

Bangor University

DOCTOR OF PHILOSOPHY

Lexical cohesion and the determination of text intelligibility by non-native readers of English

Bensemmane, Faiza

Award date:
1985

Awarding institution:
Bangor University

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LEXICAL COHESION AND THE DETERMINATION OF TEXT
INTELLIGIBILITY BY NON-NATIVE READERS OF ENGLISH

Thesis submitted by

Faïza Bensemmane

For the degree of Doctor of Philosophy in the Department
of Linguistics, University College of North Wales, Bangor.

October 1985

BEST COPY

AVAILABLE

Variable print quality

To M'hamed, Selma and Radia

To my parents

ACKNOWLEDGEMENTS

I am indebted to several persons who have provided me with encouragement and advice in preparing this thesis: to Mr. Phil Scholfield, my supervisor, who deserves my deepest gratitude for his precious guidance, advice and encouragement; to Dr. Carl James for his valuable suggestions on various drafts of the thesis; to Professor Andrew Radford, Head of the Department, for his continuous cooperative attitude and support, and to whom I express my sincere thanks, and to Professor David Wilkins for being my external examiner.

I also wish to express my gratitude to the Ministry of Higher Education in Algeria who have made this study possible by awarding me a scholarship, and to the University of Algiers for granting me a leave of absence from my teaching post. The practical part of this research was carried out in the Institut des Langues Etrangères, University of Algiers, and thanks are due to the staff and students of the Department for their cooperation during the field work.

Finally, I owe that special thanks to my husband who has supported me unfailingly throughout this research and to my two wonderful daughters. I owe them an apology if I have not always been attentive to their needs.

SUMMARY

This thesis attempts to describe lexical relationships between sentences in text and between utterances in discourse in the light of pragmatics and psycholinguistics. It was inspired by Halliday and Hasan's pioneering effort to describe relationships of cohesion in text but it goes beyond their taxonomy of lexical cohesion to include pragmatic aspects that can serve the purposes of FL reading research

Thus, the motivation of the present research is two-fold:

- i To describe lexical cohesion as a "competence" phenomenon by determining the factors contributing to its achievement in text.
- ii To provide an account of lexical cohesion as a "performance" phenomenon by investigating the factors affecting its interpretation in FL reading comprehension.

The articulation of the thesis reflects these two issues: the first Chapter is a linguistic account of lexical cohesion in English. It lays out the basis on which cohesion and lexical cohesion should be distinguished from coherence and lexical coherence, and reviews the literature which has treated these phenomena in relation to the theoretical framework adopted for the present study. Thus, two main categories are proposed for the analysis of lexical cohesion as a competence and performance phenomenon: lexicosemantic cohesion which accounts for the connectedness of "text" and lexicopragmatic coherence which is a feature of the connectedness of "discourse". The Chapter also provides detailed analysis of the lexical devices of cohesion and coherence in the light of theories of semantics and pragmatics.

Chapter Two examines the involvement of the cognitive factor in the analysis of lexical cohesion. It deals with the concept of "background" or "schematic" knowledge viewed as an essential component of the reading process and investigates the role of top-down and bottom-up processes in the making of linguistic and pragmatic inferences specifically when unknown vocabulary items are encountered in reading (comprehension). In Chapter Three an experimental investigation of the linguistic and non-linguistic features of the interpretation of lexical cohesion in reading comprehension is proposed. It seeks to inquire into the processing of lexicoreferential, lexicosubstitutional and "conjunctive" relationships of cohesion and coherence by non-native readers of English and attempts to answer the following three research questions:

- 1 How do FL learners utilise lexical resources or links of linguistic cohesion and pragmatic coherence when inferencing unknown meaning while reading?
- 2 How does the use of lexical resources of cohesion and coherence vary as a function of FL proficiency?
- 3 How does the use of lexical resources of cohesion and coherence vary as a function of language background?

Four experiments were designed to that effect and null-hypotheses were formulated to test the performance of subjects on a cloze test in four types of independent variables. The findings which reveal that the use of lexical resources of cohesion and coherence were a function of language proficiency rather than language background, bear some pedagogical and other implications which are discussed in the final Chapter.

