of "now" and "but" as sub-topic shift signalling utterance initiator is typical of a spoken text.

4.7.3.2 Features of planned written discourse in the use of sentence-initial connectives in relation to text structure

As discussed in 4.7.1 the sub-topic shift signalling utterance, generally initiated by "now" or "but", and the summary paraphrase, generally initiated by "so", are text boundary markers. They break down the text into chunks. This chunking of the text is topical, i.e. it is based on the boundaries between sub-topics. This is a manifestation of the structuring of the information content of the lecture. As pointed out earlier, systematic and explicit structuring of the content of a text would not be expected in a spontaneous spoken text. This is rather typical of a planned text, i.e. a text which has undergone forethought and organisation before it is delivered. Consequently, when connectives introduce a sub-topic shift signalling utterance they belong to an utterance which manifests a feature of a planned written text. Such a feature is also seen in the use of utterances which announce the end of the text.

4.8. Summary

In summary, features of spontaneous spoken versus planned written discourse in the use of sentence-initial connectives as dealt with here, can be presented in the following five points which relate each to the use of one (or more than one) specific connective.

1 - "Well"

As seen in 4.5.3 the functions of "well" in spoken and written language are clear cut. Most uses of "well" in this data belong to spoken language. They are: (1) "well" as editing marker of a reformulation ; (2) "well" as answer initiator following a rhetorical question.

2 - "And"

There are three aspects of the occurrence of "and" which establish it as a feature of spontaneous spoken discourse : (1) its occurrence in initial position

in the sentence ; (2) the high frequency of its occurrence in such a position (44.43% out of the total number of occurrences of all one word sentence-initial connectives found in the data (see table (12)) ; (3) one of the functions it assumes in the text, namely when it is resumptive.

3 - "So"

Its use as paraphrase initiator in the lectures is characteristic of a spoken text.

4 - "Now" and "but"

Their use as sub-topic shift signalling utterance initiator is typical of a spoken text.

5 - "Now", "but" and "so"

When "now" and "but" initiate a sub-topic shift signalling utterance, they can be said to belong to an utterance which marks off the explicit signalling of the planning of the text. In this sense, their use here is also a manifestation of a feature of a planned written text.

Finally it must be pointed out again that some of the uses of connectives are interpretable in the light of the didactic function of the lectures. These are, for example, the use of "well" as answer initiator following a rhetorical question; the use of "now" and "but" in a sub-topic shift signalling utterance and the use of "so" in a summary paraphrase.

4.9 Conclusion

Three main points emerge from this analysis of connectives:

1 - In relation to spontaneous spoken versus planned written language features of spoken language seem to prevail in the use of these sentence-initial connectives.

2 - The didactic function of the lectures is important in the way these lectures make use of features of spoken language.

3 - Sentence-initial connectives are found at different levels of the structure of the text:

- they are found at the level of the macrostructure of the text when they are used as initiator in the sub-topic shift signalling utterance and the summary paraphrase which demarcate a chunk of text.
- they are found in the internal structure of a chunk of text.

Footnotes

(1) See 3.1.1 for a discussion of the criteria used to identify sentence boundaries. We may point out again that the majority of initial "ands" are not open to doubt on grammatical grounds, as can be seen in the following example

> //the system which produces these chemical messengers is referred to as the endocrine system// and by that we mean any cell or group of cells which is able to manufacture or elaborate a secretion/ which then passes into the blood stream...../

(128418)

Here, "and" is used to introduce a new sentence, where written language would have used the anaphoric pronoun "which" to introduce a relative clause ("which refers to any cell or group of cells...."). The anaphoric pronoun would be NP subject of the verb phrase in the relative clause.

However for a comprehensive study of this question we need to investigate the use of "and" in other variaties of English, notably in conversation (where there is a change of speaker), and in descriptions of live on-going activities (such as a football match) where utterances initiated by "and" refer to different actions and events.

(2) See Sperber & Wilson (1981) who subsume all four maxims under the maxim of relevance.

(3) Bentley, P.J. (1977): "Endocrinology" in L. Goldstein (ed) Introduction to Comparative Physiology. New York: Holt Rinehart & Winston (pp 402-476). Weier, E.T.; Stocking, R.C., Barbour, M.G. & Rost, T.L. (1982) Botany. An Introduction to Plant Biology. 6th edition New York: John Wiley & Sons Inc.

CHAPTER FIVE

SUB-TOPIC SHIFT SIGNALLING IN BIOLOGY LECTURES

5.0 It was observed in the previous chapter that one function of the textual connectives 'now' and 'but' is to initiate a "sub-topic shift signalling" utterance. This chapter deals with "sub-topic shift signalling" in biology lecture discourse. This signalling is discussed in terms of the types and the grammatical and textual features of the signalling utterance. These three aspects are handled in the light of the spontaneous spoken versus planned written language dichotomy.

5.1 Sub-topic shift signalling in the lectures

5.1.0 At certain points in the lectures, the speaker announces what he is doing 'next', what the next fragment of text is about. This is referred to as sub-topic shift signalling. A close analysis of the data shows that there often is a shift of sub-topic and this shift is signalled, i.e. it has a surface manifestation. The following two extracts illustrate cases of sub-topic shift signalling:

(1) now let's see what's next... { } now what I. want to do now. is tell you a little something. of the. chemical nature of hormones.

(129111)

(2) now the function of the haemoglobin. {or let's put it another way.} the haemoglobin has two functions.

(230339)

In each extract a particular aspect of the topic of the lecture is explicitly identified as the next part to be dealt with: in (1) it is "the chemical nature of hormones" in (2), "the function of the haemoglobin".

Further analysis reveals that the signalling of a shift of sub-topic may be realised by other grammatical structures such as a question:

(3) well where did they come from. can we make any sensible suggestions about where. the flowering plants came from.
(33056)

which in this excerpt is a rhetorical question. This and other forms of signalling of shift of sub-topic are the concern of this analysis.

5.1.1 Topic and sub-topic: Scope of the sub-topic shift signalling (STSS) utterance

The notion of sub-topic is used here to refer to what a fragment of text is about. "Topic" refers to what the whole text of a lecture is about (see Brown & Yule 1983 : 70). It may also refer to the content of a series of lectures. Sub-topic is thus part of the topic of the whole lecture. Consequently two types of signalling utterance are outside the scope of the STSS utterance in this study. They are:

(i) an utterance which introduces the topic of : (a) a lecture, or (b) a series of lectures. Such an utterance is produced at the beginning or very early in the lecture. this is seen in the following example:

(4) okay. today. we begin our look. at the third and last. of the three major groups of land plant. { } and we're going to be looking at the third group. the division spermatophyta. or seed plants. and we're going to be looking at the most primitive of them. that are generally described as the gymnosperms...

(33281)

(5) I'm going to be. introducing to you. talking about. concepts of hormones for. the next five weeks..

(12831)

In excerpt (4) the utterance introduces the topic of the whole lecture as being about "the gymnosperms", in (5) it announces the topic of a series of lectures as being about "hormones".

(ii) an utterance which announces what will be done later in the lecture.

This is exemplified in the following excerpt:

(6) and again we'll come back to examples of their activities a little later..

(12859)

This utterance does not announce what is coming next in the lecture. Therefore it is not treated as a sub-topic shift signalling utterance.

On the other hand, the sub-topic shift signalling utterance is never found at the end of the text. It may occur at the beginning of the lecture but it generally belongs to the core of the text of the lecture.

5.1.2 Grammatical status of the sub-topic shift signalling (STSS) utterance

Excerpts (1) to (3) are instances of the signalling utterance where it is realised by a full grammatical sentence. The three excerpts are composed each of two sentences. Excerpt (3) of three sentences. The signalling utterance is always realised in a complete sentence which can be one or more than one clause. It is never realised by just one word or by an unfinished structure. It may contain a revised structure or a repetition of a textual item (see 5.7.1) but it is always complete as a grammatical structure. In excerpt (2) quoted above, the signalling utterance is interrupted and partly reformulated. However the reformulated utterance is a full grammatical sentence ("the haemoglobin has two functions").

5.2 Identifying the sub-topic shift signalling (STSS) utterance

This section deals with the nature and types of criteria used to identify the STSS utterance in this analysis.

5.2.1. Criteria: nature of criteria for identifying the sub-topic shift signalling (STSS) utterance

The criteria to be used to identify the signalling utterance are of six different kinds:

(i) Metatextual

The shift to a new sub-topic is explicitly verbalised. The speaker explicitly states that he is moving to a different aspect of the topic. In excerpt (1) quoted above the speaker announces that he is moving to a particular aspect of the topic of "hormones" namely that he is next dealing with "the chemical nature of hormones". The first type of criterion is then metatextual: the signalling utterance or a part of it is about the coming fragment of text.

(ii) Semantic

A shift of sub-topic means a shift in the content of the text, i.e. what is being talked about. The signalling utterance must introduce the new sub-topic. The next type of criterion to identify the STSS utterance is semantic.

(iii) Syntactic

It is often the case that there is a change of structure in the text. This change of structure constitutes a break which can suggest a shift to a new sub-topic (see Brown & Yule 1983 : 98). One type of criterion to apply along with the semantic criterion, is syntactic. The signalling utterance is realised in a grammatical structure different from that of the preceding utterance.

(iv) Phonological

The discourse of lectures being realised in the oral medium, some features of the signalling utterance are phonological. Studies of the role of prosody in marking topic boundaries have shown that pitch can mark a change from one topic to another (Brown and Yule 1983 ; Johns-Lewis 1986: xxi). The next type of criterion applied in this analysis is phonological. It applies in combination with the semantic criterion.

(v) Textual

A chunk of text often closes with a summary paraphrase (see 6.2.3.1.2) The following chunk of text may open with a signalling utterance which introduces the next sub-topic. The next type of criterion is textual i.e. whether or not the signalling utterance follows an SP. It applies in combination with the semantic and the phonological criterion. It may also apply with the metatextual criterion.

(vi) Non-verbal

Biology lectures make heavy use of visual aids (slides, transparencies, diagrams drawn on the board). A change of slide may accompany the signalling utterance. The next type of criterion is non-verbal. It applies in combination with either the metatextual criterion or the semantic and the phonological criteria.

5.2.2 Types of criteria

5.2.2.1 Metatextual criterion

The metatextual criterion applies where the signalling of a sub-topic shift is verbally explicit. The signalling utterance or part of it is about the text; it states what the coming stretch of text is about. i.e. it is metatextual. Example:

(7) so now I'm going to look at, the, { } I now want to look at the, structure and function of. the root nodule at a more. at a cellular. and anatomical level.

(230349)

The first part of the utterance which here is reformulated (underlined) indicates what the next fragment of text is about, and it indicates that it has not been dealt with before.

The metatextual criterion excludes all the other types of criterion. When it applies it is enough to qualify an utterance as a STSS utterance.

5.2.2.2 Semantic criterion

The signalling utterance introduces the next sub-topic. Therefore, it must contain new information. As previously stated (section 5.1.2) the signalling utterance is generally syntactically complete, i.e. it is realised as a full grammatical sentence. New information in the clause is that part of the clause which specifies what the sub-topic is. The rest of the clause provides given information, i.e. information which is either recoverable from the co-text or is inherently given. (see Halliday 1985: 274; Young 1985: 155). How is new information identified in this study? This is discussed next.

5.2.2.2.1 Identifying New information (NI)

1 - Problematic application of phonological criteria in identifying NI The role of phonological features in the distribution of given and new information is acknowledged. However, its application involves difficulties.

Following Halliday (1967; 1976; 1985), new and given information are used to refer to the constituents of an information unit. The information unit is independent from the grammatical units of sentence and clause and is identified by intonational features (ibid). Two intonational features are identified as playing a role in the way (i) the speaker allocates given and new information in his speech; (ii) the speaker organises his speech into information units. They are:

(i) the use of pitch prominence (Halliday: 1967). Tonic prominence, i.e. the main pitch movement in the tone group falls on the element carrying new information. It is the only part of the information unit which carries tonic prominence. However, experimental studies have shown that:

- pitch prominence may fall on more than one element in the information unit.
- phonetically trained listeners do not always agree on the allocation of a

tonic prominence. (Brown & Yule 1983: 164).

(ii) Pauses as markers of tone group boundaries: Pauses have been proposed for the recognition of tone group boundaries (Brown & Yule 1983: 161). But here again, their apparently multiple function makes them a difficult tool of analysis: there are filler pauses, planner pauses, there are pauses which occur along with non-verbal activities such as drawing a diagram on the board, or looking for the right slide, or checking the notes. It is not always clear which function(s) they are used for in a given instance. So, until a taxonomy of types and distribution of pauses is worked out, their use is problematic.

Because of the difficulties outlined above in applying phonological criteria, new information in this study is essentially identified by textual criteria.

5.2.2.2.2 Textual criteria for identifying new information

There are five kinds of textual criterion which may apply to identify new information in the signalling utterance.

(1) The first criterion is that what constitutes new information occurs in the utterance where the metatextual criterion applies,. i.e. it occurs in an utterance where the signalling of sub-topic shift is explicit. (see Excerpt 7 above). Such an utterance necessarily introduces the new sub-topic which is the new information in the utterance. This criterion, where it applies, is enough to identify new information in the signalling utterance.

(2) The second criterion is that it is not recoverable from the preceding co-text, i.e. it has not been previously dealt with. Example:

(8) now in many instances. the cells which produce these chemicals. are grouped together. in nice. clear cut discrete structures. called. glands..

(128441)

New information is underlined in the excerpt. The technical term 'glands' meaning structures made up of a group of cells is introduced for the first time in the lecture. This criterion must apply where criterion (1) does not apply.

(3) The third criterion is the use of the naming verb 'call' which provides evidence that the term(s) following "call" is (are) selected by the speaker as the new piece of information. This criterion applies along with criterion (2). in:

(9) but from another site. in the pancreas. emerges. a new hero. called. glucagon.

(129337)

The technical term "glucagon" is being introduced for the first time. In the structure of the utterance it is modified by the naming verb "call".

(4) The fourth criterion is that it is further qualified as what is going to be dealt with in a statement which follows the clause which contains new information:

(10) now the mature sporangium. has a unique structure. it looks something like. this..

(332338)

The second sentence (which starts with 'it looks') is produced with reference to a visual. It specifies further that the sub-topic is about the structure of the mature sporangium. This fourth criterion applies in combination with (b).

In summary, criterion (1) and (2) apply independently. Criterion (3) and/or (4) apply in combination with criterion (2).

5.2.2.3 Phonological criterion for identifying the sub-topic shift signalling(STSS) utterance

The next criterion to apply in identifying a shift of sub-topic is phonological. It was observed earlier, in 5.2.1, that it is generally held that a change of pitch can mark off a boundary between topics. When this criterion applies in this analysis, the sub-topic shift signalling utterance starts on a high pitch level; the preceding utterance which ends the closing sub-topic, ends on a low pitch range. (See Brown and Yule 1983: 101). In the following example:

(11) and they don't survive to the present day.. however. some of the plants. with megaphyllous leaves. some of those early tree ferns. also. independently. evolved this trick of retaining the megaspore. on the plant. and they. rapidly expanded in the late carboniferous period. and they were called. the seed ferns...

(33292)

The signalling utterance starts with 'however' which is produced on a raised pitch which decreases as the utterance proceeds. The preceding utterance ends on a low pitch.

The phonological criterion applies in combination with the semantic criteria.

5.2.2.4 Syntactic Criterion

This criterion applies where the STSS utterance is realised in a grammatical structure different from that of the utterance which precedes it:

(12) but how can we derive. the stamen the anther and the ovary. from leaf like structures..

(333532)

Here, the STSS utterance is realized in a rhetorical question which identifies the next sub-topic as being about the structure of the stamen, the anther, and the ovary

and the way they relate to leaf like structures. This type of criterion applies in combination with the semantic criteria.

5.2.2.5 Textual criteria

There are two types of textual criteria which can apply in the identification of a STSS utterance. They apply in combination with the semantic and the phonological criteria. They are:

(i) The use of a textual connective as initiator of an STSS utterance.

(ii) The occurrence of a summary paraphrase preceding the STSS utterance.

Criteria (i) applies to the following excerpt:

(13) now the. absence of a nucleus. is another curiosity.

(231627)

The utterance announces the next aspect of the structure of the sieve tube element to be dealt with, namely 'the absence of a nucleus". It is introduced by the connective now. In such a position the textual connective is produced on low pitch, which then rises on the next element.

Criterion (ii) on the other hand applies to excerpt (12) quoted above which follows the summary paraphrase:

(14) so. the flower. is four whorls. of modified leaves.. as I've said.
sepals and petals. are relatively easy. to derive from leaves..

(333530)

Excerpt (13) summarizes the preceding discussion which is about a description of parts of the flower. It ends a sub-topic. Therefore it sets the expectation that a new sub-topic is starting which might be announced in the next utterance. Excerpt (11) immediately follows the summary paraphrase and announces the next sub-topic to be dealt with.

Either of the two textual criteria discussed in this section apply along with the semantic and phonological criteria.

5.2.2.6 Non-verbal criterion

As stated above biology lectures make an extensive use of visual aids, mainly slides, transparencies and/or diagrams drawn on board as the lecture proceeds. A change of slide may announce the shift to the next sub-topic:

(15) let's have a look diagrammatically. at a cycad and look at a little.
of its life cycle.

(332947)

As the utterance is produced the new slide is made visible to the audience. Reference to the visual is lexically realised in the text by the use of the terms: diagrammatically, let's have a look at. It may also be realised by the use of a deictic element. (see section 5.5.1 below).

This criterion applies in combination with the semantic and phonological criteria.

In summary, the criteria discussed in sections (5.2.2.1/2/3/4/5/6) do not have to apply all at the same time for a given utterance to qualify as a sub-topic signalling utterance:

- criterion (5.2.2.1) (metatextual) excludes the other criteria. Where it applies, it is enough to qualify an utterance as a signalling utterance.
- criteria(5.2.2.2) and(5.2.2.3) have to apply where criteria (5.2.2.1) does not apply.
- criteria (5.2.2.4/5/6) apply each in combination with (5.2.2.1) or (/2) and (/3).

Excerpt (15) above combines the following criteria:

- It expresses 'new information; which is the life cycle of the cycad', a sub-topic which has not been dealt with before in the lecture (Criterion 5.2.2.2).
- The beginning of the utterance is marked by a change of pitch (it starts on a high pitch) (Criterion 5.2.2.3).
- The utterance is realised in a syntactic structure different from that of the utterance which precedes it. It is expressed in the imperative mood. (Criterion 5.2.2.4).
- It is accompanied by a change of slide. (Criterion 5.2.2.6).

4 types of criteria apply in this excerpt: criteria (5.2.2.2./3) (which are compulsory) criteria (5.2.2.4/6) which are optional.

5.3 Types and Distribution of the sub-topic shift signalling STSS utterance. Features of spontaneous spoken versus planned written language in the use of types of the signalling utterance

On the basis of the criteria set up in the preceding section, three types of signalling utterance are identified:

(1) - the metatextual comment

(2) - the statement

(3) - the change of structure

They are dealt with in turn, in the following sections.

5.3.1 The metatextual comment

The metatextual comment is a statement about the text. It indicates what the next fragment of text is doing; example:

(16) and. probably. in the remainder of the time. erm I think this is all I'll have time to show you. I'm going to. just give you a fairly relaxed amble. through. some of the dramatic consequences which can ensue. when these hormones. are either present in excessive quantities. or. in. diminished quantities.. (128749)

In this excerpt the utterance announces that the next part of the text is about a discussion of the consequences of hormone malfunctioning and that it is illustrated with slides.

As pointed out above, the metatextual comment is an explicit statement about what is to come next in the text. It is an explicit manifestation of the organisation of the content of the text. It would be rare, to say the least, that a participant in a conversation announces what he is going to talk about in this way. It can thus be interpreted as a feature of a text which has undergone a certain degree of planning.

5.3.2 The statement

The second type of STSS utterance is the statement. It is primarily identified by the application of the semantic and phonological criteria. The textual and/or non-verbal criteria might also apply. It introduces the sub-topic of the following stretch of text but it does not contain a metatextual element as metatextual comments do. Statements are of two types:

(1) - a type where new information as the central grammatical element in the signalling utterance is conveyed by a noun phrase (NP).

(2) - a type where new information is conveyed by the predicate.

Examples (13) (quoted earlier) and (17) below are instances of each type of statement:

(13) now the. absence of a nucleus. is another curiosity.

(231627)

(17) but the classical. demonstration. of the existence of. hormones or at least a hormone. came. from a gentleman called Bartol. in 1849.

(128632)

In (13) the NP carrying new information is "the absence of a nucleus" (underlined in the excerpt). In this instance it functions as subject of the verb, a rare case in the data. It is rather generally found to function as complement of the predicate in the

structure of the clause. In (17) on the other hand, new information is realized in the predicate (underlined in the excerpt). The new sub-topic is part of an early discussion of the discovery of hormones. The signalling utterance introduces the discussion of experiments done by the scientist Bartol which led to the demonstration of the existence of a hormone.

On the basis of the distinction between these two types of statement this type of signalling utterance will be referred to as statement^1 and statement^2 where in statement^1 NI is realized in a NP, and in statement^2 it is realized in the predicate as a whole.

The statement type of signalling utterance whereby the next sub-topic introduced is comparable to the notion of "topic sentence" discussed in studies of the structure of the paragraph in written language (see Christensen 1967). Christensen's study suggests that the use of a "topic sentence" to introduce a given paragraph is a normal feature of the structuring of the paragraph in written language. In speech, on the other hand, and as Brown & Yule (1983 : 69) observe: "Speakers often do not provide such explicit guide-lines... which can be used to demarcate topics and thus to demarcate chunks of texts". In other words conversational speech is not generally expected to use explicit signalling such as a "topic sentence" to introduce a new topic. This is not to say that the use of the "topic sentence" is exclusive to written language. Rather, it is more expected in written language. In this sense, it is typical of written language.

5.3.3 The change of structure

There are two types of change of structure which signal a shift of sub-topic; interrogative and imperative clauses. They are discussed in the next two sections.

5.3.3.1 Interrogative clause

They are either realized as (1) real questions or (2) rhetorical questions. Examples (18) and (19) below instantiate each type of interrogative:

(18) //now can you explain that// why is the sap that's bleeding from the stump here of the root system as a whole more dilute than is found than than is coming from the root nodule itself//

(230622)

(19) //so can we therefore in the face of their diversity of structure say anything about their general properties//

(129225)

In (18) the signalling utterance is formulated twice. The utterance is realized as a yes/no question followed by a WH question. In both formulations it is a question addressed to the audience. This is seen in the use of the pronoun you which addresses the audience; it is thus a real question. In (19), on the other hand, the question does not seek an answer from the audience. It is a rhetorical question (see 4.5.2.1). We note the use of the pronoun "we" to involve the audience in the on-going argument. The use of both "you" and "we" is here a manifestation of the interactive (see Sinclair 1980) aspect of these interrogatives.

Real questions (as exemplified with 18) assume the presence of the addressee or audience (Lakoff's visibility criterion 1979). You in (18) refers specifically to the audience present in thelecture(as a language event). You is here person deictic. The use of real questions is restricted to oral delivery. It is a feature of spoken language in the lectures. Rhetorical questions on the other hand can be found in a written academic text. Here, their use can be interpreted as a manifestation of pedagogic discourse (see 4.5.2.1), i.e. it is a manifestation of the way the speaker monitors his speech pedagogically.

5.3.3.2 Imperative clause

It is found that the STSS utterance is often realised in one type of imperative clause, "the first person" type of imperative (see Huddleston 1971: 53; 1984: 362). It is exemplified in the following excerpt:

(20) let's look inside. one of these. megasporangia. to complete the life cycle of the cycad..

(333024)

It is marked by the use of 'let' followed by the contracted form of "us" ('s'). "Let's" initiates the STSS utterance. Young (1985) describes this type of imperative clause as jussive inclusive imperative (Young 1985: 70). The pronoun "us" ('s') is inclusive; it refers to the speaker and the audience. (Young; Huddleston ibid). This type of signalling utterance is always accompanied with the display of a visual (slide or reference to a diagram). As the utterance here is produced, an element of the immediate environment (i.e. the slide) is visible to both the speaker and the audience. "Us" here, refers to the speaker and to the audience who are present in the same language event. Therefore, the use of this type of utterance is a feature of oral delivery. It is a feature of spoken language in the lectures.

Table(13) overleaf shows the frequency of occurrence of the three types of change of structure in the three sets of lectures.

Texts	No of Occurrences of Change of Structure	FO of Real Questions	FO of rhetorical Questions	FO of Imperative Clauses
Set 1 (3 lectures)	(6)	33.33% (2)	50% (3)	16.66% (1)
Set 2 (3 lectures)	(10)	40% (3)	55.55% (5)	11.11% (1)
Set 3 (3 lectures)	(16)	0 (0)	37.5% (6)	62.5% (10)
Total (9 lectures)	(32)	18.75% (5)	43.75% (14)	37.5% (12)

Table (13) : Frequency of occurrence of the three types of change of structure in the three sets of lectures

In the data, rhetorical questions are the most frequent (45.16 occurrences per one hundred occurrences of change of structure type of STSS utterance). The rhetorical question is also mostly used in set (1) (50%) and (2) (55.55%). With speaker (3) it is the imperative clause which is mostly used: (62.5%), this will be discussed again in Chapter seven.

5.3.4 Marginal type of the signalling utterance

There are a few cases where the signalling utterance combines the three types of STSS utterance identified in the previous sections, example:

(21) the evolution of the flower. is an extremely. complex. subject.
Statement[1] and all we are going to be able to do. this morning. is. look
Metatextual briefly at some. major trends. in the evolution of the flower.
C we cannot look at the subject in detail. so let's look. at some
Imperative S major trends.

(333829)

The signalling utterance opens with a statement (state[1]); followed by 2 metatextual comments; followed by an imperative clause (a type of change of structure). Such utterances are classified on the basis of the type of utterance which first indicates the

following sub-topic. Thus, in the excerpt quoted above the signalling utterance is a metatextual comment. The new sub-topic is expressed in the opening sentence (statement[1]) but it is specified in the metatextual comment.

5.3.5. Summary table of the frequency of occurrence of types of sub-topic shift signalling (STSS) utterance in the corpus

The distribution of the three types of STSS utterance in the corpus is presented in table (14):

Texts	Number of words	F.I. For Metatextual Comment	F.I. For Statements	F.I. For Change of Structure	Total Number of occurr of STSS
Set 1 3 lectures	14,200	(11) 0.77	(35) 2.46	(6) 0.42	(52) 3.66
Set 2 3 lectures	14,200	(16) 1.12	(19) 1.33	(10) 0.70	(45) 3.16
Set 3 3 lectures	14,600	(9) 0.61	(36) 2.46	(16) 1.90	(61) 4.17
Total 9 lectures	43,000	(36) 0.83	(90) 2.09	(2) 0.74	(158) 3.67

Table (14) : Summary table of the frequency of occurrence of each type of STSS utterance per 1,000 words in the corpus

If we look at the whole corpus (9 lectures) the frequency index for the STSS utterance is 3.67 occurrences per 1000 words. The type of STSS utterance mostly favoured is the statement (with a frequency of 2.09 per 1,000 words in the corpus). At the level of individual speakers the distribution of the 3 types of STSS utterance is comparable with the three speakers.

5.4 Textual features of sub-topic shift signalling

If the criteria set up in 5.2.2.5 apply then the STSS utterance can be:

- introduced by a sentence-initial connective

and/or

- preceded by a summary paraphrase.

This may apply to any of the 3 types of signalling utterance discussed above.

5.4.1 Sentence-initial connectives in the signalling utterance

Sentence-initial connectives which can introduce the signalling utterance generally are: "now", "but" and "and". They can occur with any type of signalling utterance. (see examples 17 and 18 quoted in 5.3.2 and 5.3.3.1).

5.4.1.1 Distribution of sentence-initial connectives in the three types of sub-topic shift signalling utterance

The frequency of occurrence of the sentence-initial connectives used in the three types of STSS utterance is as follows:

Texts	F.O. of Metatext C+C	Total no of occurrence of Metatext C	F.O. of Stat + C	Total no of occurrences of Statements	F.O.of Ch of Str + C	Total no of occurrences of Change of Structure	F.O. of STSS + C	Tot no occurrences of STSS
Set 1 3 lectures	(5) 45.54%	(11)	(23) 65.71%	(35)	(3) 50%	(6)	(31) 59.61%	(52)
Set 2 3 lectures	(9) 56.25%	(16)	(12) 63.15%	(19)	(7) 70%	(10)	(28) 62.22%	(45)
Set 3 3 lectures	(3) 33.33%	(9)	(26) 72.22%	(36)	(5) 31.25%	(16)	(34) 54.73%	(61)
Total 9 lectures	(17) 47.22%	(36)	(61) 67.77%	(90)	(15) 46.87%	(32)	(93) 58.86%	158

Table (15) Frequency of occurrence of sentence-initial connectives in the three types of STSS utterance, (per 100 occurrences of each type of STSS utterance)

The frequency of occurrence of connectives is comparable in the three sets of lectures. The statement STSS utterance is the type of signalling utterance which is mostly initiated by a connective, with the three speakers. This seems logical since the other two types of STSS show other surface features which signal the shift of sub-topic. The metatextual comment contains a metatextual element which states what is going to be done next; the change of structure is by definition structurally different from the preceding utterance. On the other hand if we compare the distribution of connectives in each type of STSS utterance among the three speakers, we can observe that it is similar.

5.4.1.2 Features of spoken language in the use of connectives in the STSS utterance

The frequency of occurrence of sentence-initial connectives as STSS utterance

initiators is a dominant feature of the signalling utterance. As observed above more than half of the total number of the STSS utterances, in the three sets of lectures, are initiated by a connective (58/.86%).

As pointed out in (4.7.3.2), this use of connectives would not be much expected in a written academic textbook where the use of a sub-title or a new paragraph is generally the device employed to signal or introduce a new topic or sub-topic. If connectives are found at all in this use, they are much less frequent. Therefore we can consider this use of connectives as typical of a spoken text.

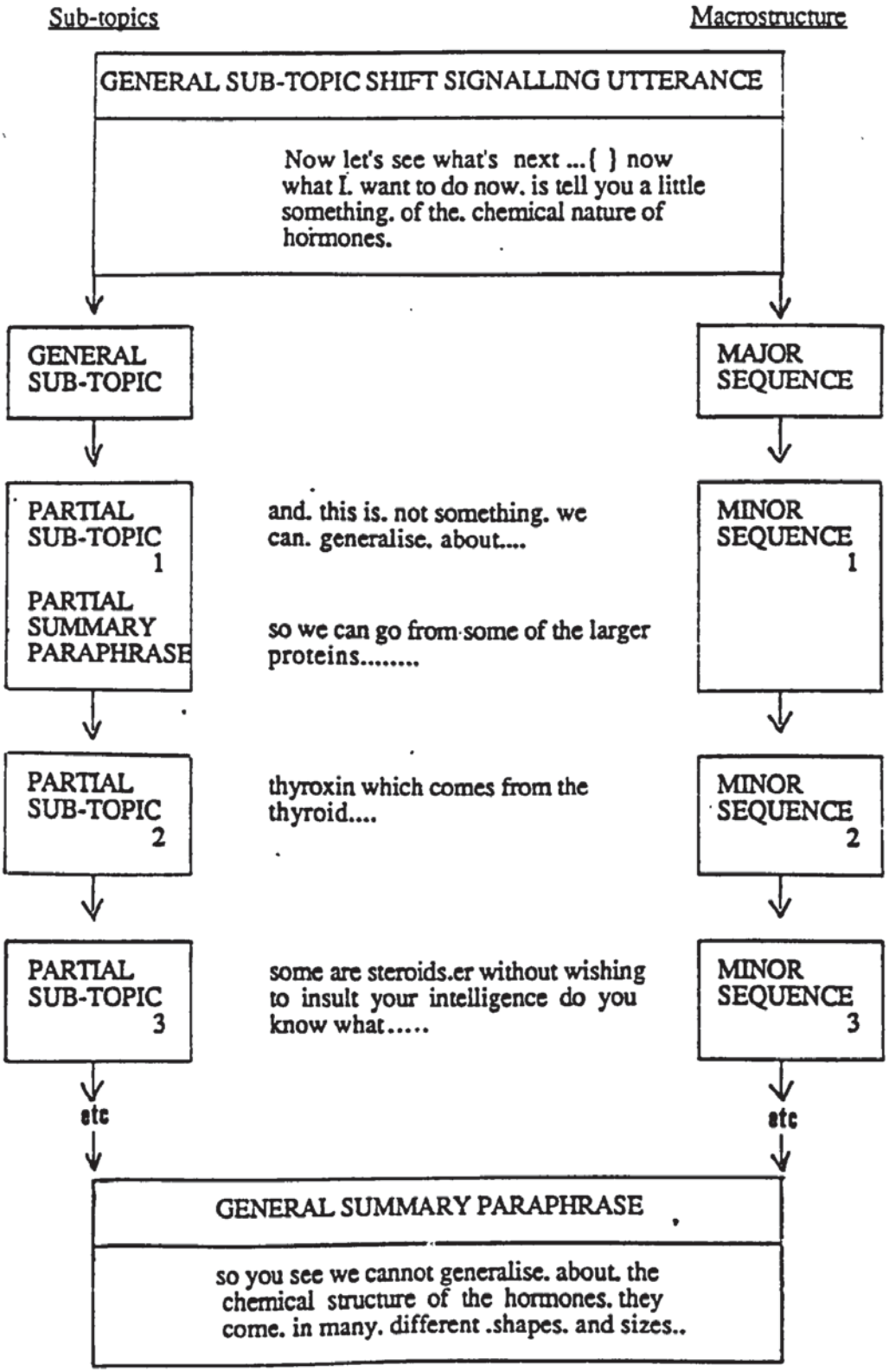
5.4.2 The STSS utterance in relation to the macrostructure of the text

5.4.2.1 General and partial sub-topic shift signalling

The relationship between sub-topics can be complex. A feature of this complexity can be seen by looking at (Figure 2), overleaf, where a sub-topic includes a number of sub-topics within it.

Figure (2)

SUB-TOPIC EXPANSION



(1) and this is not something we can generalise about....

(129114)

(2) thyroxin which comes from the thyroid in the neck....

(129139)

(3) some are steroids.....

(129149)

Therefore the sub-topic shift signalling utterance can introduce either a general or a partial sub-topic; excerpt (1), quoted earlier in (5.1.0), introduces the general sub-topic quoted in (Fig 2):

> now let's see what's next...{ } now what I. want to do now. is tell you a little something. of the. chemical nature of hormones.
>
> (129111)

This general sub-topic is about "the chemical nature of hormones". It is further broken down into partial sub-topics which are also signalled by a STSS utterance. Example:

(22) and finally. some hormones... are fatty acid. derivatives...

(129210)

The signalling utterance introduces the discussion of the structure of one type of hormone to illustrate one aspect of the "chemical nature of hormones" introduced earlier as a general sub-topic.

The general sub-topic is textually realised as a major sequence, the partial sub-topic as a minor sequence.

5.4.2.2. The sub-topic shift signalling (STSS) utterance and the summary paraphrase

As previously observed (see: 5.2.2.5) the summary paraphrase is a brief reformulation of a previous sub-topic. The analysis of the corpus shows that it is often followed by a STSS utterance, as excerpts (23) (a) and (b) below show:

(23) (a) so you see we cannot generalize. about. the chemical structure of hormones. they come. in many. different. shapes. and sizes.

(b) so can we. therefore. in the face of their. diversity of structure. say anything. about. their general properties.

(129224)

Part (a) is a summary paraphrase: it states again the main points of the preceding stretch of text or sequence (see Fig 2), which are that the chemical nature of hormones is very diverse and that it cannot be captured in terms of general descriptive features. Part (b) is a STSS utterance. It announces the sub-topic of the next sequence which is about "the general properties of hormones". Here again the sub-topic is a general sub-topic and it is realized in a major sequence. The point here is that the SP and the STSS utterance occur at sequence boundaries. The STSS opens a sequence, the summary paraphrase ends the sequence.

We saw in the previous section that the STSS utterance is either general or partial. The summary paraphrase can also be either general or partial (see 6.2.3.1.1). It either summarizes a major sequence as can be seen in excerpt (23a), or a minor sequence as exemplified in excerpt (24) below:

(24) so we can go. from some of the larger. proteins to. aggregations of only three. amino-acids.

(129137)

which summarizes the first partial sub-topic quoted in (fig 2) which was about

some types of hormones. Where the STSS utterance is preceded by a summary paraphrase it is often the case that a general summary paraphrase precedes a general STSS utterance and a partial summary paraphrase precedes a partial STSS utterance. If we consider the macrostructure of the text then a major sequence will open with a general STSS utterance and end with a general summary paraphrase. A minor sequence will have as boundaries a partial STSS and a summary paraphrase (this point will be picked up again in 6.2.3.1.1).

5.4.2.2.1 Types of sub-topic shift signalling preceded by a summary paraphrase

The analysis of the types of signalling utterance which follow a SP is tabulated in (16) overleaf. It shows that the use of a STSS following an SP is mostly exploited by speaker (1) with the three types of STSS (to be looked at again in Chapter Seven). In the corpus as a whole the distribution of STSS utterance following an SP is comparable for the three types of signalling utterance.

5.5 The sub-topic shift signalling utterance and visual aids

5.5.0 As pointed out earlier (5.2.1), biology lectures rely heavily on the use of non-verbal visual aids. A change of slide or reference to a diagram drawn on the board often accompany the signalling utterance. In the following extract:

(25) this is. er another example of. hormone excess.

(128825)

The signalling utterance is produced as the slide is shown to the audience. The deictic term this (underlined in the excerpt) is used to refer to the slide.

Texts	Total no of occur- Metatextual C	F O of Metatextual C + SP	Total no of occur- of Statement	F O of Stat + SP	Total no of occur- of Ch of str	F O of Ch of str + SP	Total no of occur- of STSS U	F O. of STSS + SP
Set 1	11	54.54% (6)	35	51.42% (18)	6	50% (3)	52	51.92 (27)
Set 2	18	11.11% (2)	18	5.55% (1)	9	33.3% (3)	45	13.33 (6)
Set 3	9	22.22 (2)	36	11.11% (4)	16	18.75 (3)	61	14.75 (9)
Total	38	26.31 (10)	89	25.84 (23)	31	26.31 (9)	158	26.58 (42)

Table (16): Frequency of occurrence of types of STSS utterance following an SP

The frequency of occurrence of each type of signalling utterance accompanied by a visual, calculated per 100 occurrences of each type of STSS utterance, is shown in Table (17):

Texts	Total no of occ of Metatext	F.O. of Metatext + V	Total no of occ of Statements	F.O. of Statements + V	Total no of occ of Change of Str	F.O. of Ch of Str + V	F.O. of STSS	F.O. of STSS + V
Set 1 3 lectures	(11)	(6) 54.54%	(35)	(15) 42.85%	(6)	(1) 16.66%	(52)	(22) 42.30%
Set 2 3 lectures	(18)	(8) 44.44%	(18)	(13) 72.22%	(9)	(6) 66.66%	(45)	(27) 60%
Set 3 3 lectures	(9)	(7) 77.77%	(36)	(34) 94.44%	(16)	(14) 87.5%	(61)	(55) 90.16%
Total 9 lectures	(38)	(21) 55.26%	(89)	(62) 69.66%	(31)	(21) 67.74%	(158)	(104) 65.82%

Table (17) : Frequency of occurrence of types of STSS utterance accompanied by a visual

The type of STSS utterance which are mostly accompanied by a visual are the statement and the change of structure in sets (2) and (3). Set (3) makes a heavy use of the visual with all types of STSS utterance. This will be looked at again in Chapter seven.

5.5.1 Reference to the visual in the sub-topic shift signalling (STSS) utterance

In excerpt (25) quoted above reference to the slide is textually realized by the use of the deictic element 'this'. Reference to the visual in the STSS utterance is mostly realized by the use of one of the deictic elements: "this, there, that", and "here". "This" is the most frequent in the whole data (65% occurrences of all deictic elements).

Reference to the visual may also be realized by the use of a lexical item such as: 'diagram', 'slide', or the prepositional verb 'look at' which signals that a visual is going to be/is being used:

(26) and one final diagram of pinus. to look at, a single ovule. prior. to fertilization.

(333223)

The use of 'diagram' and 'look at' indicate that a slide is being used as the next sub-topic is being announced.

5.5.2 Features of oral delivery in the interaction between the text and the visual

The use of non-verbal visual aids is not exclusive to spoken language. Non verbal elements (diagrams, pictures, tables) are used in written language. But where written language refers to a non-verbal element by the use of a figure between brackets, generally occurring outside the structure of the sentence in a text, spoken language uses a deictic element usually supported by showing a slide or pointing to a part of it. Thus in:

(27) the next fern that we'll look at. phyllites. this one. does not have. the characteristic. fern dissection..

(332524)

The structure "this one" would be indeterminate in written language because of the lack of what Lyons has called "the canonical situation of the utterance" (Lyons 1977 : 637). It is interpretable with reference to the slide which can be seen by the audience as the STSS utterance is produced. Further, the deictic reference of 'this' is interpretable from the view point of the speaker; it refers to something proximal to the speaker and not to the audience, what Lyons refers to as the "egocentric parameter of the canonical situation of the utterance". (ibid : 638). Another distinguishing feature is that in spoken language the visual is not part of the text. As pointed out in (5.3.3.2) it is an element of the immediate situation (see Ure& Ellis 1977:199) whereas in written language the visual element is part of the text.

We may conclude by saying that the type of interaction between the text and the visual discussed above is typical of oral as opposed to written academic discourse.

5.6 Grammatical features of the sub-topic shift signalling (STSS) utterance

5.6.0 The preceding sections (5.4 and 5.5) have dealt with aspects of the STSS utterance in relation to the use of sentence-initial connectives, of the summary paraphrase and the use of non-verbal visual aids. Features of oral versus written academic discourse have been identified. The present section will look at the grammatical features of the constituents of the signalling utterance. It will first deal with the main constituent of the signalling utterance, i.e. the constituent which carries new information. Then, the use of personal pronouns, and some aspects of grammatical complexity in the STSS utterance are investigated and discussed in relation to the unplanned spoken versus planned written discourse distinction.

5.6.1 New information carrier: types and functions

5.6.1.1 Types of new information (NI) carrier

There are four types of constituents which can carry new information in the STSS utterance: (1) noun phrase; (2) a prepositional phrase; (3) a predicate structure; and (4) a verbless clause.

1 The noun phrase (NP)

The NP carrying NI may be a technical term or a common word. A technical term used as NP can be simple, see excerpt (27) above where the technical term is "phyllites"; or compound as in:

(28) a second mechanism might be via gene activation

(130039)

'gene activation' is the compound technical term which is the next sub-topic to be dealt with. When NI is conveyed by a technical term the discussion is from the

specific (technical term) to the general (explanation/description of what is referred to by the technical term).

Common words used as NP new information carrier generally occur in compound expressions. This is exemplified in excerpt (19) quoted earlier (in 5.3.3.1), where NI is conveyed by "their general properties" (i.e. general properties of hormones). The following stretch of text discusses a number of properties of hormones. The discussion here starts with a general descriptive statement (hormones have general properties) then, it moves to the specific detailed account of each property.

2 Prepositional phrase:

Two types of prepositional phrase are found in this data to convey NI in the STSS utterance. They are the possessive genitive, and the appositive genitive. Examples (1), quoted earlier , and (29) illustrate each type.

(1) now what I. want to do now. is tell you a little something of the. chemical nature of hormones.

(129113)

(29) this morning I want first of all. to give you. a further example. of. this. concept of negative feedback...

(129610)

In (1) 'the chemical nature of hormones" is the new information. It is described as possessive genitive as it can be paraphrased as "hormones have a chemical nature". (Quirk et al: 1985: 321-322).

In (29), on the other hand, "this concept of negative feedback" is the new sub-topic. It may be reformulated as 'negative feedback is a concept'. "Negative feedback" and "concept" are semantically equivalent. (see Huddleston 1984: 270).

The prepositional constructions distinguished in (i) and (ii) realize two different relations between the NPs they relate.

3 Predicate structure

As pointed out earlier in this section, the predicate structure is one type of constituent which may convey new information. This is characteristic of two types of STSS utterance, described as statement (2), and interrogatives, example:

(30) well where did they come from. can we make any sensible suggestions about where. the flowering plants came from.

(33357)

The sub-topic announced in this signalling utterance is about the origin of flowering plants, i.e. where they come from. New information is conveyed here by the predicate.

4. Verbless clause

New information may also be conveyed by a verbless clause as the following excerpt,quoted earlier (26), shows:

> and one final diagram of pinus. to look at a. single ovule. prior. to fertilisation.
>
> (333223)

"A single ovule prior to fertilisation" is the sub-topic of the next fragment of text. It is a verbless clause. The full form would be: "a single ovule which is prior to fertilisation" or the co-ordinate construction" it is a single ovule and it is prior to fertilisation".

5.6.1.2 Functions of new information carrier

It is found that the new information carrier in the STSS utterance can assume four different functions in the structure of the clause. It can be:

(i) complement of the verb

(ii) complement of a preposition

(iii) complement of a naming verb

(iv) or it can be found in apposition

(i) Complement of the verb:

As observed earlier in (5.3.2), the new information carrier is generally found at the end of the clause (unmarked structure of the distribution of given/new information which is found in the rheme part of the clause). The constituent which carries new information (NP or PrepP) is generally complement or a constituent of the complement of the verb, as can be seen in excerpt (1) quoted above, where the prepositional phrase is part of the complement of the verb "tell". In the next excerpt:

(31) it's also characteristic of the phloem. that erm. these. vessel elements. have adjacent to them. companion cells..

(230821)

"companion cells" is an NP composed of a compound technical term; it is complement of the verb 'have' of the embedded "that-clause". 'Companion cells' is modified by "adjacent to them" which is placed here in premodifying position.

(ii) The NI carrier may be complement of a preposition:

This is exemplified in:

a second. mechanism.. might be. via gene. activation..

(130039)

"gene activation" is a constituent of the prepositional phrase "via gene activation". It is complement of the preposition via. It is also a constituent of an equative be construction.

(iii) The NI carrier may be in apposition:

(32) but. perhaps. the fern that will be. most familiar to you. in Britain. is the bracken fern. pteridium.

(33262)

'pteridium" NI carrier is in apposition to "the Bracken fern". It is in complete apposition as it can substitute the whole form "the Bracken fern". It is also constituent of a "be" equative construction.

(iv) The NI carrier may be complement of a naming verb;

This is illustrated in excerpt (8) quoted earlier

> but from another site in the pancreas emerges a new hero called glucagon
>
> (129337)

"glucagon" is complement of the naming verb "call". This is a recurrent structure in the text.

5.6.2 Personal pronouns in the sub-topic shift signallin g(STSS) utterance

There are four personal pronouns which may be found in the STSS utterance: 'I', 'you', 'we' and 'us'. Table (18) overleaf shows their distribution.

'I' is the pronoun mostly used except with one speaker who never uses it at all (to be dealt with in Chapter Seven). It is also mostly a feature of the metatextual comment type of signalling utterance, but it is also found in the statement type of STSS. Excerpt below (quoted earlier in 5.3.1) is an example of the metatextual comment signalling utterance which mostly exploits the use of pronouns.

> and. probably. in the remainder of the time. erm I think this is all I'll have time to show you . I'm going to. just give you a fairly relaxed amble. through. some of the dramatic consequences which can ensue. when these hormones. are either present in excessive quantities. or. in. diminished quantities...
>
> (128749)

As previously defined the metatextual comment is an explicit statement about the activity that is coming next in the lecture. The speaker often refers to himself (by

	Set 1			Set 2			Set 3			
Types of personal Pronouns	Personal pronoun in stat	Personal pronoun in Ch of str	Personal pronoun in metatext	Personal pronoun in stat	Personal pronoun in ch of st	Personal pronoun in metatext	Personal pronoun in stat	Personal pronoun in ch of str	Personal pronoun in metatext	Total no of occurrences of each type of pronoun
I	(5)	(0)	(9)	(5)	(0)	(18)	(0)	(0)	(0)	(37)
you	(3)	(1)	(6)	(2)	(1)	(3)	(1)	(0)	(0)	(17)
we	(2)	(2)	(3)	(2)	(2)	(4)	(3)	(3)	(5)	(26)
us	(0)	(1)	(3)	(2)	(1)	(1)	(0)	(10)	(2)	(18)

Table (18) : Distribution of personal pronouns in the three types of the STSS utterance

the use of the pronoun I) and/or addresses the audience (by the use of 'you') in stating what he is doing next.

Recurrent self-reference (i.e. use of 'I') has been described as typical of spoken language (Chafe 1982; Gruner et al 1967). It is the manifestation of the "egocentric style of spoken language" (Akinnaso 1982 : 102) where the speaker expresses his emotional state, what he thinks, using explicit forms (I) to refer to himself. In the excerpt above, it also occurs as a component of the metatextual element of the signalling utterance where reference to real time is made: "I think this is all I'll have time to show you". The speaker is explicitly referring to the period of time in which the lecture is taking place. The use of "I" here is a feature of oral delivery. The same applies to the pronoun you, used to address the audience. It is here, deictic (so is "I") it refers specifically to the audience present in the lecture (Lakoff's visibility criterion 1979). So although personal pronouns "I" and "you" are not very frequent in the STSS utterance (17.15 and 7.8 per 1,000 words respectively), nevertheless they are found in uses such as those exemplified above, which are typical of spoken language.

5.6.3 Aspects of grammatical complexity in the STSS utterance in relation to spoken versus written language

In Chapter Three some aspects of grammatical complexity in biology lecture discourse were looked at at clause level, in the whole data. This section picks up the main aspects of grammatical complexity dealt with and studies their realization in the STSS utterance to show the role of the STSS utterance in defining these lectures along the spoken versus written language continuum with respect to this feature. These aspects are:

- the use of subordination versus co-ordination.
- the degree of subordination/ embeddedness in the STSS utterance.

- the use of finite versus non finite clause in the STSS utterance.
- the most frequent types of dependent clause.

5.6.3.1 Subordination versus co-ordination in the sub-topic shift signalling (STSS) utterance

The analysis of the frequency of occurrence of subordinate clauses as opposed to co-ordinate(1) clauses shows the following results: (the frequency of occurrence of each type of clause is calculated per 1,000 words in the STSS utterance).

Texts	Total No of words in the STSS utterance	F.I. For sub clauses	F.I. For co-ordinate clauses
Set 1 (3 lectures)	787	(37) 47.01	(4) 5.08
Set 2 (3 lectures)	762	(44) 57.74	(8) 10.49
Set 3 (3 lectures)	608	(52) 85.52	(5) 8.22
Total (9 lectures)	2157	(133) 61.65	(17) 7.88

Table (19) : Frequency index for subordinate versus co-ordinate clauses in the STSS utterance

There is a definite preference for subordination as opposed to co-ordination in the STSS utterance, in the three sets of lectures. These results are similar to those obtained in (3.3) with respect to the whole data, which are that biology lecture discourse favours the use of subordination over co-ordination. As pointed out then, this is typical of written language.

5.6.3.2. Degree of subordination in sub -topic shift signalling

Table (20) shows the results of the analysis of the frequency of occurrence of sentences distinguished on the basis of their level of complexity, calculated per 1,000 words in the STSS utterance.

Texts	F.I. For 0 degree complex S M or MCM	F.I. For 1 degree complex S M + Sub or M + E	F.I For 2 degree complex S M + Sub^1 Sub^2 or M + E^1 + E^2 or M + E^1 + Sub^2	F.I. For 3 degree complex S M + Sub^1 + Sub^2 + Sub^3 or M + E^1 + E^2 + E^3 or M + E^1 + Sub^2 + E^3	Total number of words in the STSS utterance
Set 1 (3 lectures)	(20) 24.41	(10) 12.70	(9) 11.43	(1) 1.27	787
Set 2 (3 lectures)	(19) 24.93	(20) 26.24	(6) 7.87	(1) 1.31	762
Set 3 (3 lectures)	(34) 55.92	(22) 36.18	(8) 13.15	(2) 3.28	608
Total (9 lectures)	(73) 33.84	(52) 24.10	(23) 10.66	(4) 1.85	2157

Table (20) : Frequency index for sentences distinguished on the basis of their level of complexity in the STSS utterance

Lectures in the whole data mostly use (1) degree of complexity sentences. The STSS utterance shows a comparable tendency since (0) degree complexity sentences are the most frequent. This feature would put biology lectures nearer to the spoken end of the continuum. The same conclusion was arrived at in the analysis of grammatical complexity of the whole data.

5.6.3.3 Finite versus non finite clauses in the STSS utterance

The analysis of the STSS utterance in terms of finite versus non finite clauses shows the following results:

Texts	Total No of words in the STSS utt	Frequency index For finite clauses	Frequency index For non finite clauses
Set 1 3 lectures	787	(95) 120.71	(13) 16.51
Set 2 3 lectures	762	(63) 82.67	(22) 28.87
Set 3 3 lectures	608	(67) 110.19	(15) 24.67
Total 9 lectures	2157	(225) 104.31	(50) 23.18

Table (21) : Frequency index for finite/non finite clause (per 1,000 words) the STSS utterance

Finite clauses are much more used than non finite clause in the STSS utterance. These results support again the results reported in (3.4) with respect to the use of finite versus non finite clause in the whole data. As stated in (3.4) this preferential usage of finite as opposed to non finite clauses is typical of spoken language.

5.6.3.4 Major types of dependent clause

The analysis of the clausal composition of the STSS utterance shows that the type of clause mostly found is the restrictive relative clause, with a frequency of occurrence of 19 per 1,000 words in the STSS utterance. It is followed by to-

infinitive clause (8.34) and that-clause (6.95). The distribution of the three types of embedded clause is presented in table (22):

Texts	Total No of words in the STSS utterance	F.I. for the restrictive relative clause	F.I. for to-clause	F.I. for that-clause
Set 1 (3 lectures)	787	15.24 (11)	3.81 (3)	6.35 (5)
Set 2 (3 lectures)	762	20.99 (14)	10.49 (8)	9.18 (7)
Set 3 (3 lectures)	608	21.38 (16)	11.51 (7)	4.93 (3)
Total (9 lectures)	2157	19 (41)	8.34 (18)	6.95 (15)

Table (22) : Frequency index for the major types of embedded clause in the STSS utterance

The frequency of occurrence of the three types of embedded clause is comparable in the three sets of lectures.