TABLE OF CONTENTS

	<u>Page</u>
Acknowledgements	i
Summary	ii
1 Lexical cohesion and lexical coherence: theoretical background	
Introductory notes	1
1.1 Cohesion, coherence, text and discourse	3
1.1.1 Langue - competence versus parole - performance	3
1.1.2 Cohesion - text and coherence - discourse	4
1.1.3 Text - semantics and discourse - pragmatics underlying cohesion and coherence	5
1.1.4 "Rules of usage" and "rules of use"	6
1.2 On defining lexicosemantic cohesion and lexicopragmatic coherence	8
1.2.1 Some issues involved in lexical pragmatics	8
1.2.2 Cohesion links and coherence links from the writer's viewpoint	11
1.2.3 Definitional and non-definitional relations of meaning in lexical cohesion/coherence	12
1.3 Review of literature on cohesion and lexical cohesion	16
1.4 Lexical versus grammatical cohesion: similarities	20
1.5 Presupposition and implicature in lexical cohesion/coherence	23
1.5.1 Presupposition	23
1.5.2 Implicature	25
1.5.3 Leech's principles of "Textual Rhetoric": how they can explain LS cohesion and LP coherence	30

	<u>Page</u>	
1.6	An analysis of the relations underlying lexical cohesion/coherence	38
1.6.1	The referential relation in lexical cohesion/coherence	38
1.6.1.1	Anaphora	40
1.6.1.2	Cataphora	41
1.6.2	The substitutional relation in lexical cohesion/coherence	43
1.6.3	The 'conjunctive' relation in lexical cohesion/coherence	44
1.7	Semantic and pragmatic resources of lexical cohesion/coherence	50
1.7.1	Coreferential repetition of lexical meaning	51
1.7.1	Non-coreferential repetition of lexical meaning	54
1.7.3	The expression of 'special' linguistic acts/speech acts through lexical reiteration	56
1.7.4	Synonymy as a cohesive category in lexical reiteration	58
1.7.5	Hyponymy as a carrier of lexical cohesion	63
1.7.5.1	Hyponymy in the system	63
1.7.5.2	Hyponymy in use	64
1.7.5.3	Hyperonym verbs in discourse	71
1.7.5.4	Cohyponyms in discourse	75
1.7.6	Paronymy as a device of lexical cohesion	76
1.7.7	Antonymy and converseness as devices of lexical cohesion	78
1.7.7.1	Antonymy	78
1.7.7.2	Converseness	80
1.8	Functional dynamism in lexical cohesion	81
1.9	Concluding remarks	89

	<u>Page</u>
Notes on Chapter 1	91
2 Lexical cohesion in reading comprehension	
Introductory notes	94
2.1 Reading in NL and reading in FL	97
2.1.1 Top-down and bottom-up reading	97
2.1.2 Reading in NL versus reading in FL	100
2.2 A discussion of the notion of 'background knowledge' in relation to lexical cohesion	103
2.3 Content and formal schemata	105
2.3.1 Frames	107
2.3.2 Scripts	111
2.3.3 Scenarios	118
2.3.4 Summary	122
2.4 The processing of lexical relations of cohesion and coherence by native and non-native speakers readers of English	128
2.4.1 Linguistic and pragmatic inferences	128
2.4.2 Linguistic and pragmatic inferences and how they relate to lexical cohesion: literature review	130
2.4.2.1 Linguistic and pragmatic inferences involved in the processing of lexicoreferential cohesion and coherence	137
2.4.2.1.1 Effect of order of referential hyponyms in reading	145
2.4.2.1.2 Effect of distance between cohesive hyponyms in reading	146
2.4.2.1.3 Effect of background knowledge on the processing of lexicoreferential cohesion/coherence	148
2.4.2.2 Linguistic and pragmatic inferences involved in the processing of lexicoconjunctive cohesion	149