Excerpts (32) quoted earlier and (33) below exemplify these types of clause:

(32) //but perhaps the fern/ that will be most familiar to you in Britain is the bracken fern Pteridium//

(33262)

(33) //and it's usual /to remark/ that the blood concentrations are controlled by some form of negative feedback//

(129439)

In (32) the STSS utterance is grammatically realised by a sentence composed of a main clause and an embedded restrictive relative clause (underlined in the excerpt); In (33) it is realized by a main clause, a "to-infinitive clause" embedded into the main, and a that-clause embedded into to-clause (both are underlined in the excerpt).

If we compare these results with those obtained from the analysis of the whole corpus in Chapter Three, we can observe that the restrictive relative clause is again a prominent feature in these texts, in the STSS utterance. However, as pointed out in (3.6.1) the use of the restrictive relative clause is not an indicator of spokenness or writtenness in the lectures since conflicting views attribute this type of clause to both modalities.

The next two types of clause found in the STSS utterance are the non-finite to-clause and that-clause which are complement clauses. As discussed in (3.6.2). Complementizers are generally said to be typical of written language (Brown & Yule 1983 : 16). They contribute to the integrative nature of writing (Chafe 1982). The use of "to-infinitive clause" and "that-clause" is thus a feature of written language in the STSS utterance.

In summary, the analysis of grammatical complexity applied to the STSS utterance shows the following results:

(1) the STSS utterance is grammatically realised in structures readily analysable in sentence like units (see 3.1) (5.1.2).

(2) the STSS utterance uses more subordination than co-ordination

(3) the STSS utterance favours the use of a low degree of complexity, i.e. simpler sentences.

(4) the STSS utterance uses more finite than non finite clauses

(5) the analysis of the frequency of occurrence of the types of embedded finite clauses shows that the STSS utterance mostly uses the restrictive relative clause (not indicative with respect to speech and writing. See discussion above). The next two types of clauses (complement to-clause and that-clause) are typical of written language.

The results in (1) and (2) and (5) put the lectures nearer to the written end of the continuum, the results in 3, and 4, on the other hand, put them nearer to the spoken end of the continuum. As observed in (3.7), these types of results are not conclusive as to where biology lecture discourse belongs. They suggest that there are contradictory forces pulling the lectures towards both ends of the continuum.

We may conclude that the analysis of grammatical complexity whether applied to the whole data, or to sub-topic shift signalling, does not tell us much about where, on the continiuum, biology lecture discourse belongs.

5.7 Features of spontaneity in the sub-topic shift signalling (STSS) utterance

This section deals with some aspects of the STSS utterance which are not characteristic of the features of the textual and grammatical realisation of the STSS utterance as discussed so far. They are manifested in the syntactic or lexical revision of parts of the utterance; repetition of an item or structure; or a digression, interrupting the signalling utterance.

5.7.1 Syntactic revision

It is the syntactic reformulation of the sentence or part of the sentence constituent of the STSS utterance: in excerpt (2) quoted earlier:

> now the functions of the haemoglobin, {or let's put it another way.} the haemoglobin has two functions.
>
> (230339)

The signalling utterance starts with a NP which is then reformulated as a "have clause". The reformulation is preceded by the speaker's metatextual statement which specifies that a revised form of the same structure is to follow. The sub-

topic is slightly different in the two structures: it is further qualifed in the reformulation (two functions). It also puts NI in clause final position, a type of information structure favoured by the signalling utterance where new information comes at the end of the clause.

Reformulating part of an utterance is characteristic of a text in the actual process of being produced. It is a manifestation of the editing of the text produced in real time. Unlike writing where "words are laid out clinically on the page capable of being re-ordered, substituted, reflected upon " (Goody 1977 : 157), which, in other words is planned, speech is unplanned and unplannable (Ochs 1979). Reformulations as exemplified above are a manifestation of the "unplanned and unplannable" characteristics of speech. Their occurrence in the STSS utterance is thus a feature of spontaneous speech in the lecture.

5.7.2 Lexical revision

Lexical revision refers to cases where different lexical items are used in the reformulation:

(34) well now continuing this theme of. protecting the nitrogenese. I want to use I want to refer to er yet another mechanism which allows this... namely compartmentation...

(23027)

Where "to use" is reformulated as "to refer to". Here again the reformulation is a feature of sponteneous speech.

5.7.3 Reiteration of part of the signalling utterance:

It can have two manifestations. It can either be (i) reiteration of a lexical item or (ii) reiteration of part of the clause which contains NI.

(i) Reiteration of a lexical item

A lexical item may be repeated for emphasis:

(35) now we know a great deal. about. cycadella type of plants. these early seed ferns because miraculously. miraculously. representatives of them. have survived. to the present day.

(332938)

The word 'miraculously' occurs twice (prosodically it is double stressed). A different device would be used in written language such as underlining or italicising the appropriate item. Further, the use of double stress for emphasis is a medium specific feature since speech and writing are manifested through different media (the substance of speech is phonic, the substance of writing is graphic (see Berry 1975: 37). In (35) above, repeating the word "miraculously" and using stress to achieve emphasis is thus proper to speech.

(ii) Reiteration of most of the clause which contains NI in the STSS utterance

This is exemplified in:

(36) the evolution of the flower. is an extremely. complex. subject. and all we're going to be able to do. this morning. is. look briefly at some. major trends. in the evolution of the flower. we cannot look at the subject in detail. so let's look. at some major trends.

(333829)

Where most of the clause which contains new information (underlined) is repeated with the same lexical items. The reformulated clause is preceded by the contrastive clause "we cannot look at the subject in detail" which explains the selection of the

sub-topic. The sub-topic is then announced again. Here again, the reformulation is a manifestion of the unplanned nature of speech. It also occurs in a metatextual comment which uses the lexical item "this morning" (and thus is deictic) which refers to real time, and which realizes another feature of spoken language in the lectures.

5.7.4 Digression in the STSS utterance: use of a summary paraphrase interrupting the STSS utterance

The signalling utterance is interrupted by a summary paraphrase. In excerpt (6) quoted earlier:

> so now I'm going to look at.the.having looked at the root nodule, its structure, and one aspect of its structure, in relation to its functions, at er, what might be called, a molecular level. I now want to look at the. structure and function of. the root nodule. at a more. at a cellular. and anatomical level.
>
> (230349)

The speaker starts announcing the new sub-topic then interrputs his utterance before the first clause is completed to summarize the preceding sub-topic. The signalling utterance is then resumed. This interruption seems to be motivated by the need to use the main points of the preceding sub-topic as a context for the new one. The relation between the successive sub-topics is made explicit by the use of a summary paraphrase preceding the STSS utterance. A different situation is exemplified in:

> erm now let's see what's next... so you can see that, er. both, aporsity, and superfluity of, hormones can have quite, dramatic effects. now what I. want to do now. is tell you a little something. of the. chemical nature of hormones.
>
> (129111)

Here the use of the SP in the middle of the signalling utterance gives the speaker some time to decide and plan the next part of the lecture. It also allows him to keep the flow of speech going and avoid a gap of silence.

As discussed in (4.7.3.1) "digressions" can be found in a written text in the form of parenthetical remarks. However, there are three points which can be made in relation to the use of a digression (as exemplified in (6) and (1) above) which establish it here as a feature of spoken language:

(1) In (6) the digression disturbs the grammatical structure of the utterance. The first sentence in the utterance is interrupted and left incomplete (so now I'm going to look at the). It is reformulated after the digression.

(2) In (1) the digression is used as a strategy to plan the next part of the text.

(3) A summary paraphrase of a preceding part of text would not be expected to occur in the form of a parenthetical remark in a written academic text.

Therefore, the use of a digression as discussed above is typical of a spoken text.

5.8 Sub-topic shift signalling : a manifestation of the organising of the information structure of lectures

Sub-topic shift signalling is the explicit planning of the information structure of lectures. It indicates how the information is organised. Akinnaso (1982) in a review of the findings of studies of spoken and written language (see 2.4.4.) observes that:

> "written language promotes reliance on a more deliberate method of organising ideas, using such expository concepts as "..... topic sentence "....."
>
> (Akinnaso : 1982; 104)

The surface realisation of the structuring of the information is a feature of written language, (sub) topic (shift) signalling as a device is thus a feature of written language. We can further add that the frequent use of this signalling is also a feature of written language. It would be rare, that a participant in a spontaneous conversation announces what he is going to talk about and does so at frequent intervals in the exchange. Sub-topic shift signalling is thus typical of a planned text.

Further its use in the lectures can also be interpreted in terms of the pedagogical function of the lectures. As previously observed (see 4.3.2) the

purpose of lectures is to transmit information. The successful transfer of information will partly depend on how the information is structured and presented. As discussed in (2.1.2) studies in the comprehension of lectures (Chaudron & Richards 1986) with L2 learners have shown that the explicit planning of the information structure does facilitate students' comprehension of the lecture. The STSS utterance is a manifestation of such a planning. Its use helps students' processing of the text of a lecture. In other words its use promotes the transfer of information.

On the other hand, the status relationship between the speaker and the audience, which in a pedagogical situation is clear cut, also has an effect on the way the information is presented to the audience. As pointed out in (4.5.1) the speaker has the authority of transmitting knowledge to an audience of learners. He uses certain cues to facilitate their comprehension of the lecture. The signalling of the organisation of the content of the lecture in terms of sub-topic shift signalling is one of such cues. It is unlikely that, addressing an audience of peers, the speaker explicitly signals the structuring of the information he is presenting as he would do with an audience of learners. For example, it is unlikely that a type of STSS utterance such as real questions would appear in his text. So the use of the STSS utterance and aspects of its use can be seen to be motivated by pedogogical purposes.

5.9 Summary and conclusion

Features of spontaneous spoken language as opposed to planned written language in STSS can be grouped in (11) points:

1 - Use of interrogatives (real questions) and imperatives as types of change of structure signalling utterance.

2 - Frequent use of textual connectives (now, but) to initiate the signalling utterance.

3 - Use of place deictic terms to refer to the visual element.

4 - Occurrence of the personal pronoun "I" in uses typical of spoken language in the metatextual comment.

5 - Dominant use of (0) degree complexity sentences in the signalling utterance.

6 - Use of syntactic/lexical revision, repetition, interrupting structures.

7 - STSS utterance is generally realized in complete sentence like units.

8 - Prominent use of the statement (topic sentence) type of signalling utterance.

9 - Dominant use of subordination as opposed to co-ordination.

10 - Frequent use of complement clauses (non-finite to-clause and that-clause).

11 - STSS is used as a device to make explicit the planning of the information content of the text.

Features (1-6) are typical of spoken language; features (7-11) are typical of written language.

In conclusion, the findings of the whole analysis of sub-topic shift signalling in these lectures enables us to make the following points:

1 - The use of the STSS utterance is common to the three sets of lectures. Its frequency is comparable in the three sets. It is thus a feature of biology lecture discourse as a variety.

2 - The use of the STSS utterance as a device of organising the content of the lecture is a feature of planned written language. Its realisation in the text combines features of spoken and written language.

3 - Biology lecture discourse draws on features of spoken and written language for pedagogical purposes.

4 - The STSS utterance is a feature of the macrostructure of the text.

Footnotes

(1) See (3.3) for a discussion of the terms "subordinate" and "co-ordinate" as they are used in this analysis.

(2) See (3.2) for a discussion of the methodology applied in the analysis of the degree of complexity of each sentence in the texts.

CHAPTER SIX

PARAPHRASE. ITS TYPES AND TEXTUAL STRUCTURING

6.0 This chapter deals with paraphrase, the third textual feature investigated in this study. Paraphrase is first dealt with in terms of its types, its textual and grammatical features. Then some aspects of the use of paraphrase are discussed in relation to the spontaneous spoken versus planned written discourse distinction.

6.1. Paraphrase in the lectures

6.1.0 The analysis of the data reveals that some parts of the text of a lecture are reformulations of preceding stretches of text: for example:

(1) okay so that's just a brief survey. of. some of the glands in which the cells. which manufacture. there. discrete chemicals. are neatly grouped together. in discrete organs.

(128525)

This excerpt reformulates in a concise way a lengthy description of some of the glands which produce hormones (see appendix (1) :284-285). It does not say anything new. It repeats information which has been conveyed in the previous stretch of text. Being a brief reformulation, it omits certain facts and details which are found in the original stretch of text.

This is referred to as a summary paraphrase. Not all instances of reformulations found in the data are summary paraphrases, as excerpt (2a) below shows:

(2) ..... the amount of. sugar solution. well the amount of sap. which. passes through each cell. in the course of an hour. is equivalent to 100,000 times the volume of that cell.

(231830)

(2a) ..... then.. to repeat myself. it moves the sap that moves to a single cell in an hour is equivalent to 100,000 times it own volume..

(231833)

(2a) is a reformulation of (2). In this case the reformulation conveys a similar amount of information. It has the same scope in terms of information content as that of the original fragment of text. This is described as an equivalence paraphrase. The original utterance and the reformulation are equivalent in information content. Excerpt (3a) below exemplifies yet another type of paraphrase.

(3) and if we take a female cone. remember large. 90 pound structure. and remove a single cone scale. we see something like this. we see. a portion of the tip of the cone scale. which seems to represent. the reduction of a leaf.

(3a) so. the female cone scale probably evolved from a leaf like structure.

(333017)

Part (3) describes the physical structure of the female cone scale of a cycad. Part (3a) provides an interpretation of that structure in the context of a discussion of the evolution of the cycads. (3) and (3a) do not say the same thing. (3a) is not a reformulation of (3). The information conveyed in (3a) is implied in (3), or it can be inferred from (3). The utterance in (3a) is described as an inference paraphrase.

Paraphrase, as exemplified in excerpts (1), (2) and (3), refers to a phenomenon where the information content of an utterance has either been expressed earlier or is implicit in the information content of a previous fragment of text.

6.1.1 Scope of paraphrase in this study

Five different points must be made about the scope of the analysis of paraphrase in this study.

1 Paraphrase in this study is to be distinguished from paraphrase in logic

Paraphrase is looked at in ordinary language terms and not as it is used in logic where the relation between the paraphrase and the original utterance is established on the basis of the truth value of statements.

Thus, in the analysis of one type of paraphrase found in the lectures and referred to as inference pharaphrase, inference(1) is not used to refer to a strict logical relation which can be formalised. Rather, there is a case for an inference paraphrase if the paraphrastic form is a manifestation of the implication of the analysis of the semantic representation(2) of the original utterance.

2 Paraphrase is to be distinguished from expansion

In the following excerpt from the data:

(4) ... they are present in the. blood stream. normally. in very. low. concentrations..

(4a) steroid hormones for example. circulate. at concentrations. between 10^{-7}. and 10^{-9} moler. protein hormones are often present in even lower concentrations. 10^{-10} to 10^{-12} moler.

(129428)

(they in excerpt 4 refers to hormones).

Part (4) is an excerpt of a general statement about a characteristic of hormones which is that they are found in very low concentrations in the blood stream. Part (4a) gives examples of hormones and figures indicating their level of concentration in the blood. The statement is, here, specific. However, it is essentially different from instances of paraphrase described earlier in 6.1.0, in that there is additional information in (4a) which is that particular types of

hormones and their levels of concentration in the blood are specified. Because (4a) conveys information which is not found in (4), it is not considered as a case of paraphrase in this study. Thus 'paraphrase' is used here to apply only to cases where the reformulation or paraphrase form does not convey more information than contained in the original utterance. In this sense it is different from expansion which necessarily involves a gain of information (cf Longacre 1983).

3 Paraphrase may refer to more than one instance of the paraphrastic form of an utterance.

In the following excerpt:

(5) For example.. insulin from the pancreas. is normally only released. when glucose levels. in the blood go up after a meal.

(129242)

(5a) So what we've saying here. is that insulin. release. is only going to be provoked. when blood glucose levels. go. up.

(129247)

(5b) this. is the specific stimulus. for. its. release..

(129248)

The paraphrastic relation involves three utterances:

- (5a) is a reformulation of (5). Both (5) and (5a) describe two processes (insulin release and increase of glucose level in the blood which occur in sequence. (5a) can be described as an equivalence paraphrase of (5).
- (5b) states that the second process triggers the first one.
- (5b) can be inferred from either (5a) or (5) so it can be seen as the paraphrastic form of either (5a) or (5).

In this analysis an instance such as (5b) is discussed only in relation to the last preceding utterance it paraphrases because it is observed that the immediate co-text[3] of a paraphrastic form may play a role:

(i) either in the type of paraphrase used; in the above excerpt the inference paraphrase appears after the content of the original utterance is further explained and then reformulated in an equivalence paraphrase. The implication of the content expressed in (a), re-expressed in (b) is then explicitly stated (see appendix (1): 292)

(ii) or in the selection of equivalent terms used in the equivalence paraphrase. This can be seen in the following excerpt:

(6) we've just been discussing. a situation. in which a hormone interacts with the receptor at the cell surface. and that tends to be the mechanism. via which protein hormones exert their effects.

immediate co-text they are large molecules. too big. to get in the cells. into the cell normally

(6a) so they must exert their effect at the cell surface.

immediate co-text proteins are large. they are not readily lipid soluble

(6b) and so they exert their effect. at the level of plasma membrane...

(130041)

(6), (6a) and (6b) are all saying that protein hormones exert their effect at the cell surface.

The reformulation in (6b) uses plasma membrane as equivalent to "cell surface". The lexical analysis of (6b) reveals that the use of "plasma membrane" instead of repeating cell surface is not purely stylistic. The reformulation in (6b) follows a descriptive statement about the chemical property of hormones of not being readily lipid soluble. The chemical nature of plasma membrane has been dealt with earlier in the lecture (appendix (1) : 298-299) . It has been described as

having a lipid layer which forms an obstacle to hormone diffusion into the cells. "Plasma membrane" is more appropriate than "cell surface" in this reformulation as it provides further explanation as to why protein hormones do not go into the cell. The layer of lipid is an obstacle to hormone diffusion into the cell. The selection of "plasma membrane" in the paraphrastic form is thus co-text bound. This means that the sequential aspect of paraphrase is essential for its interpretation. Paraphrases are thus looked at in sequence.

4 Paraphrase does not apply across lectures

Inter-lecture reference through repetition or paraphrase (such as a reminder, at the beginning of a given lecture, of the main points dealt with in the previous lecture (see appendix (1):296) is not within the scope of this analysis. Paraphrase is only looked at within the same lecture. It is looked at as an intra-textual phenomenon (this point is further discussed in 6.1.2). The implication is that each lecture is looked at as a separate text. On the demarcation of one text from another, Ellis writes:

> "when we demarcate the flow of speech in its various manifestations into distinguishable occurrences of language, we do so partly because these occurrences have distinct situations....."

(Ellis 1976 : 91)

Language occurrences are distinguished into separate texts partly because they constitute different language events. So, "being a text" partly derives from the notion of "language event". Lectures are language events. Each lecture is looked at as being a separate language event. So each lecture has to be assumed to be a separate text(4) although the situational elements of audience, speaker and setting are held constant.

5 Paraphrase is looked at as essentially having a pedagogical function.

Longacre (1983: 115) remarks that paraphrase is used when "it is felt that saying something once is not enough". The communicator sees the need to repeat it or/and say it in different words, to ensure that the message is correctly perceived. It is expected that in a pedagogical situation such as the lecture where:

- firstly, as pointed in 5.8, the purpose of the activity is the transfer of information to an audience of learners;
- secondly, this pedogogical activity is carried out through the oral medium of language.

Paraphrase will be largely exploited. The purpose of the use of paraphrase is twofold. Firstly, the reformulation of an utterance is meant to clarify further the content of that utterance by using different words. Secondly, its recurrent use is a constant reminder of what has gone before in the text(5): Darian (1983 : 30) writes:

> "there is a limit to the complexity we can hold in the mind at one time. The speed of transmission would quickly overwhelm the listener in a highly detailed oral presentation. The only way to compensate is a level of redundancy totally unacceptable in writing.
>
> (ibid)

Repetition through paraphrase aims at facilitating students' comprehension of the lecture.

6.1.2 Defining paraphrase in this study

The description of instances of paraphrase presented in 6.1.0 allows us to adopt the following working definition; paraphrase applies where either:

(i) two or more utterances convey a similar meaning; the second and following utterances is/are saying or doing what the first utterance in the same text has said or done

or

(ii) the content of the second and following utterances can be inferred from that of the first utterance.

Four points can be made about this definition. Firstly, similarity in meaning will be qualified later in relation to the types of paraphrase it applies to in the data. Secondly, the relationship of inference is to be looked at as described in the previous section. There is a case for an inference paraphrase where an implicational relation can be seen between the first and subsequent paraphrastic forms. Thirdly, the paraphrastic form(s) occur(s) in the same text(4) as the original utterance. In this respect we are not concerned with the phenomenon of paraphrase inter-textually. It is possible to look at paraphrase across texts (see Urquhart 1982), but in this study paraphrase is treated as an intra-textual phenomenon. The implications of the occurrence of an utterance related by a paraphrastic relation to another in the same text is that it is to be read as a manifestation of continuity in the text (see De Beaugrande 1981). This continuity of the text allows us to see paraphrase as : (i) not adding new information to what the paraphrased utterance has said/done; (ii) having essentially a pedagogical purpose. Fourthly, paraphrase is being looked at as a function of the text. It applies to utterances. Its surface realisation is beyond consideration of sentence boundaries. This means that it is not treated, as in traditional linguistics, in terms of one to one correspondence between sentences (Falk 1978; Huddleston 1974; Postal 1970). Its scope applies beyond the level of the sentence: In the following excerpt:

(7) //well just as a deficiency of hormones can have quite profound effects so can excesses of hormones// and if growth hormone from that self same anterior pituitary region is produced in excess prior to puberty/ you get this enormous growth of the long bones and enormous increase in stature// and I think that this particular young man is 8 feet plus in height//

(7a) //so an excess of growth hormone prior to puberty will cause giantism//

(12906)

If the two utterances (7 and 7a) are analysed into sentence like units, the paraphrase (7a), which in this excerpt is a summary paraphrase, is realised in one sentence; the original utterance (7), into three sentences. Thus one sentence summarises the content of three sentences.

6.2 The summary paraphrase (SP)

6.2.0 This section deals with one type of paraphrase, the summary paraphrase (SP). It will first discuss the criteria used to identify the SP. Two types of SP found in the data are analysed in terms of their features of textual scope and their position in the macrostructure of the text. Ellipsis, a grammatical feature common to both types of SP, is then discussed.

As described in 6.1.0 the SP is a concise or abbreviated reformulation of the preceding chunk of text. It always involves a loss of information compared to the chunk of text it summarises. In the context of related studies in the existing literature on paraphrase and notions relating to paraphrase, the SP is comparable to the reduction of a text as described in Gopnik (1972).

6.2.1 Criteria for identifying the summary paraphrase (SP)

Two types of criteria are used to identify the summary paraphrase.

6.2.1.1 Semantic criterion

The SP must not contain any new information. It does not involve any elaboration or expansion on the content of the stretch of text which is summarised.

As stated above, it rather involves a loss of information. This can be seen in the following instance in the text:

(8) so erm. that's the ground work then for the structure of the vascular. bundle. you realise that the whole of this space is

actually filled with. tissues. of the. xylem. the phloem. and the pericycle. containing these particular. cell types..

(230929)

which summarises a lengthy discussion about the structure of the vascular bundle (see appendix (1): 307-309).

This semantic criterion is a sufficient condition to identify a SP.

6.2.1.2 Textual criteria

Two textual criteria may be used to identify the summary paraphrase.

(i) A metatextual expression is used to signal that what follows is a summary of what has gone before in the text: examples:

(9) so in summary. too little growth. hormone will give you. the achonodroplastic dwarfism. plus or minus, the lack of development. of the genitalia.. too much before puberty gives giantism. too much after puberty. gives you. accromegally...

(129038)

(10) so that's a rather long winded way. of simply pointing out to you. that this internal communication. is a joint. or collaborative responsibility. between nerves on the one hand. and chemical messengers we referred to as hormones. upon the other...

(128413)

In excerpt (9) the metatextual expression "so in summary" (underlined) identifies the utterance it initiates as an SP. The key word in the metatextual expression is "summary". In (10) the metatextual expression is "so that's a rather long winded way of simply pointing out to you" "that" (underlined in the excerpt). The deictic term "that" carries out the reference to the preceding chunk of text being summarised. The rest of the metatextual expression is a comment on that chunk of text. It implies that what is to follow is a briefer reformulation.

This textual criterion is also a sufficient condition to identify an SP.

(ii) The textual connective "so" initiates the utterance. As pointed out in 4.7.1. in this use, "so" is essentially deictic. It indicates that the utterance being introduced is a paraphrase.

This criterion applies in combination with either the textual criterion discussed in (i) or with the semantic criterion.

6.2.2 Pseudo summary paraphrase

There are rare cases of a type of utterance which is near to the topic sentence. It does not reformulate the content of a stretch of text, but rather it comments on the content of that stretch of text. In (11) for example:

(11) so that er little piece of work. by Bartol. is the great classic of. endocrinology...

(12873)

What Bartol's work is about is not repeated. It is commented on, it is qualified as "the great classic of endocrinology". This utterance does not summarize the content of the preceding fragment of text which discusses the details of experiments carried out on cockerels which enabled the scientist Bartol to discover the existence of hormones. Therefore it is not considered as a summary paraphrase.

6.2.3 Textual features of the summary paraphrase (SP)

Two types of features are looked at: the summary paraphrase as a marker of a boundary in the text and the use of initial deictic "so".

6.2.3.1 The summary paraphrase (SP) in relation to the macrostructure of the text

As mentioned earlier in (5.4.2.2), the SP has two different realisations according to the scope of its content. Consequently it is distinguished into two types : the general summary paraphrase (GSP) and the partial summary paraphrase (PSP).

6.2.3.1.1 General versus partial summary paraphrase in the lectures

Consider excerpts (8) (quoted earlier) and (12) below:

(8) so erm. that's the ground work then for the structure of the vascular. bundle. you realise that the whole of this space is actually filled with. tissues. of the. xylem. the phloem. and the pericycle. containing these particular. cell types..

(230229)

(12) er so that's what's shown in a transverse section. the characteristic. perforation of the end.. wall with the cytoplasmic strand coming through here. connecting the cytoplasm of one sieve tube element. with those above and below..

(230813)

In (8) the paraphrase summarises a general sub-topic (see 5.4.2.1) about a description of the structure of the vascular tissue. This general sub-topic is further broken down into partial sub-topics. The summary paraphrase captures in its scope all the sub-topics which make up the general sub-topic. It is referred to as a general summary paraphrase (GSP).

Excerpt (12) summarises the description of the structure of a constituent of the vascular tissue namely the sieve tube element found in the phloem tissue. The phloem tissue being composed of a type of cells constituting the endodermis which in the structure of the vascular bundle (a longitudinal strand in the vascular tissue) is a layer of cells which surrounds the vascular bundle. This type of SP is referred to as a partial summary paraphrase (PSP).

The GSP has then a wider scope than the PSP which always summarises a partial sub-topic which is part of a larger sub-topic.

The two types of summary paraphrase show the following distribution in the data.

Texts	No of occurrences of SPs in the data	F.O. of GSP	F..O. of PSP
Set 1 3 lectures	(61)	(21) 37.7%	(40) 65.57%
Set 2 3 lectures	(19)	(8) 42.1%	(11) 57.89%
Set 3 3 lectures	(17)	(7) 41.17%	(10) 58.82%
Total 9 lectures	(97)	(36) 37.11%	(61) 62.88%

Table (23) : Frequency of occurrence of general summary paraphrase and partial summary paraphrase

The results in Table (23) show that the partial summary paraphrase is more frequent than the general summary paraphrase in the three sets of lectures (the frequency of occurrence has been calculated per 100 occurrences of summary paraphrase). This distribution is logical since the GSP applies to a larger chunk of text than the PSP.

6.2.3.1.2 The summary paraphrase as a marker of a boundary in the macrostructure of the text

The summary paraphrase generally occurs at sub-topic boundaries i.e. it is followed by a new sub-topic introduced by a sub-topic shift signalling utterance . This can be seen in the following excerpts:

(13) we've seen then. that. it is possible. perhaps to. derive the stamen. consisting of an anther and a filament. from a leaf.

(333551)

(14) but how. can we. derive the. ovary. the style and the stigma. from leaves..

(333553)

Excerpt (13) is a brief reformulation of the preceding discussion of a partial sub-topic about how the stamen can be seen as structurally related to the flower. Excerpt (14) is an instance of the signalling of a sub-topic shift. It immediately follows (13) which is a PSP. Excerpt (14) is also an instance of partial sub-topic shift signalling utterance. The SP, then, closes a sub-topic.

As stated in 5.4.2.1 a sub-topic is textually realised in a sequence. This sequence is a major sequence if it textualises a general sub-topic. A major sequence often closes with a GSP, and a minor sequence, with a PSP (see fig. (2)). The SP, like the sub-topic shift signalling utterance, is a marker of a boundary at the level of the macrostructure of the text.

6.2.3.2 Use of initial "so", a feature of spoken language in the summary paraphrase (SP)

The use of deictic "so" as initiator of the summary paraphrase is a major feature of this type of paraphrase. Most SPs are introduced by "so": they are 98.36% of the total number of SPs in set (1); 89.74% in set (2) and 66.66% in set (3).

As discussed in (4.7.1) the recurrent use of "so" as paraphrase initiator is seen as a feature of spoken language. (see also Darian 1983; Farag 1986). This point will be taken up again in 6.7.2.2 in this Chapter.

6.2.4 Ellipsis in the summary paraphrase (SP)

This section deals with a prominent grammatical feature of the SP: namely the use of ellipsis.

Ellipsis is used here to refer to cases where the ellipsed constituents can be recovered from the grammatical structure of the utterance, i.e. from the grammatical structure of the summary paraphrase (Quirk et al 1972: 536, 1985: 861).

6.2.4.1. Cases of ellipted constituent in the summary paraphrase (SP)

The forms which tend to be omitted are certain uses of the verb "be".

(i) existential 'be' + article

The form "there is/are + article" is frequently ellipted in the SP as the following excerpt shows:

(15) so it looks as if in the achonodroplastic dwarf ( ) specific absence of growth hormone

(128926)

(16) so ( ) two. very different types. of seed plant. in the end of the permian period.

(332936)

(17) so ( ) a whole host. of different materials. produced. by the pituitary.

(12895)

In (15) the ellipted structure "there is a" would, if stated, precede the nominal phrase "specific absence of growth hormone"; in (16) "there are" is omitted and would occur after initial "so"; in (17) "there is" would also occur after initial "so".

In the three cases, the surface structure used is a verbless clause.

(ii) equative be + article

Example:

(18) so gene activation. ( ) typical. mechanism underlying. the action of steroid hormones.

(130129)

"be + article" is dropped preceding "typical". The resulting clause is a verbless main clause.

6.2.4.2 Ellipsis as a feature of spoken language in the summary paraphrase (SP)

The deletion of certain constituents in the summary paraphrase results in an abbreviated surface form of the utterance which reflects the nature of the summary paraphrase as being a brief reformulation of a preceding stretch of text.

It has been observed that speech shows a tendency towards achieving a certain degree of "linguistic economy by the use of ellipsis .... for transmitting information". (Darian 1983 : 38). In the summary paraphrase, the use of ellipsis seems to be motivated by two main factors:

- firstly, as observed above, the utterance where it is used is a brief recapitulation of a preceding part of the text. It is repeating information. It

tends to use only content words and omit grammatical constituents such as "existential be" as seen in excerpts (15) to (17) above.

- secondly, the implication of the preceding observation is that the semantic interpretation of the utterance is not affected by the deletion of those constituents.

As observed above, ellipsis is claimed to be typical of spoken language; however it is also found in written language. This is particularly the case of the type of ellipsis discussed above, which depends on recoverability from the linguistic context(6). However, planned written language would not allow for ellipsis of constituents which affect the grammatical well formedness of the sentence, as can be the case in spontaneous spoken language and as is observed in excerpts quoted above. So the utterance quoted earlier in excerpt (18)

> so gene activation. typical. mechanism. underlying. the action of steroid hormones.

is grammatically incomplete and would not be expected in a written academic text. It is thus a feature of spoken language in the lectures. Further, the effect of the use of ellipsis is a certain degree of fragmentariness in the surface realisation of the summary paraphrase (as excerpt (18) above exemplifies). Fragmentariness is described as being typical of spoken language (see 2.4.2) (see Chafe 1982; Tannen 1984; Johns Lewis 1987).

6.3. The equivalence paraphrase (EP)

6.3.0 This section deals with the second type of paraphrase found in the data, which is referred to as the equivalence paraphrase. It will first discuss the types of criteria set up to identify the equivalence paraphrase. Lexico-grammatical differences between the original utterance and the paraphrastic form are then discussed, with reference to particular examples. Finally, textual features of this type of paraphrase are dealt with.

As stated earlier (6.1.0) in this type of paraphrase the information content of the paraphrastic form and that of the original utterance have a similar scope. Unlike the summary paraphrase, the equivalence paraphrase says in other words what the original utterance has said with no or very little loss of information content.

6.3.1 Criteria for identifying the equivalence paraphrase

6.3.1.1 Criteria

Six linguistic criteria have been used to establish a case of an equivalence paraphrase: five of them are semantic, one is syntactic.

(i) Semantic criteria: there is a case of equivalence paraphrase if the reformulation or paraphrastic form:

(1) uses the technical/common terms which have appeared in the original utterance.

(2) uses lexical items (technical or common) which are recognised as synonyms to items used in the original utterance.

(3) uses lexical items or an expression(s) (more than one item) equivalent in meaning to items or to (an) expression(s) used in the original utterance.

(4) uses lexical items (technical or common) which have identical referents as items used in the original utterance.

(5) is a definition or an explanation of (1) technical term(s) used in the original utterance.

(6) the paraphrastic form does not say anything new.

(ii) syntactic criterion: the reformulation uses the same or equivalent constituents in different syntactic roles.

(iii) additional criterion: specialised informants (authors of the texts) have been consulted to verify that instances covered by these criteria in particular the most doubtful instances, are cases of equivalence paraphrase.

Any of the semantic criteria in combination with (6) is a sufficient condition for a reformulation to qualify as an equivalence paraphrase of a previous utterance.

Any of the semantic criteria and the syntactic criterion can combine in an instance of equivalence paraphrase.

The syntactic criterion has to combine with one (or more of the semantic criteria).

The 'additional criterion" has been used with some of them.

Three points can be made about these criteria:

Firstly, criterion (i).4 establishes that the equivalence between lexical terms (common or technical) in the reformulation and in the original utterance is handled in terms of "identity of contextual referents" (see Urquhart 1982: 75). In this analysis, the referent is either the visual display by means of slides and diagrams of the phenomenon being described, or the phenomenon when visual aids are not used. Excerpt (19) below satisfies this criterion:

(19) if we take this structure, and fold it round, so that the ends meet.

(19a) that is we take these.. two edges. fold the thing round so that these two ends fuse together.

(33364)

(19) uses the constituents "this structure", "the ends", to refer to the diagrammatic representation of a carpel (the female reproductive organ in the flowering plant). (19a) uses the constituents "the thing" instead of "this structure" and "two edges" instead of "two ends" then "these two ends" again to refer to the diagram or to the same parts of the diagram. The equivalence between these constituents can thus be established on the basis that they have identical referents.

Secondly, the syntactic criterion applies to cases where the relationship between the utterance and its equivalence paraphrase is partly syntactic (the syntactic criterion, as discussed above, applies in combination with one (or more) semantic criteria. This relationship has different manifestations:

(i) it can involve a reordering of the constituents:

Example:

(20) and thirdly. it may. inhibit .other plants. by exuding toxins. from the fronds..

(20a) so. by exuding toxins. this may kill. the competing plants. directly..

(332632)

In this excerpt the paraphrase in (20a) shows a reordering of the clauses which make up the sentence.

(20)	1 - and thirdly it may inhibit other plants	is equivalent to	(20a)	This may kill the competing plants directly
(20)	2 - by exuding toxins from the fronds	is equivalent to	(20a)	So by exuding toxins

The equivalence between the two utterances is based on the use of equivalent constituents as shown above. Except that "competing" appears in the reformulation

only. "Competing" or its equivalent is not found in the original utterance but we can say that it does not carry new information as it can be recovered from the preceding context which is about how "bracken fronds", a type of plant, eliminates other plants and invades spaces killing all plants that are in its way (see appendix (1) : 326).

(ii) it can involve (though rarely) the ellipsis of some of the constituents as the following excerpt shows:

(21) each megasporangium or ovule. is located on the upper side. of. an ovuliferous scale. protected by. a bract scale..

(333217)

(21a) megasporangia then. on the upper side. of. ovuliferous scales protected by bract scales...

(333222)

The reformulation uses the same technical terms (underlined in the excerpt), but these are used in the plural form. Consequently, the determiner preceding each one of them in the original utterance is dropped in the reformulation. These determiners are: "each", "an", "a", the other forms which are also ellipted in the reformulation are: "or ovule" occurring after "megasporangium", and 'is located' preceding "on the upper side". Ellipsis, in this case, applies to grammatical (each, an, a, is, on the) and to content words (located, upper side). However,the omission of these content words does not affect the meaning of the utterance.

Thirdly, the authors of the text (as specialist informants) have been consulted for their specialised knowledge of the subject to have their reaction about the analysis of certain utterances as equivalents in information content on the basis of the semantic and syntactic criteria set up above.

6.3.1.2 Typical examples

Some examples are used to show different combinations of the criteria discussed above:

(22) but this. hypothalamic zone. is also. in contact. via nerves. with. a gland. known. as. the pituitary gland.

(22a) and from the hypothalamus. emanate nerves. connecting it. to. a little gland. at the base of the brain called the pituitary.

(128333)

The reformulation uses technical terms which are either equivalent to technical terms used in the original utterance (hypothalamic zone/hypothalamus) or the same as those used in the first utterance: (nerves, gland, pituitary). This aspect of the equivalence between (22) and (22a) in this excerpt combine criteria (i) 1, 2 and 4. The equivalence between (22) and (22a) can be further spelt out by comparing the following equivalent segments from each part.

(22)	"is also in contact via nerves with"	is equivalent to	(22a)	"nerves connecting it to"
(22)	"known as"	is equivalent to	(22a)	"called"

the equivalence between the members of each pair is established by applying criterion (i) 3.

Further, the syntactic structure of the utterance in (22) and (22a) is different. In (22), "hypothalamic zone" is NP subject in a be construction. In (22a), on the other hand, the technical term "hypothalamus (equivalent to "hypothalamic zone") occurs in a prepositional group which is foregrounded in the structure of the

utterance. This marked foregrounding of the prepositional group and putting the (actor) subject after the verb has the function of emphasising movement and directionality in that: (1) it emphasises where the nerves originate, (2) they (as subject/actor) do the process of emanating and connecting.

In the next excerpt, only criteria (i) 3 and (iii) apply:

(23) the ground underneath the tree is always wet

(231914)

(23a) the ground underneath. the tree is. as I say continuously. damp

(231918)

The reformulation uses the same word order, the same clause constituents, but different lexical items which are roughly equivalent to the items they substitute (underlined in the excerpt):

(a)	"always"	is equivalent to	(b)	"continuously"
	"wet"	is equivalent to		"damp"

The last excerpt discussed in this section shows the application of criteria (i) 5 and (iii).

(24) angiosperms many of them. tend to have. what is regarded as being. parallel. venation.

(24a) they tend to have a series of veins which run. roughly parallel throughout the length of a leaf.

(333430)

the reformulation explains the meaning of the technical expression "parallel venation", using common words (underlined in the excerpt).

6.3.2 Types of equivalence paraphrase

On the basis of the criteria set up in 6.3.1.1, the equivalence paraphrase is distinguished in two types:

(i) the complete equivalence paraphrase

(ii) the partial equivalence paraphrase

Excerpt (23) and (24) quoted above in 6.3.1.2. are instance of (i); in each excerpt the reformulation in (23a) and (24a) applies to the whole utterance in (23) and (24). (In 23a, the comment 'as I say' is not part of the content of the utterance).

Excerpt (25) below is an instance of (ii):

(25) so when plasma calcium concentration goes up, not only does PTH. come down. but calcitonin levels. go up in the blood. calcitonin has its target. in the bone and where it. inhibits the dissolution of bone mineral.

(25a) so calcium goes up, cells in the thyroid release calcitonin, reacts on the bone, inhibits, the dissolution of bone mineral.

(129527)

The equivalent parts can be shown in pairs as follows:

(25)	so when plasma calcium concentration goes up	is equivalent to	(25a)	so calcium goes up
	calcitonin has its target in the bone			calcitonin reacts on the bone
	where it inhibits the dissolution of bone mineral			and inhibits the dissolution of bone mineral

The reformulation only applies to the above quoted segments. We can observe that part of the utterance in (25a) which is "cells in the thyroid release "calcitonin" is not

expressed in (a). However this is not additional information as it can be recovered from the context immediately preceding the first utterance in excerpt (24):

> there is another hormone which comes in fact from the thyroid gland which is called calcitonin
>
> (129526)

Excerpt (25) is an instance of a partial equivalence paraphrase.

Excerpts (23), (24) and (25) show that the equivalence paraphrase can either be complete or partial. It is complete if the whole of the second utterance is the paraphrastic form of the totality of the first utterance, the equivalence is partial if the equivalence only partly applies.

6.3.3 Motivations for lexical and syntactic differences between the utterance and its reformulation in the equivalence paraphrase

6.3.3.0 This section tries to identify reasons for using equivalent words or alternative syntactic structure, in the paraphrastic form of a previous utterance.

6.3.3.1 Motivations for lexical differences in the equivalence paraphrase (EP)

Lexical difference refers to the use of an alternative word or expression (technical or common) in the paraphrase equivalent in meaning to a word or expression (technical or common) used in the original utterance.

If we look again at excerpt (23) (quoted in 6.3.1.2)

(23) the ground underneath the tree is always wet

(23a) the ground underneath the tree is erm is is as I say continuously damp

the lexical differences between (23) and (23a) in the excerpt is the use of near synonyms 'continuously' and 'damp' as substitute for "always" and 'wet'.

The reason which seems to suggest itself for using these equivalent words in the paraphrase is to avoid pure repetition of the utterance. The use of the comment "as I say" preceding that part of the utterance where equivalent words are used, can be seen as a strategy used by the speaker not to interrupt the flow of speech while looking for the right formulation. It can also be seen to be used to signal that a "repetition" is to follow. Therefore it also realizes a metatextual function.

Different reasons seem to underlie the use of an equivalent technical expression in the following excerpt (6) (quoted in 6.1.1)

(6) we've just been discussing. a situation. in which a hormone interacts with the receptor at the cell surface. and that. tends to be the mechanism. via which. protein hormones exert their effects. they are larger molecules. too big. to get in the cells. into the cells normally

(6a) so they must exert their effect at the cell surface

(6b) and so they exert their effect. at the level of plasma membrane...

In this instance, the lexical difference between (6a) and (6b) is the alternative use of the expression or compound "plasma membrane" as a substitute for "cell surface." (The equivalence between "plasma membrane" and "cell surface" is established on the basis that they have the same referent, i.e. a particular part of the cell).

As previously discussed (6.1.1) the reformulation in (b) comes after further explanation has been given as to why protein hormones do not go into the cell, namely that they are not readily lipid soluble. The chemical property of 'plasma membrane' of having a lipid layer has been discussed previously in the lecture (see appendix (1) : 298-299). So the occurrence of 'plasma membrane' in the paraphrastic form highlights the property of "plasma membrane" of having a lipid layer which inhibits protein hormone moving into the cell. So the selection of

'plasma membrane' as equivalent to 'cell surface' is determined by the co-text of the paraphrastic form.

6.3.3.2 Motivations for syntactic differences in the equivalence paraphrase

The syntactic difference between the utterance and its paraphrastic form is often the syntactic reordering of the structure of the utterance. This can be seen in the following excerpt:

(26) and since. specificity is at the heart. of this recognition process.. you might. deduce. or suspect. that. the recognition mechanism has got something to do with the shapes. of the molecules.. and therefore. the receptors. should be endowed. with chemical properties. which would. enable them. to discriminate. on the basis. of the shapes of. the hormones. and the best molecules. for taking up specific configurations in space. with an appropriate orientation. are proteins..

(26a) the proteins are extremely adept. at forming. specific configurations in space. into which. a specific hormone. shape..... (129832)

Part (a) is a long explanation of the recognition mechanism which is said to take place between the hormone and the cell. It states that this recognition mechanism essentially depends on the shape of the hormones. The discussion ends with the statement that proteins are the best type of molecules for having appropriate shapes to interact with a hormone. In (26a) the statement about proteins is reformulated. The paraphrastic form (26a) begins with the lexical (technical) term 'proteins' which in (26) comes last, at the end of a long explanation. The reformulation follows a certain progression. The new information in the immediately preceding utterance is used as theme (i.e. starting point) in the reformulation. The reformulation however does not include new information in the rheme (i.e. the predicate). What we notice is that the reordering has essentially the pedagogical purpose of explaining the phenomenon from a different angle.

6.3.4 Textual features of the equivalence paraphrase (EP)

6.3.4.0 Three textual features of the equivalence paraphrase as looked at, they are:

- the distribution of the EP in the structure of the text
- the distribution of the EP in relation to the original utterance
- the relationship between the EP and the preceding co-text

6.3.4.1 The equivalence paraphrase (EP) is a feature of the sequence

The EP (and the original utterance) are generally found within the boundaries of a sequence which is formally marked by the sub-topic shift signalling utterance (see 5.4.2.1) and the summary paraphrase (see 6.2). The EP tends to be found in the main part or towards the end of the sequence, preceding the summary paraphrase. For example, excerpt (6) quoted earlier is found within a minor sequence which begins with the partial sub-topic shift signalling utterance:

> a second mechanism might be via gene activation that means switching on of specific portion of DNA
>
> (130039)

and ends with the PSP:

> so gene activation. typical. mechanism. underlying. the action of steroid hormones.
>
> (130129)

The sub-topic in this sequence is the discussion of 'gene activation', as being one way in which hormones exert their effect inside the cell, with specific reference to the action of steroid hormones. The EP occurs early in the discussion. It is a feature of the internal structure of the sequence, or a feature of the development of the sequence (cf Longacre 1983 : 122).

6.3.4.2 Distribution of the equivalence paraphrase (EP) in relation to the original utterance

The original utterance and its paraphrastic form are often not immediately sequential in a running stretch of text. The formulation does not occur immediately following the part of the text it paraphrases.

This aspect of the distribution of the EP can have an effect on the formulation of the paraphrastic form. As discussed in 6.1.1 and 6.3.3.1, the choice of (an) equivalent lexical item(s) or expression(s) can be motivated by the co-text preceding the paraphrastic form. (see 6.3.3.1 and discussion in relation to excerpt (6)).

The relationship between the EP and the preceding co-text can have yet another manifestation. This is discussed next.

6.3.4.3 Dependence on co-text

There are cases of EP where the paraphrastic relation can only be established with reference to the co-text which precedes the two utterances i.e. the original utterance and its paraphrastic form. This aspect of paraphrase has been discussed in Urquhart (1982:84), where a distinction is made between co-text free and co-text bound paraphrases. Both types of paraphrases are found in this data: some EP's are co-text free as the following excerpt shows:

(27) there was a gradual reduction. in the number. of. parts. in the flower.

(27a) the most primitive flowers. tended to have. large numbers. of different parts. and the most advanced flowers. tend to have relatively few. parts..

(333837)

The equivalence of the two utterances can be established without resort to the co-text: the paraphrastic form explains the content of the original utterance which is about an aspect of the evolution of the flower.

However most EP's tend to be co-text bound as can be seen in the following excerpt:

(28) but in some way. that. process. has ultimately to be switched off again.

(129732)

(28a) there must be a termination. to. the. hormone effect.

(129734)

(28) follows a discussion of the action of hormones in metabolic processes (see appendix (1) : 297) . "That process" in (28) refers to hormone action or effect on a metabolic process. The equivalence of (28) and (28a) depends on the co-referentiality of "that process" and "hormone effect" which can only be recovered from the preceding co-text. So the paraphrastic relation between (28) and (28a) depends on the preceding co-text.

A similar case of EP can be seen in:

(29) this is a very big advantage for a plant in the tropical forest. to be able to do.

(29a) so if you can climb around your neighbours. then. that. is a big advantage.

(332541)

The paraphrastic relation can only be established if the referent of the anaphoric pronoun "this" is recovered from the preceding co-text:

Lygodium. can. climb up other plants. in the tropical forest.

(332540)

The referent of "this" is "the characteristic of a plant (such as Lygodium) to climb up other plants in the tropical forest". The equivalence between (29) and (29a) can then be seen as represented in the following pairs:

(29)	this (which refers to the characteristic of a plant to climb up other plants in the tropical forest)	is equivalent to	(29a)	"if you can climb around your neighbours
(29)	"is a very big advantage"	is repeated in	(29a)	then that is a big advantage

This dependence on co-text can be looked at as a manifestation of textual continuity. This point will be discussed later in 6.5.

6.4 The inference paraphrase

The present section deals with the last type of paraphrase looked at in the data, referred to as the inference paraphrase. The criteria set up to identify the inference paraphrase are discussed first with typical examples, then the textual features of this type of paraphrase are dealt with.

In this type of paraphrase, the paraphrastic form is analysed as a possible inference(7) of the information content of the original utterance. It does not reformulate the information content of the original utterance. Rather, the information content of the paraphrastic form and that of the original utterance are related by means of an implicational relation.

6.4.1 Criteria for identifying the inference paraphrase

Three linguistic criteria are used to identify the inference paraphrase:

(i) Semantic criteria

There is a case for an inference paraphrase if the information content of the paraphrastic form can be inferred from the content of the original utterance. It must be implicit in the content of the original utterance;

- the paraphrastic form does not say anything new

(ii) Textual criteria

- the paraphrastic form is initiated by the textual connective "so" or "therefore"

(iii) Additional criteria

Specialized informants (authors of the text) have been consulted about doubtful cases of inference paraphrase.

Criteria (i) and (ii) are both necessary conditions for identifying a case of inference paraphrase. Criteria (iii) has been used for some instances of inference paraphrase.

The following two examples show how criteria (i) and (ii) apply:

(30) only. when there is an increase in follicle stimulating hormone. from the anterior pituitary. will. the body increase its oestrogen production. from. the ovaries.

(30a) so the specific stimulus in this instance. is follicle stimulating hormone.

(129251)

While (30) describes two sequential processes, (30a) states that the second process is the cause of the first. Thus we can say that (30) and (30a) do not say the same thing: (30a) is implied by the meaning of (30), i.e. it can be inferred from (30). Because it is implied by the meaning of (a), it does not say anything new.

(30a) is introduced by the textual connective "so" which in this function is:

(i) deictic as it signals the paraphrase, (this was pointed out in 4.7.1).

(ii) a marker of the implicational relation between the utterance and its paraphrastic form

In this next excerpt:

(31) it's known for example. that if animals eat bracken fronds. which they may do at times of drought. it kills them.

(31a) so there are definite toxins. and poisons. in. the bracken fronds.

(332633)

(31) says that animals die when they eat bracken fronds, (31a) states explicitly that bracken fronds have poisons and toxins. (31a) is implicit in (31). There is thus an implicational relation between (31) and (31a) which is marked by the use of "so" which, as discussed above, is also deictic.

As stated in 6.4.0, the inference paraphrase does not reformulate the content of the original utterance as it is found in the previous type of paraphrase i.e. the equivalence paraphrase. If we consider again a typical instance of equivalence paraphrase quoted earlier (in excerpt 21).

(a) each megasporangium or ovule. is located on the upper side. of. an ovuliferous scale. protected by.a bract scale..

(b) megasporangia then. on the upper side. of. ovuliferous scales protected by bract scales...

(333217)

Here the paraphrastic form (b) reformulates the content of the first utterance (a), using the same technical terms (underlined) in a modified syntactic structure: the constituents "each", "an", "a" or ovules, "located" are dropped in the reformulation. (b) is not an inference of (a). It repeats the context expressed in (a) in a different syntactic structure. The inference paraphrase, on the other hand, does not repeat the content of the original utterance. Unlike the equivalence paraphrase, the inference paraphrase does not depend on lexical or syntactic relationships

between the first utterance and its paraphrastic form. In other words, it does not depend on the occurrence of equivalent lexical constituents or reoccurrence of the same lexical constituents in an identical or different syntactic structure.

6.4.2 Textual features of the inference paraphrase (IP)

6.4.2.0 Two textual features of the inference paraphrase are looked at; they are:

- the distribution of the inference paraphrase and the original utterance in the structure of the text.
- the relationship between the inference paraphrase and the preceding co-text

6.4.2.1 The inference paraphrase (IP) is a feature of the sequence

The inference paraphrase and the original utterance share the distributional feature of the equivalence paraphrase and the utterance it paraphrases in that it is generally found within a sequence. For example, excerpt (30) above, occurs in a sequence which opens with the following partial sub topic shift signalling utterance:

> now since. we've seen. that they have. such dramatic effects if they're present in excess or. in less than normal quantities. it seems. fairly obvious. that. their concentrations in the blood stream or the plasma. should only be varied. in response. to. specific requirements.. and indeed. all hormones.. are only released.. in. greater. than normal amounts. in response to very. highly. specific. stimuli..
>
> (129236)

This announces that the next sub topic is about "the release of hormones in response to specific stimuli." The sequence ends with the partial summary paraphrase:

> so each of the hormones. has its own. specific. trigger..
>
> (129254)

The inference paraphrase is thus a feature of the structure of the sequence; like the equivalence paraphrase it is a feature of the development of the sequence (cf Longacre 1983: 122). On the other hand, its distribution in the sequence in relation to the original utterance is that it always immediately follows it. In other words The original utterance and its paraphrastic form are always sequentially simultaneous.