	<u>Page</u>
2.4.2.2.1 Evidence from research in NL	151
2.4.2.2.2 Evidence from research in SL/FL	152
2.5 Conclusion	154
Notes on Chapter 2	159
5 An experimental investigation of the awareness of lexical relationships of cohesion and coherence by non-native adult readers of English on an inferencing task	
Part One	162
3.1 The rationale for empirical research	163
3.2 Utilizing inferencing as an experimental task	165
3.3 The verbalization and cloze procedures	168
3.3.1 The verbalization procedure	168
3.3.2 The cloze procedure	169
3.3.3 Cloze procedure combined with verbalization in the Pilot Test	171
3.4 The experiments: Aim and nature and research questions	174
(i) The hypotheses	178
(ii) Some considerations on clozefication (or item deletion)	187
3.4.1 The subjects and the Placement Test	188
3.4.1.1 The subjects	188
3.4.1.2 The Placement Test	188
3.4.2 Selection for experiments	190
3.5 The design	191
3.5.1 The variables	191
3.5.2 The procedure	192
3.5.3 The scoring method for the dependent variable	196
3.6 The experimental variables	199
3.6.1 The moderator variables	199
3.6.1.1 The yeartime independent variable	199
3.6.1.2 The language dominance independent variable	200

3.6.2	The materials	200
3.6.2.1	Selection	200
3.6.2.2	Text design for each experiment	201
3.6.2.2.1	Control of 'kind' of clue	201
3.6.2.2.2	Control of 'type' of text	205
3.6.2.2.3	Some considerations on text editing	205
3.7	The results	207
3.7.1	Experiment 1: 'Order of clue'	208
3.7.1.1	Comparing scores for the IV 'experimental condition'	208
3.7.1.2	Comparing scores for the 'yeartime' independent variable	210
3.7.1.3	Comparing scores for the language dominance independent variable	213
3.7.1.4	Interaction between 'yeartime' and 'language' (YEARTIM BY AF)	214
3.7.1.5	Interaction between 'yeartime' and 'condition' (YEARTIM BY CONDITIO)	216
3.7.1.6	Interaction between 'language' and 'condition' (AF BY CONDITIO)	218
3.7.2	Experiment 2: 'Distance of clue'	223
3.7.2.1	Comparing mean scores for experiment condition IV	224
3.7.2.2	Comparing scores for the 'yeartime' variable alone	224
3.7.2.3	Comparing the scores for the 'language dominance' IV alone	225
3.7.2.4	Interaction between 'year' and 'language'	226
3.7.2.5	Interaction between 'yeartime' and 'condition'	228
3.7.2.6	Interaction between 'language' and 'condition'	231
3.7.3	Experiment 3: 'Conjunctive clue'	234
3.7.3.1	Comparing mean scores for the 'experimental condition' IV	234
3.7.3.2	Comparing mean scores for the yeartime IV alone	235
3.7.3.3	Comparing scores for the 'language dominance' variable alone	236

	<u>Page</u>	
3.7.3.4	Interaction between 'yeartime' and 'language'	236
3.7.3.5	Interaction between 'yeartime' and 'condition'	238
3.7.3.6	Interaction between 'language' and 'condition'	242
3.7.4	Experiment 4: 'Linguistic clue'	244
3.7.4.1	Comparing scores for the 'experimental condition' IV	244
3.7.4.2	Comparing scores for the 'yeartime' IV	245
3.7.4.3	Comparing scores for the 'language dominance' IV	246
3.7.4.4	Interaction between 'yeartime' and 'language' IV	246
3.7.4.5	Interaction between 'year' and 'condition'	249
3.7.4.6	Interaction between 'language' and 'condition'	251
3.7.5	Concluding remarks	254
 Part Two		
3.8	Miscue Analysis	261
3.8.1	Preliminaries	261
3.8.2	Miscue research and analysis	262
3.8.2.1	Inadequate clue-searching strategy and inadequate response	262
3.8.2.2	Adequate clue-searching strategy but inadequate response	263
3.8.2.3	Reading without reasoning	264
3.8.2.4	Inadequate linguistic knowledge	266
(i)	Syntactic knowledge	266
(ii)	Semantic knowledge	266
(iii)	Discourse thematicity	267

	<u>Page</u>
(iv) Discourse functions	267
3.8.2.5 Inadequate pragmatic knowledge	268
3.9 Concluding remarks	270
Notes on Chapter 3	273
4 Pedagogic and other implications	277
4.0 Introduction	277
4.1 Implications for FL teaching: developing lexicocommunicative competence	280
4.1.1 Analysis of items involved in cloze experiments: a problem in recognizing discourse functions	280
4.1.2 Recognizing discourse functions and using them in reading and writing: on the role of discourse markers	283
4.1.3 Some problems on FL vocabulary teaching	287
4.1.3.1 Vocabulary development and vocabulary recognition	287
4.1.4 Toward 'communicative' vocabulary teaching	289