6.4.2.2 Dependence on co-text

As discussed in the analysis of the equivalence paraphrase, some instances of inference paraphrase are co-text dependent, i.e. the paraphrastic relation can only be established with reference to the co-text of the members of the paraphrastic relation. This can be seen in the following pairs:

(32) it proved by electro-microscopy it proved by. examination. that there are cytoplasmic connections between the infected cells. and the uninfected cells of the cortex. there are cytoplasmic connections across the endodermis. and into the pericycle.

(32a) so there is nothing to impede. the flow of organic nitrogen from the infected cells. to the pericycle. in the simplast.

(231131)

(33) unfortunately you've got to be very careful. about the use of both of these compounds. they are very toxic. and there are many other organisms as well. which may suffer. from their application.

(33a) so you have to chance. getting rid of the bracken. with the possibility. of side effect.

(332644)

The excerpt in (32) follows a discussion about the circulation of organic nitrogen from infected cells to uninfected cells. (32) says that there are cytoplasmic connections between infected cells and uninfected cells (see appendix (1): 311). (32a) states that the movement of organic nitrogen across cells can freely take

place. (32a) is implicit in (32). The implicational relation between (32) and (32a) depends on the preceding co-text where it is explained that organic nitrogen moves across cells. In (33), on the other hand, the inference paraphrase is preceded by a discussion of the effect of 'bracken fronds' on neighbouring plants, namely that it eliminates them, and the use of types of chemicals, i.e. herbicides, to get rid of bracken. It is explained that these herbicides can eliminate other plants as well, when applied to bracken fronds. (33) in this excerpt states the main points again of the preceding discussion, namely that other plants may suffer from the application of these herbicides (the term "compounds" is used as an equivalent term to herbicides). (33a) states that the use of herbicides to kill bracken has side effects. (33a) is implicit in (33). The implicational relation between (33) and (33a) can only be established by reference to the preceding co-text, because (33) does not say that the use of compounds (i.e. herbicides) is to kill bracken. This information has to be recovered from the preceding discussion.

The dependence on co-text can have a different manifestation, as the following excerpt shows:

(34) but they do exhibit. specificity. and they interact. with. receptors. which are located. in the cytoplasm. of. the cells..

(34a) so we get. within the cytoplasm. a steroid. receptor complex..

(130052)

In this instance, the co-text will establish that they (two occurrences underlined twice in (34) and steroids (underlined in 34a) are co-referential (Halliday & Hasan 1976:3). The implicational relation between (34) and (34a) depends on the phoric relation between the anaphoric pronoun they and the technical term "steroids" (130051). As pointed out in 6.3.4.3 in relation to the (EP), dependence on co-text can be seen as a manifestation of textual continuity. This is discussed next.

6.5 Textual continuity and paraphrase

As discussed in 6.1.1 (4th point) each lecture is a language event, so each lecture is assumed to be a 'text'. It is also assumed that 'text' has continuity. Paraphrase can be looked at as a manifestation of this continuity in the text. This manifestation of textual continuity in the use of paraphrase has two different aspects:

Firstly it is seen in the repetition of:

(i) Information content. This is characteristic of paraphrase, although the paraphrastic relation is different in each type of paraphrase, as discussed in 6.1.0, 6.2.0, 6.3.0, and 6.4.0.

(ii) Lexical terms or use of equivalents in the paraphrastic form. This is mostly a feature of the summary paraphrase or the equivalence paraphrase. In the following two excerpts: (quoted earlier in 6.1.2 and 6.3.1.1).

(7) well. just as. a. deficiency. of hormones can have. quite profound effects. so. can. excesses of hormones. and if. growth hormone from that self same. anterior pituitary region. is produced in excess.. prior. to puberty. you can. get. this enormous growth of the. long bones. and enormous increase in. stature. and I think that. this particular young man is. 8 feet plus. in height.

(7a) so an excess of growth hormone.. prior to. puberty. will. cause. giantism..

(12906)

(21) each megasporangium or ovule. is located on the upper side. of. an ovuliferous scale. protected by. a bract scale..

(21a) megasporanga then. on the upper side. of. ovuliferous scales protected by bract scales...

(333217)

Excerpt (7a) is an instance of summary paraphrase; (21a) an instance of an equivalence paraphrase. In each excerpt instances of lexical repetitions or use of

lexical equivalents are underlined. The inference paraphrase can also show this feature although the paraphrastic relation, in this type of paraphrase, does not depend on lexical relations between the constituent of the two utterances: in excerpt (31) quoted in 6.4.1, the technical term 'bracken fronds' appear both in the original utterance and in the paraphrastic form.

Secondly, manifestation of textual continuity is seen in the relationship between the co-text and paraphrase. This relationship has two manifestations:

(i) It is realized in the role of the co-text in the selection of lexical items, in the paraphrastic form, as discussed in 6.1.1; 6.3.3.1. This type of relationship between paraphrase and the co-text, is only found with the equivalence paraphrase.

(ii) It is also realized in the dependence of the paraphrastic relation on the preceding co-text (see 6.3.4.3; 6.4.2). In each case, the equivalence paraphrase and the inference paraphrase can only be established by resort to the co-text preceding the utterances related by a paraphrastic relation. If we look at the following instance of an equivalence paraphrase.

> but in some way. that. process. has ultimately to be switched off again.
>
> there must be a termination. to. the. hormone effect.
>
> (129732)

We can see that the paraphrastic relation cannot be established unless the co-referentiality of "that process" occurring in the original utterance and "hormone effect" occurring in the paraphrastic form is recovered from the preceding co-text. The interpretation of the paraphrase depends on phoric relations.

So whether dependence is in terms of repetition of lexical items (or use of equivalents) or in terms of phoric relations, it is seen as a manifestation of continuity in the text.

6.6 Distribution of the three types of paraphrase in the corpus

To complete this description of the phenomenon of paraphrase in these lectures the frequency index for each type of paraphrase (per 1,000 words in the corpus) is presented in table (24) below:

Texts	Total Number ofwords	F.I For SP	F.I For EP	F.I For IP	F.I For 3 types of paraphrase
Set 1 (3 lectures)	14.200	(61) 4.2	(11) 0.8	(16) 1.12	(88) 6.19
Set 2 (3 lectures)	14.200	(19) 1.33	(9) 0.7	(3) 0.21	(31) 2.18
Set 3 (3 lectures)	14.600	(17) 1.16	(19) 1.30	(6) 0.41	(42) 2.9
Total (9 lectures)	43.000	(97) 2.25	(39) 0.9	(25) 0.5	(161) 3.7

Table (24) : Frequency index for the summary parphrase (SP), the equivalence paraphrase (EP) and the inference paraphrase (IP) in the corpus

The results show that, in the data as a whole, paraphrase is mostly used in set (1) (with a frequency of occurrence of 2.25 per 1,000 words) (to be picked up again in the next chapter). The type of paraphrase mostly used in the three sets of lectures is the SP, except in set (3) where it is nearly as frequent as the EP.

6.7 Features of unplanned spoken versus planned written discourse in paraphrase

6.7.0 This section identifies and deals with those features of the textual realisation of paraphrase which either belong to unplanned spoken or to planned written discourse. These features are of two types: situational and non-situational features.

Situation features are either: (1) features of the interaction between the text and non-verbal visual aids; or (2) they are audience sensitive features, i.e. they are features of the interaction between the speaker and the audience. Non-situational features, on the other hand, are:

(i) The use of certain grammatical constructions, namely subordination versus co-ordination, finite/non finite clauses, which have been used as indices for measuring grammatical complexity at clause level and differentiating between spoken and written language (see Chapter Three).

(ii) The use of reformulation, which is either: (a) textual, as realised in the recurrent use of paraphrase, in particular the recurrent use of the summary paraphrase: or (b) sentential, ovserved in the repetition of lexical terms and syntactic structures within a clause or a sentance.

(iii) The use of ellipsis

(iv) The use of deictic initial 'so'

6.7.1 Situational features as markers of unplanned discourse in paraphrase

6.7.1.1 Features of the interaction between the textual realization of paraphrase and the visual element

6.7.1.1.1 Reference to the visual element in paraphrase in the lectures

Reference to the visual can be seen in the following instances of a summary paraphrase, and of an equivalence paraphrase (quoted earlier in 6.3.1.1):

(35) so if this is. the hypothalamic zone. up here nerves emanating from this area, could pass. down tracks. linking it. to the. posterior... part of this gland, or alternatively. minute blood vessels. may link. the hypothalamus. with the. front portion. of the gland. referred to. as. the anterior.. pituitary....

(128340)

(19) if we take this structure. and fold it round. so that the ends meet.

(19a) that is we take these.. two edges. fold the thing round so that these two ends fuse together.

(33364)

In each instance, (35) and (19a), the paraphrase is produced with continued reference to the visual which has been used with the previous discussion. Reference to the visual is realized by the use of the deictic terms (underlined in the excerpts): "this" (used 3 times); "up here"; "these" (used twice); "the thing"; to refer to the appropriate parts of the slides.

Table (25) overleaf shows the frequency of occurrence of each type of paraphrase accompanied by a visual per 100 occurrences of instances of each type of paraphrase in the data.

Texts	F.O. of SP + V	Tot no of occ of SP	F.O. of EP + V	Tot no of occ of EP	F.O. of IP + V	Tot no of occ of IP	F.O. of Para-phrase + V	Tot no of occ of P
Set 1 3 lectures	(20) 32.78	(61)	(4) 36.36%	(11)	(2) 12.5%	(16)	(26) 29.54%	(88)
Set 2 3 lectures	(9) 47.36%	(19)	(7) 77.77%	(9)	(0) (0)%	(3)	(16) 51.61%	(31)
Set 3 3 lectures	(7) 41.17%	(17)	(16) 84.21%	(19)	(3) 50%%	(6)	(26) 61.90%	(42)
Total 9 lectures	(36) 37.11%	(97)	(27) 69.23%	(39)	(5) 20%	(25)	(68) 42.23%	(161)

Table (25) : Frequency of occurrence of each type of paraphrase accompanied by a visual in the corpus.

The (EP) is the type of paraphrase most often accompanied by the use of a visual with the three speakers. The use of the visual is mostly exploited in set (3) where it is highly recurrant with all types of paraphrase. (To be taken up again in Chapter Seven).

6.7.1.1.2 Features of oral delivery in the reference to the visual in paraphrase in the lectures

As observed in Chapter Five (5.5.2) non-verbal elements (i.e.pictures, diagrams.) are also used in written language. It was pointed out then, that reference to the visual element is realised in a different way in each modality. The distinctive characteristics of the interaction between the text and the visual in

written and spoken language can be, at the risk of repetition, briefly stated again as follows:

- reference to the non-verbal is generally carried out by the use of a figure between brackets in a written text; in a spoken text it is realised by place deictic terms.

- the non-verbal element is part of the text in written language, but outside the text in spoken language.

- the deictic terms which are used to refer to the visual are only interpretable by reference to elements of the immediate environment. (See 5.5.2 for a full discussion of this point); In a written text, the interpretation of a reference to the non-verbal is intra-textual.

As observed in (5.5.2) these features of the reference to the visual in the lectures are characteristic of oral delivery. They are characteristic of spoken as opposed to written academic texts.

6.7.1.2 Audience oriented features in paraphrase

There are constituents in paraphrase, mostly the summary paraphrase, which are used to involve the audience in the on going text. In the following examples (36), and (8) (quoted earlier in 6.2.1.1).

(36) so a whole host. of different materials. produced. by the pituitary. so you can imagine, that chopping this out. is going to have widespread effects. on all. of these. target tissues.. so not only. might you expect those little dogs. to have growth problems. you might expect them to have problems. in all these other areas as well.

(12895)

(8) so erm. that is the ground work then for the structure of the vascular. bundle. you realise that the whole of this space is actually filled. with tissues. of the. xylem. the phloem. and the pericycle. containing these particular. cell types..

(230929)

"So you can imagine", "might you expect", "you might expect" in (36) and "you realise that" in (8), are used to invite audience participation. They are essentially interactive (see Sinclair 1981) as they are extrinsic to the conceptual content of the summary paraphrase. Their unique function is to involve the audience in the on-going text. The use of the pronoun 'you' is a marker of such a function. Furthermore, the referent of you, here, is a specific group of individuals, i.e. "you" here is person deictic. This use of person deictic "you" is a feature of oral delivery since when it is used in writing, the speaker is a specific individual whereas the reader could be anybody (Lakoff's visibility criterion 1979).

6.7.2 Non-situational features in relation to unplanned spoken versus planned written discourse in paraphrase in the lectures

6.7.2.1 Aspects of grammatical complexity in paraphrase

As dealt with in (5.6.3), in relation to the sub-topic shift signalling utterance, grammatical complexity as measured in:

- the use of subordination versus coordination (see 3.3 for the discussion of these terms),
- the degree of subordination,
- the use of finite versus non finite clauses,
- most recurrent types of clause,

is analysed in paraphrase to identify the role of paraphrase in defining the lecture along the spontaneous spoken versus non-spontaneous written continuum with respect to this feature.

6.7.2.1.1 Subordination versus coordination in paraphrase

The analysis of the frequency of occurrence of co-ordinate versus subordinate clauses shows the following results:

Texts	Total no of words in the 3 types of paraphrase	F.I For co-ord cl	F.I For sub/emb cl
Set 1 (3 lectures)	1791	16.20 (29)	53.63 (96)
Set 2 (3 lectures)	936	5.34 (5)	61.96 (58)
Set 3 (3 lectures)	990	14.14 (14)	49.49 (49)
Total (9 lectures)	3716	12.91 (48)	54.62 (203)

Table (26) : Frequency index for co-ordinate versus subordinate clauses in paraphrase

The three sets of lectures favour the use of subordination as opposed to coordination in paraphrase. So, the characteristic of lectures to use more subordination than coordination is found at the level of paraphrase, as is found in the sub-topic shift signalling utterance. As observed in (2.4.4) and (3.3), this tendency to favour subordination over coordination is typical of written language.

6.7.2.1.2 Depth or degree of subordination in paraphrase

Table (27) overleaf, shows the frequency of occurrence of sentences with different degrees of subordination in paraphrase, in each set of lectures(8).

Texts	FO of 0 degree complex s	FO of 1 degree complex s	FO of 2 degree complex s sent	FO of 3 degree complex s	FO of 4 degree complex s	FO of 5 degree complex s	No of words in the 3 types of P
Set 1	21.21 (38)	22.33 (40)	8.93 (16)	0.55 (1)	0	0	1791
Set 2	17.09 (16)	18.16 (17)	7.47 (7)	3.20 (3)	1.06 (1)	1.06 (1)	936
Set 3	24.24 (24)	21.21 (21)	5.05 (5)	1.01 (1)	1.01 (1)	0	990
Total	19.64 (73)	20.99 (78)	7.53 (28)	1.34 (5)	0.53 (2)	0.26 (1)	3716

Table (27) : Frequency index for sentences distinguished on the basis of their degree of complexity in paraphrase (calculated per 1,000 words in paraphrase)

Sentences in paraphrase, mostly have (0) or (1) degree complexity. This tendency is found when comparing individual sets of lectures. Sentences with (1) degree complexity are mostly used, except with set 3 where (0) degree complexity sentences come first. (However the difference between the FO of (0) degree complexity sentences and (1) degree complexity sentences is negligible).

The conclusion which can be drawn from these results is that paraphrase shows a clear tendency to use simpler sentences, a feature also found in the analysis of the degree of complexity of sentences in the whole data. As observed then, these findings put the lectures nearer to the spoken end of the continuum.

6.7.2.1.3 Finite versus non-finite clauses in paraphrase

The analysis of the frequency of occurrence of finite and non-finite clauses per 1000 words in the paraphrase, shows the following results:

Texts	F.I For finite clause	F.I For non-finite clause	Total no of words in the 3 types of paraphrase
Set 1	(151) 84.31	(28) 15.65	1791
Set 2	(93) 99.35	(13) 13.88	936
Set 3	(107) 108.08	(15) 15.15	990
Total	(351) 94.45	(56) 15.06	3716

Table (28) : Frequency index for finite and non-finite clauses in paraphrase

The results in this table show that finite clauses are far more frequent than non-finite clauses. Again, the finding applies to the sub-topic shift signalling utterance as it indeed applies to the whole data. As pointed out in Chapter Three the preferential usage of finite over non finite clauses is a feature of spoken language.

6.7.2.1.4 Favoured types of dependent clause in paraphrase

The analysis of the clausal composition of paraphrase in the data, shows that the most recurrent types of dependent clause are the embedded finite clauses "that clause" , and the restrictive relative clause (with a frequency of occurrence of 19.70 and 16.74 respectively per 1,00 occurrences of dependent clauses in paraphrase).

The use of "that clause" in paraphrase is exemplified in the following excerpt of a summary paraphrase:

(37) so you can see that er both aporsity and superfluity of hormones can have quite dramatic effects

(129111)

The underlined clause is embedded as object into the main clause "so you can see"

Table (29) shows the frequency of occurrence of that-clause in the 3 sets of lectures, per 100 occurrences of subordinate clauses in paraphrase:

Texts	No of occur sub/cl in P	FO of "that-clause"
Set 1 (3 lectures)	96	18.75% (18)
Set 2 (3 lectures)	58	15.51% (9)
Set 3 (3 lectures)	49	26.53% (13)
Total (9 lectures)	203	19.70% (40)

Table (29) : Frequency of occurrence of "that -clause" in paraphrase

These results show that "that-clause" has a comparable frequency of occurrence in the 3 sets of lectures. It can be looked at as a feature of paraphrase in the lectures. So as discussed in (3.6.2 and 5.6.3.4) the use of complement clauses (that-clause) is interpreted as a feature of the integrated nature of planned written language (Chafe 1982).

As explained in 3.6.1. and 5.6.3.4, the restrictive relative clause cannot be used as indicator in relation to the polarity. Therefore it will not be dealt with any further in the analysis of paraphrase.

The findings of this analysis of grammatical complexity applied to paraphrase can be summarised as follows:

1 Paraphrase uses more subordination than coordination

2 Paraphrase favours the use of sentences with a low degree of complexity, i.e. simpler sentences

3 Paraphrase uses more finite than non finite clauses

4 Paraphrase favours the use of complement "that-clause"

We can add, as a further point, that paraphrase is grammatically realised in structures which are readily analysable in sentence like units. Excerpt (8) discussed earlier is structurally realized into two sentences.

> //so erm that is the ground work then for the structure of the vascular bundle// you realise that the whole of this space is actually filled with tissues of the xylem the phloem and the pericycle containing these particular cell types//
>
> (230229)

This last point and the result in (1) and (4) are interpreted as features of written language (see 3.7). The results in (2) and (3) are features of spoken language (see 3.2 and 3.4) Again, the tendency of lectures to combine features of spoken and written language is the outcome of this analysis of grammatical complexity, applied to paraphrase. These results just echo, those found in Chapter Three. We can observe again that grammatical complexity, whether applied to the whole lecture or just to a textual feature, does not tell us much about where biology lecture discourse belongs on the spontaneous spoken versus non-spontaneous written continuum.

6.7.2.2 Deictic 'so' as paraphrase initiator

Deictic 'so' as paraphrase initiator is a feature of two types of paraphrase discussed here: the summary paraphrase (see 6.2.3.2) and the inference paraphrase. Deictic "so" introduces every instance of inference paraphrase. It also introduces the majority of SP's in the 3 sets of lectures (82.19%).

6.7.2.3 Ellipsis, a feature of spoken language in paraphrase

As discussed in 6.2.4, ellipses of certain words or structures is a characteristic feature of the summary paraphrase. It was observed then, that the inherent characteristic of the summary paraphrase of being a briefer reformulation of a fragment of text, favours the use of ellipses. Condensation of information is achieved at the expense of the grammatical structure (6.2.4.2).

Ellipsis is also found in the equivalence paraphrase, but it is rather rare. The following excerpt is one of the few instances of equivalence paraphrase which show this feature:

> each megasporangium or ovule is located on the upper side of an ovuliferous scale protected by a bract scale
>
> megasporangia then on the upper side of ovuliferous scales protected by bract scales
>
> (333217)

Where the grammatical constituents 'each', 'an' and 'a' are dropped, (underlined in the excerpt), the structure 'is located' is also ellipted. The ellipsis of all of these constituents does not affect the information content of the utterance.

As observed in 6.2.4 the use of ellipsis is a feature typical of spoken language. It is a manifestation of the tendency of spoken language to show a "certain degree of linguistic economy "(Darian 1983: 38). This is manifested in paraphrase, mostly the summary paraphrase, when omitting constituents does not have any or has little effect on the meaning of the utterance.

6.7.2.4 Repetition as a feature of oral delivery in the lectures

Repetition in the lectures is manifested at 2 levels:

1 at the level of the text: in the frequent occurrence of paraphrase, whereby an utterance is reformulated as a paraphrase

2 at the level of the sentence: where the reformulation applies to a constituent(s) of a sentence

(1) Repetition as paraphrase of utterance

As dealt with throughout this chapter, paraphrase is a type of repetition of a stretch of text. This repetition, because it is highly recurrent in the lectures introduces a redundancy in the text, which is generally avoided in written language. A written text provides accessibility to the reader which partly eliminates the need for repetition. A spoken text, on the other hand, unfolds in time, what is said is not accessible again to the listener unless it is repeated either through exact repetition or reformulated in other words, i.e. through paraphrase (see also discussion in 6.1.1).

This is not to say that there is no paraphrase at all in academic written language. We can find the summary paraphrase, the inference paraphrase, as discussed here, and a type of equivalence paraphase, namely paraphrase which explains a technical term or expression, in a written academic text. However the recurrent reformulation of a stretch of text as observed in the use of the summary paraphrase here, is not expected in a written academic text. Therefore it is a feature of spoken language in the lectures.

(2) Repetition as repetition or reformulation of a constituent(s) of a sentence

It can have different manifestations:

(i) It can be a case of lexico-syntactic repetition of part of a clause. This

can be seen in the following examples:

(38) so this work with aphids in particular encourages us er encourages us in the belief that .....

(231921)

(39) so what I'm trying to. say to you is that. in the control of something. apparently as trivial .as blood glucose levels. we have insulin, we have glucagon, we have adrenalin, we have ACTH, we have growth hormone, and indeed some of the releasing hormones from the hypothalamus .....

(129419)

Excerpt (38) and (39), which are both instances of a summary paraphrase, exemplify different phenomena. In (38), the phrase "encourages us" (underlined in the excerpt) is repeated. This repetition seems to have been used by the speaker to avoid a gap of silence while planning the rest of the utterance. It can also be interpreted as a manifestation of hesitation phenomenon. In either case it is typical of a text which is produced under the constraints of real time.

Excerpt (39), on the other hand, shows a more interesting phenomenon. The same syntactic structure: NP+V+NP (instances underlined in the excerpt) is repeated five times running in the text, for a rhetorical purpose: the speaker is making the point that as many as five hormones act simultaneously, in the control of a single function in the body, (namely the regulation of blood glucose levels).

This type of repetition can be described as being partly a manifestation of structural parallelism (cf Quirk al 1985: 1427) where the repeated structures show the features of clause structure, tense and word order. In addition, they use the same constituents: "we" as NP subject; "have" as a main verb, and a technical term as NP complement. The same structure is repeated here for emphasis.

Rhetorical emphasis can be a feature of a written academic text. However, it would be expected to have a different textual manifestation. A written academic

text would omit the constituents "we" and "have", and use a series of co-ordinated NP complements (as technical terms) instead:

> "we have insulin, glucagon, adrenalin, ACTH and growth hormone.....:

So we can say that repeating the whole clause as exemplified in (39) is typical of oral delivery.

(ii) Repetition in paraphrase can also be a case of lexical revision of a constituent of a clause; as exemplified in:

(40) so Munch envisaged. erm. a circulation. of water. and sugar solution. the water. in the simplast of the xylem vessel. the sugar. sorry the the water in the apoplast of the xylem vessels. the sugar. in the simplast of the living cells .....

(231326)

Where the speaker reformulates part of his utterance by using a different technical term as he reformulates part of the description of the phenomenon (about the movement of sugar solution in the root nodule). This is a feature of a text which has not been edited, which is in the actual process of being produced. This feature is typical of speech.

6.7.2.5 Interrupted structures in paraphrase

Interrupted structures in paraphrase refer to cases where the utterance is interrupted at a certain point either by an aside or a metatextual comment. The following examples illustrate each case:

(41) so collectively... {sorry}.. collectively. the glucagon. the adrenalin. and these glucocorticoids .....

(129411)

(42) and as a consequence blood levels go up..and as we remarked a moment or two ago. this is the specific stimulus for the release of insulin

(129318)

In excerpt (41) the beginning of the summary paraphrase is interrupted by an aside (underlined in the excerpt). The initial word in the utterance is then repeated when the summary paraphrase is resumed. The aside "sorry" is produced as the speaker selects the next slide to be shown to the audience (the wrong slide is picked up) (analyst's observation). It is some kind of comment on an extra-linguistic activity. It is outside the content of the summary paraphrase. This type of 'digression' is typical of oral delivery and has no counterpart in a written text.

Excerpt (42) is interrupted by a metatextual expression (underlined in the excerpt) which qualifies the rest of the utterance as a repetition of something that has been said before. It is possible to find this type of comment in a written text, i.e. a metatextual statement which refers to a previous part of the text. However, the type of lexical items used here would not be used in a written text, namely "a moment or two ago", which refers to real time. The feature of oral delivery, here, is the use of this type of lexical word.

6.8 Conclusion

In conclusion four points can be made:

- Firstly, the frequent use of paraphrase, in particular the summary paraphrase, which introduduces redundancy in the text is a feature of spoken language.
- Secondly, the summary paraphrase being the most recurrent type of paraphrase in the three sets of lectures, is a prominent feature of the lecture as a variety.
- Thirdly, while the summary paraphrase is used to demarcate sequence boundaries, and hence, is a feature of the macrostructure of the text, the equivalence and inference paraphrase are both found within the boundaries of

the sequence. They are both features of the internal structure of the sequence.

- Fourthly, the grammatical realisation of paraphrase in general and summary paraphrase in particular, shows features of spoken language (such as ellipsis, repetition of lexical or grammatical constituents, interrupted structures). However, we also note that paraphrase is analysable in sentence -like units. The analysis of grammatical complexity, finally, reveals features of both spoken and written language.

Footnotes

(1) cf Sampson (1980) who defines paraphrase as inferrable statements only.

(2) See Urquhart 1982.

(3) See Strang (1962); Lyons (1977).

(4) The notion of 'text' as part of a linguistic theory is by no means easy. It has been expressed succinctly by Lotman (1977), who lays two prerequisites for the notion of text, namely that a text is distinguished from a 'non-text' on the one hand, and 'another text' on the other. Ellis (1976) has linked these two conditions to the idea of continuity as a fundamental aspect of 'text' as a linguistic unit. Continuity has been assumed to be an essential process in text by De Beaugrande & Dressler (1981), who are concerned about the actual processing of text (i.e. performance or instance) rather than the theoretical delimitation as discussed in Lotman (ibid) and Ellis (ibid).

(5) See also Cook (1975 : 51)

(6) See Quirk et al (1972 : 541 ; 1985 : 861) for a discussion of another type of ellipsis which depends on extra textual recoverability and which is restricted to spoken language.

(7) While we object to Sampson's use of inference as the essential element in a paraphrase, we are arguing here that inference paraphrase is one type of paraphrase.

(8) See 3.2 for a discussion of the analysis of the degree of complexity of sentences in the corpus.

CHAPTER SEVEN

SOME ASPECTS OF INTER-SPEAKER VARIATION IN THE USE OF GRAMMATICAL COMPLEXITY, CONNECTIVES, SUB-TOPIC SHIFTING AND PARAPHRASE

7.0 The previous chapters have each dealt with a textual feature common to all three sets of lectures. This chapter concentrates on differences between the three sets, observed in certain aspects of the realization of the textual features investigated. It will first report those differences; then, it will look into possible factors to account for those differences.

7.1 Comparability between speakers

Before we consider these differences, it is useful to briefly state the elements of homogeneity: which are found in the textual realization of the features investigated in the previous chapters.

The first point which must be made is that the overall picture is that of homogeneity: the findings of the investigation of the three features are generally comparable. All three speakers use sentence-initial connectives, the sub-topic shift signalling utterance and paraphrase. They all make use of the same types of connectives, sub-topic shift signalling utterance and paraphrase, dealt with in this study. The second point is that the conclusions drawn from the analysis of grammatical complexity in the data, are all based on features which are common to the three sets of lectures. It is found that

- Firstly, the distribution of subordinate as opposed to coordinate constructions: is comparable in the three sets. The preferred use of subordination is clear cut in each set. As Table (4) shows, the frequency

index for subordinate in set (1) is 53.59 per 1000 words as opposed to 9.43 per 1000 words for coordinate clauses; in set (2), it is 58.73 as opposed to 7.18 and in set (3), it is 49.86 as opposed to 9.45.

- Secondly, the embedded clause is more used than the subordinate clause by all three speakers. Table (6)shows that the frequency index for embedded clause in set (1) is 38.02 per 1000 words as opposed to 15.56 for subordinate clauses; in set(2), it is 40.21 as opposed to 18.52 and in set(3)it is 37.87 as opposed to 11.98.

- Thirdly, the use of finite clauses contrasted with non finite clauses is equally comparable. Finite clauses show a frequency index of 93.09 per 1000 words as opposed to 17.72 for non finite clause in set (1); in set (2) they show a frequency index of 93.52 as opposed to 16.61 and in set (3) their frequency index is 105.13 as opposed to 18.49. (see table 8).

- Fourthly, the distribution of the types of favoured embedded clauses is also similar in the three sets. As seen in 3.6, the types of embedded clauses mostly found are the restrictive relative clause, with a frequency of occurrence of 18.80 per 1000 words in set (1); 18.45 in set (2) and 17.53 in set (3); and that-clause; with a frequency of occurrence of 8.38 per 1000 words in set (1) ; 9.64 in set (2) and 7.46 in set (3) (see table 9).

These results are evidence of a high degree of homogeneity in the three sets, in features of biology lecture discourse discussed in this work. However, variation in some aspects of those features is also observed between the three sets of lectures. This is dealt with next.

7.2 Variation between speakers

7.2.1 Variation in the use of zero degree grammatical complexity

As discussed in 3.3.2, the overall tendency of lectures is to use one degree complexity sentences (sentences which have the structure of a main clause followed by a subordinate or an embedded clause, or a main clause followed by two or more coordinated subordinate or embedded clauses). However, when we look at the use of different levels of complexity in each set of lectures, we can see that zero degree complexity sentences (i.e. sentences which are just a main clause or two or more coordinated main clauses) show the highest score of frequency of occurrence. Speaker three exploits it most (with a frequency index of 24.24 per 1,000 words, compared with a frequency index of 13.16 in set 2 and 14.64 in set 1) (see Table 5) the chi squared technique has proved that the difference in frequency of occurrence of zero degree complexity sentences, between the three sets of lectures is significant ($X^2 = 60.04$ with 2 degrees of freedom at the 1 per cent significance level).

It is interesting to note that a significant difference in the use of zero degree complexity sentences is found again between the three sets, at the level of the textual realization of sub-topic shift signalling utterance ($X^2 = 12.1$ with 2 degrees of freedom at the 1 per cent significance level). Set (3) is again singled out in its marked use of zero degree complexity sentences which show a frequency index of 55.92 per 1000 words as opposed to 24.93 in set (2)and 24.41 in set (1). (see Table 20).

7.2.2. Variation in the frequency of occurrence of one-word connectives "but", "now", "so" and "well"

Speakers vary in the extent to which they use connectives "and", "but", "now", "so" and "well" discussed in Chapter four, only connective "and" shows a comparable distribution in all three sets of lectures. Connectives "but, "so", "now" and "well" are exploited to different degrees. The difference between the three sets in the frequency of occurrence of these four connectives has been proved to be significant when applying the chi squared technique. Thus for "but" the chi square is $X^2 = 9.8$; for "so" it is $X^2 = 47.96$; for "now" it is $X^2 = 16.22$ and for "well" it is 17.15; each chi square value is with 2 degrees of freedom at the 1 per cent significance level).

Set (2) stands out with the highest frequency of "now" and "well". "Now in this set has a frequency of occurrence of 25.27 per cent out of the total number of all one word connectives in the same set, compared with 11.07 per cent in set (1) and 15.96 per cent in set (3); and "well" has a frequency of occurrence of 9.74 per cent in set (2), compared with 2.21 per cent in set (1) and 3.36 per cent in set (3). (See table 12). Set (2) also stands out with the lowest frequency of occurrence of connective "but", which has a frequency of occurrence of 4.33 per cent compared with 9.59 per cent in set (1) and 12.18 per cent in set (3). (See Table 12).

Set (1), on the other hand heavily uses connective "so" which shows a frequency of occurrence of 36.16 per cent out of the total number of all connectives as opposed to 15.16 per cent in set (2) and 10.08 per cent in set (3). (See Table 12).

7.2.3 Variation in the use of thetextual structure: sub-topic topic shift signalling following a summary paraphrase (SP & STSS)

It was observed in 5.4.2.2, that the sub-topic shift signalling utterance is often preceded by a summary paraphrase. Both textual features mark off the boundaries of a sequence (see 5.4.2.2.). This textual structure is mostly found in set (1) where the frequency of occurrence of STSS & SP is 51.92 per cent out of the total number of all STSS utterances in the same set, compared with 13.33 per cent in set (2) and 14.75 in set (3). The difference between the three sets in the frequency of occurrence of the structure STSS & SP has been found significant (X^2 = 18.73 with 2 degrees of freedom at the 1 per cent significance level.)

7.2.4 Variation in the use of personal pronoun "I" in the sub-topic shift signalling utterance

The obvious difference between the three sets in the use of personal pronouns in the sub-topic shift signalling utterance is the absence of the pronoun "I" in set (3). (See Table 18). As observed in 5.6.2, "I" is the personal pronoun mostly used in the STSS utterance in the whole data. However, it is non-existent in set (3). There seems to be a preferance for using"we" instead.

7.2.5 Variation in the frequency of occurrence of paraphrase and the summary paraphrase

A significant difference between the three sets in the distribution of paraphrase is found (X^2 = 34.99 with 2 degrees of freedom at the 1 per cent level of significance). Set (1) uses it most where the frequency index of paraphrase is 6.19 per 1,000 words compared with 2.18 in set (2) and 2.9 in set (3). (See Table 24). When we consider types of paraphrase, a noticeable difference is found in the

frequency of occurrence of the summary paraphrase. The chi squared technique has proved that the difference is significant ($X^2 = 39.2$ with 2 degrees of freedom at the 1 per cent significance level). Set (1) uses it most. It shows a frequency of occurrence for this type of paraphrase of 4.2 per 1,000 words compared with 1.33 in set (2) and 1.16 in set (3).

7.2.6 Variation in the use of the visual in paraphrase

It was observed in (6.7.1.1.1), that the use of the visual element (slides and diagrams) in paraphrase is mostly used in set (3) and set (2), with a frequency of occurrence of 61.90% (per 100 occurrences of paraphrase) and 51.61% respectively, compared with set (1) where the frequency of occurrence is only 29.54%. The difference between the three sets in the frequency of occurrence of the visual in paraphrase has been proved to be significant ($X^2 = 7.84$ with 2 degrees of freedom ,but at the 5 per cent significance level).

7.2.7 Non significant difference in the use of types of change of structure in sub-topic shift signalling (STSS) and in the use of the visual element in the signalling utterance

A difference was observed in the distribution of types of the change of structure signalling utterance in the three sets of lectures (see table 14), and in the use of the visual element in the STSS utterance (see Table 17). However, the chi squared technique has shown that these differences are not significant. The chi squared for the distribution of types of change of structure is $X^2 = 5.01$; for the frequency of occurrence of the visual element in the STSS utterance in the three sets, it is $X^2 = 4.61$ (with 2 degrees of freedom at the 5 per cent significance level).

7.3 Accounting for variation between speakers

There does not seem to be one simple element which would explain all aspects of variation described in the preceding section. A few factors may be suggested.

7.3.1 Personal style

Three features seem to be nearest to a manifestation of personal style. Firstly, at the level of grammar, the extensive use of zero degree complexity sentences in set (3). The selection of clause types is to a large extent evidence of individual performances as choices in grammar are open to all speakers (bearing in mind that the contextual parameters of setting, academic level of audience, general topic are comparable). Although the selection of sentence structure could be determined by certain functions such as expressing cause and effect, still there is room for individual preferences. Evidence of personal choice could be seen in the use of zero degree complexity structures, such as instantiated in the following excerpt from set (3):

> //You take a fern plant/2 and you turn it over/3 and you will see on the underside a whole series of little kidney shaped scales//
>
> (332324)

which has the structure of three coordinated main clauses. Evidence of personal choice can also be seen in a combination of clause types resulting in high degree complexity sentences. This can be illustrated in the following excerpt from set 2:

> //They withdraw their stylets/ and they scramble over the plant/ and insert their stylets in other plants/ because they are searching for a better source of food/ indicating perhaps/ that the flow of the xylem erm flow of the erm flow of the sap in the sieve tube elements sieve tubes is er reduced/ when the leaves wilt/which

> makes sense/ because when the leaves wilt/ it means the leaves have become flaccid/ and that means/ the pressure in the supplying cells has fallen//
>
> (231819)

which has the structure : McMcM Sub^1 Sub^2 E^{t3} Sub^4 Sub^5 Sub^6 Sub^7M Emb^{t1} cM$Emb^{t1(2)}$ the sentence has a score of 7 degree complexity. Although rare, this type of structure appears in the data, in particular with speaker (1) and (2) (see Table 5.).

Secondly, personal style is also sometimes manifested in the use of personal pronouns. As discussed in 7.2.4 personal pronoun "I" is not used in the STSS utterance in set (3) where the pronoun found in the same function is "we"(3). (See Table 18). As observed in 5.6.2 the frequent use of personal pronoun "I" is generally regarded as being typical of informal unplanned spoken language. This could be interpreted as evidence of the fact that. with respect to the use of pronouns, the 3 sets of lectures exploit different levels of formality, with set (3) favouring the more formal level. This feature can be attributed to personal style.

Thirdly, personal style can also explain the greater use of "well" in set 2. As discussed in 4.5.3, the use of "well" in this data, is a feature of unplanned spoken discourse in the lectures. It was then observed that "well" mostly appears in two uses:

- as editing marker to reformulate a statement, or part of a statement. This is seen in:

> //now if if this amount of sap was moving in a single sieve tube element that means in a tube composed of single cells placed in two ends it would mean that the amount of sugar solution well the amount of sap which passes through each cell.....
>
> (231828)

where "well" precedes the reformulation of part of the sentence.

- as answer initiator, following a rhetorical question, as seen in:

> //now what are the tissues of transport// well its er the carbohydrate brought into the erm nodule through the sieve tubes of the phloem tissue
>
> (230442)

"well" here initiates the sentence which follows a question described in the analysis as a change of structure STSS utterance. (see 5.3.3). As pointed out in 4.5.2.3 this use of "well", (prominent in set 2), manifests a feature of spoken language in the lectures.

7.3.2 Discourse topic

Another motivating factor for the use of a feature seems to be discourse topic. The choice of sentence structure can also be influenced by topic. Thus, the use of zero degree complexity sentences in set 3 may be motivated by the topic when it is the description of the physical layout of a plant or part of a plant. this can be seen in the following excerpt from set (3).

> //[1]here is the egg cell//[2] you can even see the nucleus of the egg cell inside the cytoplasm/embedded in the tissues of the gametaphyte//[3] here are the wall cells//[4] this one has not yet opened up//[5] you can still see in here neck canal cells//
>
> (332454)

Out of these five sentences, only sentence (2) has a score of one degree complexity (its structure is M $Emb^{en(2)}$). The other four sentences are simple main clauses. The excerpt is a description of a diagrammatic representation of a plant as shown on a slide.

A typical instance where topic seems to be the motivating factor for the use of a feature is in the use of paraphrase in set (1), in particular the summary paraphrase,

which is by far the most frequent type of paraphrase (see Table 24) and in the use of the visual in paraphrase in set (3) and set (2).

In set (1) the abstract topic of endocrinology (theory of the nature, properties and functions of hormones) seems to call for a greater use of summary paraphrase which functions as a reminder or a kind of aide-memoire to help the audience (i.e. listeners) through the process of explaining aspects and relations of the topic. The summary paraphrase is used to function as an aid to comprehension. In set (3) and set (2) on the other hand, the topic of plant biology (description of the morphology and physical layout of plants) requires a visual illustration by means of slides and/or diagrams drawn on board to support the description of the structure of a plant or part of a plant.

7.3.3 Structuring strategies

We have seen that there are idiosyncrasies in the choice of sentence structure, the use of particular personal pronouns and the use of a particular connective; we have also seen that topic may explain certain differences such as using more summary paraphrase, or exploiting visual aids more in the summary paraphrase as it may partly explain the choice of a particular sentence structure. Further, it seems to be the case that structuring strategies are also used to carry out certain features in the lectures. This could be seen in the extensive use of connective "So" in set (1), and "now" in set (2) (see 7.3.2). "So" in set (1) is relatable to the use of paraphrase as it is mainly used to introduce paraphrase. "Now", prominent in set (2), is mostly used to initiate sub-topic shift signalling (see 5.4.1.1). Both "now" and "so" signal the boundary of a sequence (see 4.7.1).

The use of "so" and "now" as discussed above are manifestations of structuring strategies, as they are used to signal textual features and relations.

A final textual feature which can be explained as evidence of text structuring is the use of summary paraphrase immediately followed by a signalling utterance. As discussed in 7.3.5, this is mostly found in set (1).

7.4 Conclusion

Although the general picture is that of similarity among the three sets of lectures, we find that there are significant differences in the elements that realize a certain feature. These differences can be summarized in the following three points:

1 - Differences between the three sets are manifested at the level of sentence structure (greater use of zero degree complexity sentences in set (3), though the total is comparable. This is a reflection of singularity .

2 - The summary paraphrase is mostly used in set (1) and the use of the visual in paraphrase is mostly found in set (3). Topic is the motivating factor for both the frequent use of the summary paraphrase in set (1), and the use of the visual in paraphrase in set (2) and set (3).

3 - Different types of connectives (now, well, so) are used, partially as a strategy for structuring and organising discourse; and partially to introduce a given feature (i.e. a summary paraphrase or a sub-topic shift signalling utterance).

It is true that the lecture being the product of a given context of situation manifests a high degree of homogeneity at the level of grammar, and the three

features studied here. But it is also important to recognize the differences that are found in different individual lecturers, differences that may be determined by personal style, discourse topic or structuring strategies.

Footnotes

(1) Sample of the derivation of a chi squared value, applied to the frequency of occurrence of "now" in the three sets of lectures (see table 12)

1.

30	70	38	138
271	277	238	786

2. $E = \frac{138 \times 271}{786} = 47.58$

$E = \frac{138 \times 277}{786} = 48.63$

$E = \frac{138 \times 238}{786} = 41.78$

3. $X2 = \frac{(30 - 47.58)^2}{47.58} + \frac{(70 - 48.63)^2}{48.63} + \frac{(38 - 41.78)^2}{41.78}$

$= 6.49 + 9.39 + 0.34 = 16.22$

$X^2 = 16.22$ (with 2 degrees of freedom at the 1 per cent significance level).

(2) M stands for "main" clause; Emb, for embedded clause and Sub, for subordinate clause. Emb^t indicates "embedded that-clause"; Emb^{en}, embedded non finite en-clause'; c stands for co-ordinating conjunction.

(3) We notice that "we" is used interchangeably with "I" as in this following example from set (3):

> it is to the ferns that we're going to. look at. Today let's look at a typical fern..
>
> (332141)

this use of "we" which involves the addressee is also reinforced by the use of jussive inclusive imperative "let's" in the following sentence. (see our comment on this type of imperative and its use in the present data in (5.3.3.2).

CHAPTER EIGHT

SUMMARY AND CONCLUSIONS

THEORETICAL IMPLICATIONS

8.0 This chapter presents a summary of the research and findings of the present work. It discusses the theoretical implications of these findings. It also discusses their potential pedogogical applications. Finally, it puts forward suggestions for further research.

8.1 Summary of research

At the beginning of this investigation we set ourselves the task of identifying biology lecture discourse along the unplanned spoken versus planned written language continuum. Chapter one presents the main question of the investigation and describes the present corpus of biology lectures (nine lectures -4,3000 words). Chapter two which states the concerns of this study, asks three main questions: firstly, whether biology lecture discourse can be established to be nearer to the spoken mode or to the written mode; secondly, whether the investigation of the features of grammatical complexity, sentence-inital connectives, sub-topic shift signalling and paraphrase will enable us to answer that question; thirdly, what insights into the nature of the discourse of biology lectures will be gained from the investigation of these features.

As a first step taken towards answering these questions, Chapter three undertakes an analysis of some aspects of grammatical complexity. Endorsing the views that: (1) subordination as opposed to co-ordination is an indicator of higher complexity in language, (2) finite as opposed to non finite clauses are typical of spoken language, (3) some types of clause are more associated with one mode or the other, this Chapter presents an investigation of the clausal structure of sentences

in this data. Complexity is mainly dealt with in terms of (a) frequency of subordinate construction as opposed to co-ordinate constructions (2) depth or degree of subordination/embeddedness, (3) the most frequent types of embedded clause.

The next three Chapters take the analysis beyond the boundaries of the sentence, at the level of discourse. Some features of the utterance are investigated and aspects of their textual realization studied in the light of the question of the relationship between biology lecture discourse and the unplanned spoken versus planned written language dichotomy.

Chapter four studies sentence-initial connectives. Their use in these lectures is basically examined in relation to the concept of relevance, whether it is semantic or pragmatic relevance. Semantic relevance is described in terms of semantic progression which is seen in the thematic structure of a given stretch of text. It is also identified as manifested in the use of an utterance initiated by a connective to paraphrase a preceding stretch of text or to focus on a particular aspect of the sub-topic under discussion. The pragmatic relevance of the use of a connective is revealed in the speaker's strategy to monitor his speech pedagogically by the use, for example, of sentence initial "well" immediately following a rhetorical question. It is also revealed in the use of a connective initiated sentence to invite audience participation, as instantiated in "and you may be familiar with some of these:" (128442).

The use of textual connectives is also examined in relation to the text structure. The role of "so" "and" and "now" as markers of a boundary in the text is highlighted.

Features of semantic and pragmatic relevance are assessed in the light of the unplanned spoken versus planned written language distinction. These features

include, for example, the use of "well" as editing marker of a reformulation and the high frequency of "and" and "so" in sentence-initial position, and in the summary paraphrase which are associated with unplanned spoken language.

Chapter five deals with the second textual feature analysed in this study, namely sub-topic shift signalling. It first identifies the phenomenon and establishes three types of STSS: they are the statement, the metatextual comment and the change of structure. It then moves on to examine the textual features and grammatical realization of STSS. Three textual features are dealt with: (1) the use and frequency of occurrence of connectives as initiators of STSS; (2) the role of STSS in the structure of the text, which is discussed in relation to the summary paraphrase and its role in text structuring; (3) features of the interaction between the text and the visual element in the STSS utterance. These features are also discussed in relation to the unplanned spoken versus planned written dichotomy.

The analysis of the grammatical realization of STSS concentrates on three aspects. Firstly, the new-information carrier as the central grammatical constituent of the signalling utterance is described in terms of its types and functions. Secondly, the use of personal pronouns in STSS is dealt with. Thirdly, aspects of grammatical complexity in the signalling utterance are examined. The use of personal pronouns and aspects of grammatical complexity are handled in the light of the main issue of this work, namely the relationship between biology lecture discourse and the unplanned spoken versus planned written language distinction. The analysis of aspects of grammatical complexity is applied to the signalling utterance in order to assess the role of STSS, with respect to this feature, in qualifying the lectures under study as being nearer to one pole or the other of the unplanned spoken/planned written language continuum.

Finally, features typical of unplanned spoken language such as syntactic and lexical revision, reiteration of part of the signalling utterance; and features typical of planned written language such as the use of STSS as a manifestation of the explicit planning of the text, are discussed.

In Chapter six, the third and last textual feature investigated here, namely paraphrase, is dealt with. Paraphrase is first identified and classified into types. thus, the summary paraphrase, the equivalence paraphrase and the inference paraphrase are established as types of paraphrase used in this data. The summary paraphrase is further divided into general and partial SP; the equivalence paraphrase, into partial and complete equivalence paraphrase. Paraphrase is then discussed in relation to the structure of the text. It is found that the summary paraphrase is a marker of sequence boundaries, while the equivalence and the inference paraphrase are features of the internal structure of the sequence. The next textual feature of the EP and IP dealt with, is the dependence of these 2 types of paraphrase on their co-text. It is pointed out that the choice of particular equivalent terms, for example in an EP, is motivated by the co-text where this type of paraphrase belongs. Finally, paraphrase as a marker of textual continuity is also illustrated.

Features of unplanned spoken as opposed to planned written language in paraphrase are discussed. Features of the interaction between the text and the non-verbal element in paraphrase are identified. Further some features of paraphrase which have been studied as part of the description of the phenomenon are assessed in relation to the polarity. Thus, the frequent use of connective ":so" to introduce SPs and IPs; the use of ellipsis in the SP, are interpreted as being more expected in

unplanned spoken discourse - the investigation of grammatical complexity in paraphrase confirms the findings of Chapter Three where grammatical complexity in the whole data was examined. Finally typical features of unplanned spoken language as realized in the use of repetition; which is partly discussed as being instantiated in the occurrence of paraphrase itself, and as realized in the use of interrupted structures are identified.

Each chapter has so far studied a feature of the discourse of these lectures under investigation and examined its realisation in relation to features of unplanned spoken versus planned written language opposition. Each of the main features and their types are found in the three sets of lectures. However, idiosyncrasic differences between the three sets of lectures were also identified and had to be accounted for. Chapter Seven is a kind of net which tries to capture those idiosyncrasies and assesses whether inter-speaker differences in the use of a particular feature are significant by applying the Chi squared technique. It establishes that: (1) set (3) is singled out among the three sets in its use of zero degree grammatical complexity in the whole data (Chi squared $X^2 = 60.04$ with 2 degrees of freedom at the 1 per cent significance level) and its non use of personal pronoun "I"; (2) set (2) shows a marked preference for using the connectives "well" and "now" (Chi squared for "well" is $X^2=17.15$; for "now" $X^2 = 16.22$); (3) finally set (1) uses mostly the connective "so" ($X^2 = 47.96$), it also favours the textual structure SP + STSS ($X^2= 18.73$ and it shows the highest frequency of occurrence of SPs ($X^2= 39.2$). It is then suggested that these differences between the three sets of lectures are potentially explainable by three factors, namely personal style, discourse topic and structuring strategies.

8.2 Summary of findings

The main findings of this investigation are grouped under the sub-headings of "features of unplanned spoken language", and "features of planned written language" which constitute the central issue of this work, and "other features" which relate to secondary issues.

8.2.1 Features of unplanned spoken language

This investigation has established the following characteristics of biology lecture discourse.

1 Clear tendency to use simpler sentences; one degree grammatical complexity sentences are favoured.

2 Use of finite clauses more than non-finite clauses.

3 Recurrent use of the connective "and" in sentence initial position; "so" as paraphrase initiator.

4 Use of connective "well" editing marker of reformulation and response initiator following a rhetorical question; use of connective "and" as additive and resumptive.

5 Use of syntactic/lexical revision, interrupted structures, digressions in the realisation of STSS and paraphrase, in particular the summary paraphrase which also shows a recurrent use of ellipsis.

6 Use of deictic expressions in STSS and in paraphrase to refer to the visual element in the text.

7 Use of paraphrase as a type of repetition and in particular the frequent use of the summary paraphrase which introduces redundancy in the information conveyed in the text.

8.2.2. Features of planned written language

Biology lecture discourse also shows:

1 A dominant use of subordination as opposed to co-ordination.

2 Favoured use of complement clause (that clause).

3 Readiness of the text to be syntactically analysed into sentence like units, i.e. full grammatical sentences.

4 Use of sub-topic shift signalling as a device for explicitly signalling the organisation of the content of the text.

5 Use of the statement (or topic sentence) as the most recurrent type of STSS utterance.

It should be emphasized that most of these features are not exclusive to the spoken or the written mode. However, they are more expected in one mode or the other.

8.2.3 Other features

This study has also established the following features of biology lecture discourse which relate to the secondary issues of inter-speaker variation, text structure and the pedagogical function of the variety.

1 Variation between speakers, although acknowledged is minimal for two main reasons: (1) it is not prominent in the use of the features investigated. Differences between speakers only appear in certain aspects of the textual realisation of a given feature (such as the frequency of occurrence of connective "so" in set (1)); (2) variation is not solely attributable to personal style; discourse topic and structuring strategies can also be motivating factors for variation.

2 Connectives are found both at the level of the macrostructure of the text (as STSS and summary paraphrase initiators), and within a sequence.

3 STSS and the SP are sequence boundary markers. They are features of the macrostructure of the text. The EP and the IP are features of the internal structure of the sequence.

4 Features of spoken/written language are used for pedagogical purposes (such as the frequent use of summary paraphrase and the use of sub-topic shift signalling).

8.3 Theoretical implications

This section discusses the theoretical implications which can be drawn from the evidence obtained in the present investigation. These implications relate firstly to the central issue of the dimension of mode in this variety; secondly to the issue of inter-speaker variation; thirdly to the issue of the function of the variety, and finally to the question of text structure.

8.3.1 Variety and mode

As we have seen in 2.3 Gregory and Carroll (1978) suggest that mode in a variety can be approached as a combination of both speech and writing

> "the text which has been written may be written to be spoken as if not written, ... or written to be spoken"
>
> (1978: 42)

That is speech and writing combine in different ways in describing a variety. Gregory and Carroll suggest the use of the modifier "plus/minus spontaneity" (ibid: 40) to describe mode in a variety according to the amount of planning it undergoes before it is delivered. The implication is that a text which is least

spontaneous would show more features of writing than speech; and a text which is most spontaneous would show more features of speech than writing. Further spontaneous discourse versus non-spontaneous discourse form the ends of a "cline" (ibid: 40) on which varieties are to be placed according to their degree of planning.

As stated in 8.1, the central issue of the present work was to identify biology lecture discourse along the unplanned spoken versus planned written language continuum. The results of this investigation indicate that this variety does not show any clear cut tendency towards one pole or the other of the continuum. Neither features of unplanned spoken discourse nor features of planned written discourse are used exclusively or decisively prevail in biology lecture discourse. In fact, it rather combines both types of features. Whether at sentence level, in features of grammatical complexity, or at discourse level in the use of sub-topic shifting, paraphrase and sentence-initial connectives, the evidence of this investigation is that features of both speech and writing are found in the same utterance.

If we maintain the polarity and the idea of a continuum then the conclusion to be drawn from these findings is that there are contradictory forces pulling biology lecture discourse towards both ends of the continuum. This is a conclusion we arrived at at the end of the analysis of grammatical complexity in Chapter three. Such a conclusion has motivated us to take the investigation a step further, to the level of discourse with the aim of gaining more insights into the speech/writing dichotomy. But, upon the investigation of the three textual features, biology lecture discourse has been found not to belong more to spoken language or more to written language. Both types of features co exist in this variety.

What has emerged from the findings of the present investigation is that the polarity of speech versus writing is not very useful except in the demarcation of

registers (see Ure & Ellis 1977). The approach to mode in a variety, as suggested by Gregory and Carroll as a potential combination of speech and writing is not helpful either because it maintains the polarity. There is a need to go beyond this polarity and to approach the investigation of mode from a different theoretical perspective.

The present work shows that aspects of mode are found at the level of textual features such as connectives, sub-topic shift signalling and paraphrase. This will have two theoretical implications. Firstly, the investigation of mode related aspects of a variety, should necessarily be undertaken at the inter-sentential level of discourse, an idea which is found in Kroll (1977) and Chafe (1980;1982) who went to the notion of "idea unit" as the unit of analysis in their investigation of features of speech versus writing in selected samples of discourse. The present study shows that the syntactic description of sentences will not give us the evidence we need in our endeavour to identify mode related aspects in biology lecture discourse. In fact, such a description has only been helpful in revealing certain major trends in this variety, such as the use of subordination as opposed to co-ordination and the use of low degree complexity structures. (See Chapter Three). But the study demonstrates that invaluable insights into the speech/writing or mode related aspects of biology lectures have been gained by taking the analysis beyond the level of the sentence. Secondly, we should ultimately look for those mode related aspects of a variety within this narrow scope of a feature (or an utterance) rather than a general description which labels a given variety by using a combination of "spoken" and "written" and some qualification such as spontaneous or planned. This low level analysis, (which is the "act" in Sinclair 1975, and "move" in Swales 1981) typical of discourse analysis, takes the utterance as the domain for the investigation of mode related aspects of a variety.

In this line, this study is a contribution to the elaboration on what aspects of mode are found in the features investigated in these biology lectures (cf Murray 1988)[1].

On the other hand, the evidence indicates that these mode related aspects are primarily, but not exclusively, carrying out the interactive function of language suggested by Sinclair (1981). (See also Berry 1981). This is seen in the modality of some utterances which instantiate a sub-topic shift signalling such as the injussive imperative clause "let us consider pine..." (333139), or the question "have you been introduced to a gentleman....." (12875). this interactive function of language is also realised in the use of personal pronouns, in a summary paraphrase "so what I'm trying to say to you is that ...." (129419).

There are also textual and metatextual elements which are primarily addressed to the listeners (thus interactive) and have to do with text making. Here we find that sub-topic shift signalling and the summary paraphrase carry out a textual function as they are organisational in that they signal a stage in the text. Deictics "now", "and", and "so" also have a textual function when they are resumptive and when they are text boundary markers as initiators of a STSS or SP. On the other hand we find that in the summary paraphrase or sub-topic shift signalling there are metatextual elements, such as the above quoted part of a summary paraphrase "so what I'm trying to say to you", which announces what is to follow.

This characteristic of the utterances investigated here, to carry out more than one function lends support to Sinclair's observation that the two major functions of language (namely the interactive and the autonomous) are realised in one and the same utterance. In one utterance we find elements which primarily function textually/metatextually or interactively. All these functions are subsumed by the

interactive function. Hence a relationship between all the features we have studied and one macro function of language can be established, namely the interactive function.

8.3.2 Variety and inter-speaker variation

We have seen in this study that singularity as manifested in individual occasional features is present in biology lecture discourse. There are some differences between speakers in the use of a given feature. However, we have also seen that factors other than individual style can motivate variation between speakers. Thus discourse topic can explain the prominent use of summary paraphrase in set one; structuring strategies can explain the frequent use of "now" in set two as a sub-topic shift signalling initiator or as deictic to signal return to main discourse. Further, as observed earlier, variation is occasional in that it is only observed in a few aspects of the features investigated. Comparability between the three sets of lectures is the dominant feature. This is seen in the fact that the three features investigated here are found in the three sets of lectures. It is also seen in aspects of grammatical complexity drawn from the analysis of the whole data and which are found to be common to the nine lectures investigated; namely that subordination is overwhelmingly more frequent than co-ordination; that finite clauses are more used than non finite clauses and that the two most frequent types of embedded clauses are found to be dominant in the three sets of lectures. This comparability or homogeneity in the lectures is interpretable in relation to an important characteristic of the language of biology lectures, namely that it is restricted language, that is, it is specialized scientific language. By virtue of its being specialized language, it has achieved a high level of standardization.

Standardization means institutionalized situations which co-relate with certain linguistic features. Thus standardization shows an advanced stage in development in both social institutions and their linguistic manifestations. A standard variety such as scientific papers, lectures, news bulletin, will have predictable features of lexicon, syntactic structures, and text structuring. Standardization is manifested in scientific language in the existence of a fixed repertoire of terminologies. That is to say the lexicon or inventory of technical terms or concepts used to refer to the phenomenon is established by conventions. Standardization is also manifested at the level of grammar in some features such as the use of simple present (see Lackstrom et al 1970 ; 1973) or the choice of types of clause such as the relative clause which has been found to be the most common type of clause in scientific English (Swales 1971: 54).

The implications of these characteristics of specialised restricted language is that they leave little room for individual optional variation. Scientific language is less favourable to variation than general language. As Milroy and Milroy observe "strictly speaking standardization does not tolerate variability" (1985: 22). Standardization of scientific language implies constraints on the lexical and syntactic choices open to the users of the language. In other words, standardization implies that variation, if allowed, is limited. Homogeneity is thus a normal feature of scientific language.

8.3.3 Variety and purpose

It has been observed earlier (2.6.3.2; 4.3.2) that the lecture is an institutionalized type of language event with a fixed setting and well defined participant roles. The audience and the speaker are of different academic status. The audience is a group learners, the speaker is the teacher. He/she has the

authority to impart knowledge. Audience and speaker perform different roles. The role of the audience is to listen and assimilate information, the role of the speaker is to hold the floor all through the event and transmit information. Further, audience and speaker have different purposes. One is to learn, the other is to teach. (Ure & Ellis 1977: 200).

On the other hand, the purpose of the event as a whole is pedagogical; however, it is socially determined in that its purpose is defined from the "community standpoint" (Dell Hymes 1972).

We have seen in this study that the use of the textual features investigated as well as aspects of the realization of these features are partially explainable in the light of the purpose of the activity. Thus the explicit signalling of the organisation of the information content of the lecture as manifested in the recurrent use of sub-topic shift signalling; the use of the equivalence and the inference paraphrase as a device to achieve clarity and explicitness; the frequent use of the summary paraphrase to help the audience keep track of what has preceded in the text; the speaker's use of the textual connective "well" in some rhetorical patterns to monitor his speech, are all interpretable in terms of the "didactic function" (Ure & Ellis 1977: 200) of the lecture. As we have argued in the discussion of these features (5.8; 6.1.1; 4.5.2.1) mode related aspects are exploited for pedagogical purposes. Thus the linguistic features of the lectures investigated are relatable to both mode and purpose of the variety. This lends support to Ure & Ellis' observations that the situational dimensions of a variety "may have an effect on the language" (ibid: 202). Thus 'purpose' as one of the situational dimensions has a linguistic manifestation in the variety.

This is different from Swales' definition of "purpose", which is posited as the essential factor in his consideration of "genre", excluding all other factors (such as mode). We may add that Swales main concern is to identify those features in language which would help in teaching English for specific purposes.

So purpose is seen in the "setting" (cf Dell Hymes 1972) of the event and in linguistic features such as paraphrase as repetition, which ensure that the purpose of the activity i.e. the "transmission of information" is being carried out.

8.3.4 Text structure and the present study

The present investigation has shown that sub-topic shift signalling and the summary paraphrase are text boundary markers. They signal the beginning and the end of a sequence (see 5.4.2). It has also shown that the connectives "now", "but " and sometimes "and" initiate sub-topic shift signalling; the connective "so" initiates the summary paraphrase. In such uses, these connectives are also text boundary markers.

We know that there is a strict sense of the notion of text structure as suggested in Sinclair and Coulthard (1975) where the same criteria applied to sentence grammar are maintained and where text structure is described in terms of a model in which every level or unit of the structure is accounted for. The description presented here does not take this strict sense of text structure. The main concern here was to identify the role of the features analysed in demarcating text boundaries.

If we adopt a "loose" sense of text structure, the question which remains to be answered, however, is how to define the boundaries of the main unit of structure i.e. the text. As previously discussed (6.1.1) the unity of text is primarily founded on extra-linguistic factors. The text belongs to a language

event, the lecture. Each lecture is a separate text. However, if we look at the linguistic realization of the text, we also find formal signalling or indications of the lecture boundaries as text. The beginning and end of the lecture are indicated by forms like "okay. we may well start". (33211); "and I'll continue on that theme next Friday" (129552).

If we look at the structure of the text, we find that it can be described in terms of sequences which are either major sequences or minor sequences (see Fig. 2). The beginning of the sequence whether it is a major sequence or a minor sequence is almost invariably signalled by STSS. The end of the sequence is mostly signalled by a SP. Where a SP does not appear, the occurrence of STSS indicates retrospectively the end of the preceding sequence. Thus STSS and the SP are features of the macrostructure of the text. The internal structuring of the text, on the other hand, is partly revealed by the use of the other two types of paraphrase dealt with in this study, namely the equivalence and the inference paraphrase.

At both levels of the structure of the text, we find sentence-initial connectives. They initiate STSS as they can initiate any type of paraphrase. The connective "so" introduces most instances of a summary paraphrase, it also introduces every instance of inference paraphrase. In addition, the deictic use of "and" (namely when it is resumptive) is also a marker of a textual relation within the sequence.

A major characteristic of the three features studied here is that they have a predictive power (see Sinclair 1980; Tadros 1982; Al-Shabab 1986). They indicate what is to come in the text. Sub-topic shifting, when it introduces a major sequence (as seen in "so can we therefore in the face of their diversity of structure

say anything about their general properties"(129225), is nearest to one of Tadros' types of predictive structure, namely what she has labeled "enumeration" (Tadros: ibid: 143). It commits the speaker to deal with more than one aspect of the sub-topic. However, our description of text structure is not exhaustive. The utterances investigated here are not discussed in relation to all the other utterances in the sequence. We do not go into the details of studying every utterance in the text and its position in the structure as Sinclair & Coulthard (1975) have suggested. Nor do we study all the utterances in a part of the text as Swales (1981) has done. Nevertheless, as far as the macrostructure of the text is concerned, the features investigated here have enabled us to establish two kinds of sequence: a major sequence and a minor sequence. In other words, they have enabled us to establish two components in the structure of the text realizing a hierarchical relation: a minor sequence is established by virtue of its being a component of a major sequence. The textual features studied here thus contribute to our understanding of the structure of the text in biology lecture discourse.

8.4 Potential applications

The experimental work of testing the role of the textual features investigated here in learners' comprehension of biology lectures is not the concern of this study. Therefore, the following suggestions for the implementation of the present findings in the construction of teaching materials for L2 (and also L1) learners are tentative.

There are three broad areas of potential applications of the results of the present investigation: (1) listening comprehension: (2) note taking: and (3) teacher training programming.

Listening comprehension and note taking have been identified as major problem areas encountered by students in the processing of lecture discourse,

(James 1975). Although this observation was made with L2 learners in mind, it can also apply to L1 learners. As Murphy &Candlin remark, in relation to note taking: "it should not be assumed that because their language control is fluent they need no help in note-taking" (1979: 70).

1 Listening comprehension

Listening comprehension exercises could be set up to focus on developing the students' awareness of the structure of biology lectures. Students could be trained to identify features of the overall organisation of the text such as sub-topic shift signalling and the summary paraphrase which highlight the major information blocks of the lecture. Indeed, as previously mentioned in 5.8, some experimental work has been carried out to test the effect of "discourse markers" on comprehension. Chaudron and Richards (1986) have analysed the role of what they have called "micro-markers" and "macro-markers" in L2 comprehension of history lecture discourse (see 2.1.2). Within their framework, our STSS would be nearer to the sub category of "macro-markers". The results of their experiment suggest that "macro-markers" enhance students' comprehension of the lecture.

A type of exercise which would be set up is to get the students to listen to a recorded authentic lecture and ask them questions which would incite them to identify those parts of the text which announce a shift of sub-topic. As we are dealing with a type of lecture which uses slides and diagrams,the visual element should be used in the exercise as it is complementary to the text (see Murphy &Candlin 1979). In the same orientation, students could be trained to listen for the summary paraphrase, i.e. for those parts of the text where the main points are stated again.

A further aspect of text structure and content organisation of biology lectures is revealed in the use of textual connectives. One could help the students to develop an awareness of the functional meaning of connectives in different contexts.

2 Note taking

Listening comprehension exercises should pave the way for the training of students in note taking. Indeed, students are generally required to perform both tasks at the same time, so the two skills go hand in hand.

The types of exercises suggested above should get the students to be sensitive to signals of main points in the text as realised in sub-topic shift signalling and the summary paraphrase. Note-taking exercises concentrating on those features, in particular the summary paraphrase, would reinforce their ability to identify the main points in the text and extract information. Students could be trained to reproduce the main points using their own words.

One could also draw students' attention to certain features typical of spoken language, such as interrupting structures, hesitation, syntactic/lexical revision which introduce redundancy in the text. Students can be taught how to identify these features and discard them. They would rather learn to exploit them to complete their notes of the main text.

3 Teacher training programme

The results of the present investigation could also be used in a teacher training programme. This would have as a primary objective to bring to

teachers' awareness the role of the features studied here, in potentially facilitating students' comprehension.

We have seen that the features we have investigated are interpretable in the light of the purpose of the activity. That is they are interpreted as being motivated by the pedagogical purpose of the activity. We suggest that they should be highlighted in the making of a didactic text such as the lecture.

One could probably also argue that a text which shows typical features of spoken discourse such as repetition, digression and hesitation phenomena would be easier to follow than a text which lacks such features. For one thing, the presence of such features relieves the load of information content that students have to assimilate. Further, a lecture which avoids any redundancy would hardly be tolerable. Therefore, one could draw trainee teachers' attention to the part that features of redundancy may have in lecturing. We suggest that they should not be discouraged to use them.

8.5 Suggestions for further research

The present investigation has attempted to explore aspects of the dimension of mode in selected features of biology lecture discourse. It has revealed what mode related aspects are found in the features studied in this variety. However, the present work has also touched on certain issues which have not been fully pursued. In this light, the following recommendations for further research can be made:

1 The features studied here, could be investigated in a written academic text. Such an investigation would show whether sub-topic shift signalling, paraphrase or textual connectives are used; to what extent they are used; and what grammatical realization they may have. This will

provide more data that can be compared with the results obtained in this work.

2 Text structure in biology lectures could also be a topic for an independent work. Montgomery's investigations of science lectures (1977) has identified elements which can be used in the study of text structure. The present work has also identified two sequences of structuring biology lecture discourse. However, there is a need for this aspect of science lecture discourse to be thoroughly investigated.

3 It would also be interesting to study further the dimension of singularity in this or other varieties. Using a bigger population of lecturers, idiosyncrasies in the use of a textual feature would be more accurately assessed and a more significant indication would be given of the part of individual variation in a given variety.

4 Finally, we can suggest that a replicate of this investigation could be carried out in other varieties or text types. Such studies would contribute to our understanding of the part that mode has to play in the identification and description of varieties. They would also give us an indication of what aspects of speech and writing tend to co-exist in a given variety. This would provide further insight into this dimension of variety description.

Footnotes

(1) Murray (1988) who has investigated "mode/medium" switching in computer language, comes also to the conclusion that the approach to speech versus writing as dichotomous, and the idea of a continuum should be discarded. However, she suggests that "medium" should be investigated in terms of some aspects of the "context of situation", namely setting and audience versus non-audience.

BIBLIOGRAPHY

AKINNASO, F.N., (1982) : "On the Differences between Spoken and Written Language" in Language and Speech, Vol. 25/2. (97-125).

AL-SHABAB, O.S., (1986) : Organisational and Textual Structuring of Radio News Discourse in English and Arabic. Unpublished Ph.D. thesis, University of Aston in Birmingham.

AUSTIN, J.L., (1962) : How to do Things with Words, Oxford : Oxford University Press.

BARBER, C.J., (1962) : "Some Measurable Characteristics of Modern Scientific Prose" in Contributions to English Syntax and Phonology, Stockholm.

BEAMAN, K., (1984) : "Co-ordination and Subordination Revisited : Syntactic Complexity in Spoken and Written Narrative Discourse" in Tannen, D. (Ed.) Coherence in Spoken and Written Discourse, Norwood, New Jersey: Ablex, (45-80).

BERRY, M., (1975) : An Introduction to Systemic Linguistics. 1. Structures and Systems, London : Batsford.

(1977) : An Introduction to Systemic Linguistics. 2. Levels and Links, London : Batsford.

(1979) : "A Note on Sinclair and Coulthard's Classes of Acts including a Comment on Comments", in Nottingham Linguistic Circular, 8, (49-59).

(1981) : "Polarity, Ellipticity, Elicitation and Propositional Development, Their Relevance to the Well-Formedness of an Exchange", in Nottingham Linguistic Circular, 10 (36-63).

(1981) : "Systemic Linguistics and Discourse Analysis: A Multi-Layered Approach to Exchange Structure", in Coulthard, H. & Montgomery, M. (Eds.), Studies in Discourse Analysis, London : Routledge & Kegan Paul, (120-45).

(1982) : "Review of Halliday 1978", in Nottingham Linguistic Circular, 11, (64-94).

BLIGH, D., (1972) : What's the Use of Lectures. London : Penguin.

BOLINGER, D., (1977) : Meaning and Form, London : Longman Group Limited

BREMOND, C., (1970) : "Morphology of the French Folktale" in Semiotica, 2, (247-276)

BROWN, G., (1972) : "Learning about the Structure of Lectures" in Lecturing and Explaining. London : Methuen

(1977) : Listening to Spoken English, London : Longman.

BROWN, G. & YULE, G., (1983) : Discourse Analysis, Cambridge : Cambridge University Press.

BROWN, G., (1989) : "Making Sense : The Interaction of Linguistic Expression and Contextual Information" in Applied Linguistics, Vol. 10/1, (97-108).

BROWN, R. & BELLUGI, V., (1964) : "Three processes in the child's acquisition of Syntax" in Harvard Educational Review, 34, (133-157).

BURTON, D., (1980) : Dialogue and Discourse, London : Routledge & Kegan Paul

BURTON, D., (1981) : "Analysing Spoken Discourse" in Coulthard, M. & Montgomery, M., (Eds.) Studies in Discourse Analysis, London : Routledge & Kegan Paul, (61-81).

CANDLIN, C.N., BURTON, J. & COLEMAN, H., (1980) : Dentist-Patient Communication. A Report to the General Dental Council, Lancaster : University of Lancaster

CARRELL, P.L., (1983) : "Some Issues in Studying the Role of Schemata, or Background Knowledge in Second Language Comprehension" in Reading in a Foreign Language, Vol. 1/2, (81-92).

CHAFE, W., (1970) : Meaning and the Structure of Language, Chicago and London : The University of Chicago Press.

(1979) : "The Flow of Thought and the Flow of Language" in Givon, T., (Ed.) Syntax & Semantics Vol. 12 : Discourse & Syntax, Academic Press Inc., (159-180).

(1980) : "The Deployment of Consciousness in the production of a Narrative" in Chafe, W, (ed) The Pear Stories : Cognitive, Cultural, and Linguistic Aspects of Narrative Production, Norwood, NJ, (9-50)

(1982) : "Integration and Involvement in Speaking, Writing and Oral Literature" in Tannen, D. (Ed.) Oral and Written Discourse : Exploring Orality and Literacy, Norwood, N.J., (35-53).

(1985) : "Linguistic Differences Produced by Differences between Speaking and Writing" in Olson, D., Torrance, N. & Hildyard, A., (Eds.) Literacy, Language and Learning. The Nature and Consequences of Reading and Writing, Cambridge : Cambridge University Press, (105-123).

CHAUDRON, R., (1983) : "Simplification of Input : Topic Reinstatements and their Effects on L2 Learners' Recognition and Recall", TESOL Quarterly, 17/3, (437-58).

CHAUDRON, C. & RICHARDS, J.C., (1986) : "The effects of Discourse Markers on the Comprehension of Lectures", Applied Linguistics Vol. 7/2, (113-127).

CHERRY, C., (1957) : On Human Communication, Massachussetts M.I.T. Press. (Reprinted New York : Science Editions, 1959).

CHRISTENSEN, F., (1965) : A Generative Rhetoric of the Paragraph" in College Composition and Communication, 16/3. (144-156).

CICOUREL, A., (1980) : "Three Models of Discourse Analysis" in Discourse Processes, Vol. 3, (101-133).

COOK, S., (1975) : A Communicative Approach to the Analysis of Extended Monologue Discourse, Unpublished M.Litt. thesis, Edinburgh University.

COOK-GUMPERZ, J. & GUMPERZ, J., (1981) : "From Oral to Written Culture : The Transition to Literacy", in Whitehead, H.F. (Ed.) Writing : The Nature, Development and Teaching of Written Communication, Vol. 1 Variation in Writing: Functional and Linguistic : Cultural Differences. Hillsdale : N.J. Erlbaum (89-109).

COULTHARD, H., (1971) : An Introduction to Discourse Analysis, London : Longman.

COULTHARD, H. & MONTGOMERY, H., (1981) : Studies in Discourse Analysis, London : Routledge & Kegan Paul.

CRYSTAL, D. & DAVY, D., (1969) : Investigating English Style, London : Longman.

DARIAN, S., (1983) : "Spoken and Written Communication: their Forms and Functions", ITL Review 62, (27-51)

DAVIDSON, O., (1968) : "On Saying That", Synthese, 19 (130-144); and in Martinich, A.P. (Ed.), The Philosophy of Language, New York, Oxford : Oxford University Press (1985) (339-348).

DE BEAUGRANDE, R., (1980) : Text, Discourse and Process. Towards a Multidisciplinary Science of Texts, London : Longman.

DE BEAUGRANDE, R. & DRESSLER, W., (1981) : Introduction to Text Linguistics. London : Longman. German edition (1972).

DE VITO, J.A., (1966) : "Psychogrammatical factors in Oral and Written Discourse by Skilled Communicators" in Speech Monographs, 33, (73-76).

DEYES, A.F., (1975) : Notes and Discussion : Towards a Linguistic Definition of Functional Varieties of Written English IRAL, Vol. 16/4, (313-329).

DRESSLER, W.U., (Ed.), (1978) : Current Trends in Text Linguistics, Berlin & New York : Walter de Gruyter.

DUDLEY-EVANS, A. & JOHNS, T., (1979) : A Team Teaching Approach to Lecture Comprehension for Overseas Students in ELT Documents Special : The Teaching of Listening Comprehension. The British Council.

EDMONSON, W., (1981) : Spoken Discourse : A Model for Analysis, London : Longman.

ELLIS, J., (1966) : "On Contextual Meaning" in Bazell, C.E. & al (Eds.) In Memory of J.R.Firth, London : Longman.

& URE, J., (1974) : The Contrastive Analysis of Language Registers. Unpublished draft.

(1976) : "The Role of the Concept of Text in the Elaboration of Linguistic Data" in York Papers in Linguistics.

FALK, J.S., (1978) : Linguistics and Language. A Survey of Basic Concepts and Implications, New York : John Wiley & Sons.

FARAG, S.M., (1986) : A Linguistic Analysis of Spoken and Written Narrative Discourse, Unpublished Ph.D. thesis. The University of Aston in Birmingham.

FILLMORE, G.J., (1971) : "Towards a Theory of Deixis" The PCCLL4 Papers. Department of Linguisitics, University of Hawaii.

FIRTH, J.R., (1957) : Papers in Linguistics (1934-1951), Oxford : Oxford University Press.

FREI, H., (1944) : "Systemes de Deictiques" in Acta Linguistica, Vol. 4, (111-129)

GIVON, T., (1983) : "Topic Continuity in Discourse : An Introduction" in Givon, T. (Ed.) Topic Continuity in Discourse : A Quantitative Cross Language Study, Amsterdam, Philadelphia : John Benjamins Publishing Company.

GOFFMAN, E., (1981) : Forms of Talk, Oxford : Basil Blackwell.

GOLUB, L.S., (1969) : "Linguistic Structures in Students' Oral and Written Discourse" in Research in the Teaching of English, 3, (70-85).

GOODY, J., (1977) : The Domestication of the Savage Mind, Cambridge : Cambridge University Press.

GOPNIK, H., (1972) : Linguistic Structures in Scientific Texts, Janua Linguarum Seri Minor, 129, The Hague : Mouton.

GRAVER, B.C., (1971, 1963) : Advanced English Practice, London : Oxford University Press.

GREGORY, M., (1967) : "Aspects of Varieties Differentiation" in Journal of Linguistics, 3, (177-198)

GREGORY, M. & CARROLL, S., (1978) : Language and Situation : Language Varieties and their Social Context. London : Routledge & Kegan Paul.

GRICE, H.P., (1975) : "Logic and Conversation" in Cole, P. & Morgan, J.L. (Eds.) Syntax and Semantics. Vol. 3 : Speech Acts, New York : Academic Press.

GRIMES, J., (1975) : The Thread of Discourse, The Hague : Mouton Publishers.

GRUNER, C.R. & al, (1967) : "A Quantitative Analysis of Selected Characteristics of Oral and Written Vocabularies" in Journal of Communication, 17, (152-158).

GUMPERZ, J.J. & HYMES, D., (Eds.), (1964) : The Ethnography of Communication, Special Issue of American Anthropologist, Vol. 66, 16, Part 2.

GUMPERZ, J.J., (1982) : Discourse Strategies : Studies in Interactional Sociolinguistics, Vol. 1, Cambridge : Cambridge University Press.

HALLIDAY, M.A.K., (1961) : "Categories of the Theory of Grammar" in Word, 17/3, (241-292).

HALLIDAY, M.A.K., McINTOSH, A. & STREVENS, P.D., (1964) : The Linguistic Sciences and Language Teaching, London : Longman.

HALLIDAY, M.A.K., (1967-1968) : "Notes on Transitivity and Theme in English" Parts 1-3 in Journal of Linguistics (1967) : (37-81); (1967) : (199-244); (1968) : (179-215).

HALLIDAY, M.A.K., (1973) : Exploration in the Functions of Language, London: Arnold.

HALLIDAY, M.A.K. & HASAN, R., (1976) : Cohesion in English, London : Longman Group Limited.

HALLIDAY, M.A.K., (1978) : Language as Social Semiotics : The Social Interpretation of Language and Meaning, London : Edward Arnold.

(1979) : "Differences between Spoken and Written Language : Some Implications for Literacy Teaching", in Page, G., Elkins, J. & O'Connor, B. (Eds.) Communication through Reading : Proceedings of the fourth Australian Reading Conference, Adelaide, S.A. : Australian Reading Association, Vol. 2, (37-52).

(1985) : An Introduction to Functional Grammar, London : Edward Arnold.

HARRIS, Z.S., (1952) : "Discourse Analysis" in Language, 28, (1-30).

HARTMAN, R.R.K., (1977) : Contrastive Textology, Heidelberg: Julius Groos Verlag.

HOBBS, J.R., (1982) : "Towards an Understanding of Coherence in Discourse" in Wendy, G.L. & Martin, H.R. (Eds.) Strategies for Natural Language Processing, London, New Jersey : Lawrence Elbaum Association Publishers.

HOCKETT, C.F., (1985) : A Course in Modern Linguistics, New York : Macmillan.

HOEY, H., (1979) : "Signalling in Discourse", E.L.R. Mimeograph, Birmingham Instant Print Limited.

(1983) : "On the Surface of Discourse, London : George Allen & Unwin.

(1983) : "A Tentative Map of Discourse Studies and their Place in Linguistics" in Quaderni di Linguistica, Vol. 1, (7-20).

HOROWITZ, H. & NEWMAN, J., (1964) : "Spoken and Written Expression : an Experimental Analysis" in Journal of Abnormal and Social Psychology, 68/6, (640-647).

HUDDLESTON, R.D. & al., (1968) : Sentence and Clause in Scientific English, Mimeograph, Department of General Linguistics, University College, London University.

HUDDLESTON, R.D., (1971) : The Sentence in Written English: A Syntactic Study Based on an Analysis of Scientific Texts, Cambridge : Cambridge University Press.

HUDDLESTON, R.D, (1984) : Introduction to the Grammar of English, Cambridge : Cambridge University Press.

HUDSON, R., (1980) : Sociolinguistics, Cambridge, London : Cambridge University Press.

HUTCHINS, J., (1977) : "On the Structure of Scientific Texts", Journal of Linguistics, 5, (18-39).

HYMES, D., (1964) : "Towards Ethnographies of Communication : the Analysis of Communicative Events". Excerpts from Hymes, D., Introduction : Toward Ethnographies of Communication, American Anthropologist, Vol. 66/6, part 2, (12-25).

HYMES, D., (1972) : "Models of the Interaction of Language and Social Life" in Gumperz, J.J. & Hymes, D., (Eds.), Directions in Sociolinguistics: the Ethnography of Communication, New York : Holt, Reinhaut & Winston, Inc., (35-71).

JAMES,K., (1975): "Note-taking in lectures: Problems & Strategies", in English for Academic Purposes, BAAL/SELMOUS.

JOHNS-LEWIS, C., (1985) : "Prosodic Differentiation of Discourse Modes" in Johns-Lewis, C., (Ed.), Intonation in Discourse, London : Croom Helm, (199-219).

JOHNS-LEWIS, C., (1987), : "Linguistic Differentiation of Spoken and Written Language : Implications for Language Teaching and Learning", in Ager, D.E. & Townson, H. (Eds.), The Written Skills in the Undergraduate Curriculum in Foreign Languages. CILT/AMLC

JOHNS-LEWIS, C. & SKELTON, J., (1987) : "Variation in Discourse", Paper presented at conference on Variation in Second Language Acquisition, English Language Institute, Michigan University, Oct. 9-11 (1987).

JOOS, M., (1961) : The Five Clocks, New York & London : A Harvest / HBJ Book, Harcourt Brace Jovonovich.

KEENAN, E.D. & SCHIEFFELIN, B.B., (1976) : "Topic as a Discourse Notion : a Study of Topic in the Conversation of Children and Adults", in Li, C., (Ed.), Subject and Topic, New York : Academic Press, (337-385).

KELCH, K., (1985) : "Modified Input as an Aid to Comprehension" in Studies in Second Language Acquisition, 7/1, (81-9).

KEMPSON, R.M., (1977) : Semantic Theory, Cambridge, London : Cambridge University Press.

KING, P., (1987) : Spoken and Written Science and Engineering Text. Unpublished paper, E.L.R. Birmingham University.

KIRSNER, R.S., (1979) : "Deixis in Discourse : An Exploratory Quantitative Study of the Modern Dutch Demonstrative Adjectives" in Givon T., (Ed.), Syntax and Semantics Vol. 12 : Discourse and Syntax, Academic Press Inc., (355-373).

KROLL, B., (1977) : "Ways Communicators Encode Propositions in Spoken and Written English " A Look at Subordination and Co-ordination" in Keenan, E.O. & Bennett, T., (Eds.), Discourse across Time and Space, Southern California Occasional Papers in Linguistics, 5, (Los Angeles), (69-108).

KUNO, S., (1975) : "Three Perspectives in the Functional Approach to Syntax". in Grossman, R.E., San, L.J. & Vance, T.J. (Eds.), Papers From the Parasession on Functionalism, Chicago : The Chicago Linguistics Society.

LABOV, W., (1972) : Language in the Inner City, Pennsylvania : University of Pennsylvania Press.

LACKSTROM, J., SELINKER, L. & TRIMBLE, L., (1970) : "Grammar and Technical English" in Lugton, R., (Ed.), English as a Second Language : Current Issues, Philadelphia, p.a., The Center of Curriculum Development Inc., (101-133), also in English Teaching Forum (1972), 10/5, (3-14).

(1973) : "Technical Rhetoric Principles and Grammatical Choice", in English Teaching Forum, June-August 1973, (1-7).

LACKSTROM, J., (1977) : "The Comprehension of EST Arguments and Definitions", in Language Learning, (47-66).

LAKOFF, R., (1972) : "Language in Context", in Language, Vol. 48, 4, (907-927).

(1973) : "Questionable Answers and Answerable Questions", in Kachru, B. et al (Eds.), Papers in Linguistics in Honour of Henry and Renee Kahane, Urbana, IL : University of Illinois, (453-67).

(1979) : Expository Writing and the Oral Dyad as Points on a Communicative Continuum : Writing anxiety as the result of mistranslation. Unpublished manuscript.

LEECH, G.N., (1969) : Towards a Semantic Description of English, London : Longman.

(1971) : Meaning and the English Verb, London : Longman.

(1974) : Semantics, Harmondsworth : Penguin.

(1983) : Principles of Pragmatics. London : Longman.

LEVINSON, S.C., (1979) : "Activity Type and Language" in Linguistics, 17, (365-399). The Hague : Mouton Publishers.

(1983) : Pragmatics, Cambridge : Cambridge University Press.

LILLEY, A.D., (1976) : Towards the Analysis of a Communicative Situation : the Science Lecture. Mimeograph. University of Essex.

LONG, M.H., (1985) : "Input and Second Language Acquisition Theory" in Gass, S.M. & Madden, C., (Eds.), Input and Second Language Acquisition, Rowley, H.A. : Newbury House.

LONGACRE, R.E., (1974): "Narrative Versus Other Discourse Genre" in Brend, R.M, (ed) Advances in Tagmemics, Amsterdam: North Holland, (357-376).

(1982) : "Discourse Typology in Relation to Language Typology" in Allen, S., (Ed.), Text Processing : Text Analysis and Generation, Text Typology and Attribution, Proceedings of Nobel Symposium 51 (1982), Almovist & Wiksell International.

(1983) : The Grammar of Discourse, New York and London : Plenum Press.

LOTMAN, J.H., (1977) : "Problems in the Typology of Texts", in Lucid, D.P., (Ed.), Soviet Semiotics : An Anthology, Baltimore : The John Hopkins University Press (119-124).

LYONS, J., (1968) : Introduction to Theoretical Linguistics, London & New York : Cambridge University Press.

(1977) : Semantics. Vol. I & II, Cambridge : Cambridge University Press.

(1981) : Language, Meaning and Context, Bungay, Suffolk, England : Fontana Paperback.

MALINOWSKY, B., (1930) : "The Problem of Meaning in Primitive Language", in the second edition of Ogden, C.K. & Richards, I.A., The Meaning of Meaning, London : Routledge & Kegan Paul.

MAYNARD, D.W., (1980) : "Placement of Topic Changes in Conversation" in Semiotica, 30, (263-90).

MILROY, J. & MILROY, L., (1985) : Authority in Language. Investigating Language Prescriptivism and Standardization, London, New York : Routledge & Kegan Paul.

MINSKY, M., (1980) : "A Framework for Representing Knowledge", in Metzing, D., (Ed.), Frame Conceptions and Text Understanding, Berlin, New York : Walter de Gruyter.

MONTGOMERY, M.M., (1977) : Some Aspects of Discourse Structure and Cohesion in Selected Science Lectures, Unpublished M.A. Dissertation. University of Birmingham.

MURPHY, D.F. & CANDLIN, C.N., (1979) : "Engineering Lecture Discourse and Listening Comprehension", in Practical Papers in English Language Education, Vol. 2, Lancaster : University of Lancaster.

MURRAY, D.E., (1988) : "The Context of Oral and Written Language : A Framework for Mode and Medium Switching", in Language Society, 17, (351-373), Cambridge University Press.

MYERS, G., (1989) : "The Pragmatics of Politeness in Scientific Articles" in Applied Linguistics, Vol. 10/1, (1-35).

OCHS, E., (1979) : Planned and Unplanned Discourse, in Givon, T., (Ed.), Syntax and Semantics, Vol. 12 : Discourse and Syntax, Academic Press Inc.

O'DONNOLL, R.C., (1974) : "Syntactic Differences between Speech and Writing", in American Speech, 49, (102-110).

OLSON, D.R., (1977) : "From Utterance to Text : The Bias of Language in Speech and Writing", in Harvard Educational Review, 47, (257-281).

ONG, W.J., (1982) : Orality and Literacy, London and New York : Methuen.

OWEN, M., (1981) : Conversational Units and the use of "Well..", in Werth, P., (Ed.), Conversation and Discourse. Structure and Interpretation, London : Croom Helm (99-116).

PALMER, F.R., (Ed.), (1968) : Selected Papers of J.R.Firth 1952-59, London : Longman & Bloomington Ind : Indiana University Press.

(1976) : Semantics:A New Outline, Cambridge : Cambridge University Press.

PARTEE, B., (1973) : "The Syntax and Semantics of Quotation", in Anderson, R. & Kiparski (Eds.), A Festchrift for Morris Halle, New York : Holt, Reinhart and Winston, Inc. (410-418).

PHILLIPS, H., (1985) : Aspects of Text Structure. An Investigation of the Lexical Organisation of Text, Amsterdam, New York, Oxford : North Holland.

POOLE, H.E. & FIELD, T.W., (1976) : "A Comparison of Oral and Written Code Elaboration" in Language and Speech, 19, (305-312).

QUIRK, R. et al, (1972) : A Grammar of Contemporary English, London : Longman Group Limited.

(1985) : A Comprehensive Grammar of the English Language, London : Longman.

ROBINS, R.H., (1964) : General Linguistics : An Introductory Survey, London : Longman.

ROE, P., (1977) : Scientific Text, Discourse Analysis Monograph 4 E.L.R. The University of Birmingham

RUBIN, A., (1980) : "A Theoretical Taxonomy of the Differences between Oral and Written Language", in Spiro, R.J. et al, (Eds.), Theoretical Issues in Reading Comprehension : Perspectives from Cognitive Psychology, Linguistics, Artificial Intelligence and Education, Hillsdale, N.J., (411-437).

RUMELHART, D., (1984) : "Understanding Understanding", in Flood, J., (Ed.), Understanding Reading Comprehension : Cognition, Language and the Structure of Prose, U.S.A. International Reading Association, (1-20).

(1977) : "Understanding and Summarising Brief Stories", in Laberge,D. & Scemuels, J. (Eds.), Basic Processes in Reading Perception and Comprehension, Hillsdale, N.J. : Lawrence Erlbaum Associates.

SACKS, H., SCHEGLOFF, E.H. & JEFFERSON, G., (1973) : "A Simplest Systematics for the Organisation of Turn-taking for Conversation", in Language, 50, (695-735).

SAMPSON, G., (1980) : Making Sense, Oxford & New York : Oxford University Press.

SAVILLE-TROIKE, H., (1982) : The Ethnography of Communication : an Introduction, Oxford : Basil Blackwell.

SAWA, G.H., (1985) : Modified Input as an Aid to Comprehension, Manuscript, University of Hawaii at Manoa.

SCHANK, R.C., (1977) : Rules and Topics in Conversation, in Cognitive Science, 1, (421-42).

SCHEGLOFF, E.A., (1968) : "Sequencing in Conversational Openings", in American Anthropologist, 70, (1075-95).

SCHEGLOFF, E.A. & SACKS, H., (1973) : "Opening up Closing", in Semiotica, 8, (289-327).

SCHIFFRIN, D., (1987) : Discourse Markers, Studies in Interactional Sociolinguistics, 5, Cambridge University Press.

SEARLE, J., (1969) : Speech Acts, Cambridge : Cambridge University Press.

(1975) : "Indirect Speech Acts" in Cole, P. & Morgan, J.L., (Eds.), Syntax and Semantics 3 : Speech Acts, New York : Academic Press, (59-82).

SELINKER, L., (1979) : "On the Use of Informants in Discourse Analysis", IRAL, 27/3 (189-215).

SINCLAIR, J.M., (1980) : "Some Implications of Discourse Analysis for ESP Methodology" in Applied Linguistics, Vol. 1/3, (253-261).

(1981) : "Planes of Discourse", First Draft Mimeograph Birmingham University. Published in Rizvi, S.N.A., (Ed.), The Twofold Voice : Essays in Honour of Rakesh Mohan. Saltzburg Studies in English Literature, University of Salzburg, (1983).

SINCLAIR, J.M. & COULTHARD, R.M., (1975) : Towards an Analysis of Discourse : The English Used by Teachers and Pupils, London : Oxford University Press.

STRANG, B.H.H., (1962) : Modern English Structure, London : Edward Arnold, 2nd edition, (1968).

STUBBS, H., (1983) : Discourse Analysis : The Sociolinguistic Analysis of Natural Language, Oxford : Basil Blackwell.

SVARTVIK, J., (1980) : "Well in Conversation", in Greenbaum, S., Leech, G. & Svartvik, J., (Eds.), Studies in English Linguistics, London : Longman Group Limited, (167-177).

SWALES, J.M., (1974) : "Notes on the Function of Attributive en-Participles in Scientific Discourse", in Papers for English for Special Purposes No.1 ELSU : University of Khartoum.

(1981) : Aspects of Article Introductions, Aston ESP Research Report No.1, Language Studies Unit, University of Aston in Birmingham.

SWALES, J.M., (1986) : "Citation Analysis and Discourse Analysis", in Applied Linguistics, Vol. 7/1, (39-56).

TADROS, A.A., (1981) : Linguistic Prediction in Economic Texts, Unpublished Ph.D thesis, the University of Birmingham.

TANNEN, D., (1982) : "Oral and Literate Strategies in Spoken and Written Narrative" in Language, 53/1, (1-21).

(1982) : "The Oral / Literate Continuum in Spoken Discourse", in Tannen, D., (Ed.), Spoken and Written Language Exploring Orality and Literacy, Norwood, New Jersey : Ablex, (1-16).

TANNEN, D., (Ed.), (1982) : Spoken and Written Language : Exploring Orality and Literacy, Norwood, New Jersey : Ablex.

(Ed.), (1984) : Coherence in Spoken and Written Discourse : Advances in Discourse, Norwood, New Jersey : Ablex.

(1984) : "Spoken and Written Narrative in English and Greek" in Tannen, D., (Ed.), Coherence in Spoken and Written Discourse : Vol. 2, Advances in Discourse Processes, Norwood, New Jersey : Ablex (21-41).

(1985) : "Relative Focus on Involvement in Oral and Written Discourse in Olson, D.R., Torrance, N. & Hildyard, A., (Eds.), Literacy, Language and Learning, Cambridge : Cambridge University Press, (124-147).

ULLMAN, S., (1962) : Semantics : An Introduction to the Science of Meaning, Oxford : Basil Blackwell.

URE, J. & ELLIS, J., (1977) : "Register in Descriptive Linguistics and Linguistic Sociology" in Uribe-Villegas, O., (Ed.), Issues in Sociolinguistics, The Hague, Paris, New York : Mouton, (197-243).

URQUHART, A.H., (1981) : "Operating on Learning Texts" in Selinker, L. et al, (Eds.), English for Academic and Technical Purposes. Studies in Honour of Louis Trimble, Rowley, Massachussetts : Newbury House Inc.

(1982) : "Paraphrase in Discourse", in MALS Journal, Vol. 7, Spring 1982.

VAN DIJK, T.A., (1972) : Some Aspects of Text Grammars. The Hague : Mouton.

(1977) : Text and Context : Exploration in the Semantics and Pragmatics of Discourse, London : Longman.

WERTH, P., (1981) : "The Concept of Relevance in Conversational Analysis" in Werth, P., (Ed.), Conversation and Discourse, Structure and Interpretation, London : Croom Helm, (129-153).

WIDDOWSON, H.G., (1973) : An Applied Linguistic Approach to Discourse Analysis. Unpublished Ph.D. thesis. University of Edinburgh.

(1978) : Teaching Language as Communication, Oxford : Oxford University Press.

WIDDOWSON, H.G., (1980) : "Conceptual and Communicative Functions in Written Discourse" in Applied Linguistics, Vol. 12/3, (234-243).

WIJASURIYA, B., (1971) : The Occurrence of Discourse Markers and Inter-Sentence Connectives in University Lectures and their place in the Testing and Teaching of Listening Comprehension in English as a Foreign Language. MEd thesis, University of Manchester.

WILSON, D. & SPERBER, D., (1981) : "On Grice's Theory of Conversation" in Werth, P., (Ed.), Conversation and Discourse, Structure and Interpretation, London : Croom-Helm, (155-179).

WINTER, E.O., (1974) : Replacement as a Function of Repetition: A Study of Some of its Principal Features in the Clause Relations of Contemporary English. Unpublished Ph/D. thesis, University of London.

(1977) : "A Clause Relational Approach to English Texts : A Study of Some Predictive Lexical Items in Written Discourse", in Instructional Science, 6, (1-92).

(1978) : Fundamentals of Information Structure, Hatfield School of Humanities. Hatfield Polytechnic, 3rd edition.

(1982) : Towards a Contextual Grammar of English, London, Herts : George Allen & Unwin Publisheres Ltd.

YOUNG, D., (1985) : The Structure of English Clauses, London : Hutchinson.

Aston University

Pageă removed for copyright restrictions.

Appendix (2) (Appendix to Chapter Three)

1. Table (1) : Frequency indices for number of clauses joined by co-ordination in the same sentence (calculated per 1,000 words)

Text	No of Words	Number of co-ordinate clauses per sentence		
		2	3	4
Set 1 3 lectures	14,200	(105) 7.39	(21) 1.47	(8) 0.05
Set 2 3 lectures	14,200	(83) 5.84	(16) 0.07	(3) 0.02
Set 3 3 lectures	14,600	(115) 7.87	(19) 0.06	(4) 0.02
Total	43,000	(303) 7.04	(56) 1.3	(15) 0.34

Table (2) : Frequency indices for co-ordinate clauses in compound and complex sentences (calculated per 1,000 words)

Texts	No of words	F.I. for Co-ord clauses in compound sentence	F.I. for Co-ord clauses in complex sentence
Set 1 (3 lectures)	14,200	(37) 2.60	(97) 6.38
Set 2 (3 lectures)	14,200	(31) 2.18	(71) 5
Set 3 (3 lectures)	14,600	(51) 3.49	(87) 5.95
Total	43,000	(119) 2.76	(255) 5.93

2. Examples and discussion of types of subordinate/embedded clauses which are not dealt with in the main text.

1 - Nominal clauses

(i) - Nominal relative clauses:

They are always introduced by a WH-element. They can normally be paraphrased by a noun phrase containing a postmodifying relative clause. In the following excerpt from the data:

> //and if you look at/ where the sporangia are located/ you find that...../
>
> (333216)

The nominal relative clause "where the sporangia are located can be paraphrased as "the place where the sporangia are located".

(ii) - WH interrogative clauses:

They are also always introduced by a WH-element. "They resemble WH question in that they leave a gap of unknown information, represented by the WH element" (Quirk et al 1972 : 736). The following excerpt from the data exemplifies a case of WH interrogative clause (underlined).

> //but we don't have to rely only/ on looking at this/ to see/ what a ginkgo forest was like...../
>
> (33317)

(iii) - to-infinitive clauses:

They are non-finite clauses. In the following example:

> //and it is usual/ to remark that the blood concentrations are controlled by some form of negative feedback.//
>
> (129439)

to-infinitive clause is used as adjectival complement (it is complement of the adjective "usual").

(iv) - ing-clause:

They are also non-finite clauses. They do not include ing-clauses of verbal progressive constructions like "now those carbon atoms are coming from the carbohydrate of the host plant" (330429). The following example is an instance of an ing-clause found in the data : (underlined the the excerpt).

> //and they will stimulate the enzyme responsible/ for converting things like lactic acid and glycerol into glucose in the liver//
>
> (129439)

The ing-clause is used here as a complement of the preposition for

2 - Adverbial clauses:

Adverbial clauses found in the data include clauses like:

(a) //the erm the phloem sieve tubes are not the are not erm not conspicuous channels for the conduction of organic nitrogen during the autumn/ when leaves fall//

(231854)

(b) //and if growth hormone from that self same anterior pituitary region is produced in excess prior to puberty/ you can get this enormous growth of the long bones and enormous increase in stature//

(12908)

In (a) the adverbial clause is introduced by the subordinator "when". It is used as a temporal clause. In (b) it is an if-conditional clause.

Appendix (3) (Appendix to Chapter Six)

Edited transcript of a structured interview with specialist informants, namely the authors of the text, Dr. Perris and Dr. Armstrong. (Dr. Hughes was no longer a member of staff at Aston University at the time when the interview was conducted. Dr. Armstrong, who took over Dr. Hughes' part of the course on Plant Biology, kindly accepted to deal with Dr. Hughes' texts).

Extracts of paraphrase were given to the authors of the texts to be verified as cases of either equivalence or inference paraphrase, as identified by the criteria set up in the analysis. (see overleaf).

Reference in the corpus	Original utterance	Paraphrastic form	Reaction of specialist informant
(129510)	it (PTH) provokes the dissolution of bone mineral. which is a calcium phosphate complex called apatite.	so it stimulates the breakdown of bone material, releasing calcium into the blood.	inference paraphrase
(129527)	so when plasma calcium concentrations go up. not only does PTH. come down, but catonin levels. go up in the blood. calcitonin has its target. in the bone where it. inhibits the dissolution of bone mineral.	so calcium goes up. cells in the thyroid release calcitonin. react. on the bone. inhibit the dissolution of bone mineral.	equivalence paraphrase
(129540)	now the enzyme which catalyses the second hydroxilation, not surprisingly is called hydroxylase. and is activated by. low. plasma calcium concentrations.	so when calcium concentrations go down. we get an enhanced production of 1,25. diahydroxy-cholical ciferol..	inference paraphrase
(129832)	and since, specificity is at the heart, of this recognition process.. you might. deduce. or suspect. that. the	The proteins. are extremely adept. at forming. specific configurations in space. into which. a specific hormone. shape. can. come to	equivalence paraphrase

Reference in the corpus	Original utterance	Paraphrastic form	Reaction of specialist informant
	recognition mechanism has got something to do with the shapes. of the molecules.. and therefore. the receptors. should be endowed. with chemical properties. which would. enable them. to discriminate on the basis. of the shapes of. the hormones.. and the best molecules. for taking up specific configurations in space. with an appropriate orientation. are proteins..	lie. or interact..	
(130041)	We've just been discussing. a situation. in which a hormone interacts with a receptor at the cell surface. and that. tends to be the mechanism. via which. protein hormones exert their effects. they are large molecules. too big. to get in the cells. into the cells normally.	so they must exert their effect at the cell surface. proteins are large they are not readily lipid soluble. and so they must exert their effect. at the level of plasma membrane..	equivalence paraphrase

Reference in the corpus	Original utterance	Paraphrastic form	Reaction of specialist informant
(130052)	but they (steroids) do exhibit. specificity. and they interact. with. receptors. which are located. in the cytoplasm. of. the cells..	so we get. within the cytoplasm. a steroid. receptor complex..	inference paraphrase
(230343)	but in addition. to protecting the bacterium. the haemoglobin is a source. of. oxygen. at low partial pressures...	so oxygen is available. both to the bacterium. and to the host plant. at low partial pressure.	equivalence paraphrase
(231131)	it proved by electron microscopy it proved by. examination. that there are cytoplasmic connections between the infected cells. and the uninfected cells of the cortex. there are cytoplasmic connections across the endodermis. and into the pericycle.	so there is nothing to impede. the flow of organic nitrogen. from the infected cells. to the pericycle. in the simplast	inference paraphrase
(231710)	and er. and water would be attracted to the sieve tube element. er. in by osmosis.	and therefore the sieve tube element will become inflated with water..	inference paraphrase

Reference in the corpus	Original utterance	Paraphrastic form	Reaction of specialist informant
(332212)	and we can interpret this fern frond. as a stage in the development of the megaphyl. in which the webbing. has not completely joined. all. the lateral axes.	You are looking at a branch system then. with some webbing on it. but that webbing has not become complete. has not. formed. er completely.	equivalence paraphrase
	you are looking at a branch system then. with some webbing on it. but that webbing has not become complete. has not. formed. er completely.	The ferns. are stuck if you like. in an evolutionary. intermediate stage.	This was a suspected instance of inference paraphrase. It was disconfirmed by the author of the text on the grounds that it requires the audience to have a solid knowledge of the subject. The utterance conveys new information. Therefore it is not a case of paraphrase.
(333336)	today. the vegetation of the earth generally is not dominated by conifers. they are important. in their zone the boreal forest. which is a single area in the. northern hemisphere and in the southern hemisphere.	but most of the rest of the world. is dominated by quite different plants than the gymnosperms	equivalence paraphrase

Reference in the corpus	Original utterance	Paraphrastic form	Reaction of Specialist Informant
(333017)	and if we take a female cone remember large. 90 pound structure. and remove a single cone scale. we see something like this. we see. a portion of the tip of the cone scale. which seems to represent. the reduction of a leaf.	so. the female cone scale probably evolved from a leaf like structure.	inference paraphrase
(332644)	unfortunately you've got to be very careful. about the use of both these compounds. they are very toxic. and there are many other organisms as well. which may suffer. from their application.	so you have to chance getting rid of the bracken, with the possibility. of side effects.	inference paraphrase
(333437)	and a third. extremely common pattern. is called palmately. reticulate.	the leaf is lobed. in this sort of fashion. a single vein enters the leaf. branches. and the branches. are connected. together by. a complex. network. of veins..	equivalence paraphrase