	<u>Page</u>	
4.1.5	Developing inferencing strategies as part of "receptive" communication strategies for reading comprehension	294
4.1.5.1	On "receptive" communication strategies	294
4.1.5.2	Improving inferencing strategies for reading comprehension: some practical exercises	298
4.2	Implications for applied linguistics research	304
4.3	Concluding remarks	306
Notes on Chapter 4		309
Appendices		
Appendix I	Transcription and pronunciation table	313
Appendix II	Texts	315
Appendix III	Sample of miscue inventory: "A" and "F" Expected and Observed Responses	334
Appendix IV	Students' data on cloze experiments	341
Appendix V	SPSS programme used to compute students' responses	342
Appendix VI	Results of statistical analysis	343
Appendix VII	Sample of students' verbal protocol on Pilot Test	347
Appendix VIII	Pilot Test	350
Appendix IX	Sample of A and F students' responses on Placement Test	352
Bibliography		353
List of Diagrams, Graphs and Tables		
Diagrams		
Diagram 1	Cohesion and coherence	15
Diagram 2	Lexical cohesion and lexical coherence	50

		<u>Page</u>
Diagram 3	<u>Flower</u> and its hyponyms	63
Diagram 4	<u>Get</u> and its hyponyms	71
Diagram 5	Structure of a restaurant script	112-14
Diagram 6	Thorndike's (1976) model of generation of inferences	132
Diagram 7	Model of generation of inferences with unknown vocabulary items	156
Diagram 8	Experimental Subjects	191
Diagram 9	Experimental Conditions	193
Diagram 10	Kind of clues	202
Graphs		
Graph 01	Experiment 1: year x language	215
Graph 02	Experiment 1: year x condition	217
Graph 03	Experiment 1: language x condition	219
Graph 04	Experiment 2: year x language	227
Graph 05	Experiment 2: year x condition	229
Graph 06	Experiment 2: language x condition	231
Graph 07	Experiment 3: year x language	237
Graph 08	Experiment 3: year x condition	239
Graph 09	Experiment 3: language x condition	243
Graph 010	Experiment 4: year x condition	247
Graph 011	Experiment 4: year x condition	249
Graph 012	Experiment 4: language x condition	251
Tables		
Table 01	The procedure	193
Table 02	Mean scores obtained by Y1, Y2, Y3 in A and F	214
Table 03	Mean scores obtained by Y1, Y2, Y3 in C1 and C2	216

		<u>Page</u>
Table 04	Mean scores obtained by A and F in C1 and C2	218
Table 05	Mean scores obtained by Y1, Y2, Y3 in A and F	226
Table 06	Mean scores obtained by Y1, Y2, Y3 in C1 and C2	228
Table 07	Mean scores obtained by A and F in C1 and C2	231
Table 08	Mean scores obtained by Y1, Y2, Y3 in A and F	237
Table 09	Mean scores obtained in C1, C2, C3 by Y1, Y2, Y3	238
Table 010	Mean scores obtained by A and F in C1, C2 and C3	242
Table 011	Mean scores obtained by Y1, Y2, Y3 in A and F	247
Table 012	Mean scores obtained by Y1, Y2, Y3 in C1 and C2	249
Table 013	Mean scores obtained by A and F in C1 and C2	251
Table 014	Summarising table	255

CHAPTER 1

LEXICAL COHESION AND LEXICAL COHERENCE: THEORETICAL

BACKGROUND

"However luxuriant the grammatical cohesion displayed by any piece of discourse, it will not form a text unless this is matched by cohesive patterning of a lexical kind" (Halliday and Hasan, Cohesion in English, 1976, p292).

Introductory notes

The extension of linguistics beyond the sentence has generated acute interest in the study of the connectedness of text and of its interpretation as discourse. The beginning of interest in cohesion theory which can account for the connectedness of text, dates back from the mid fifties with the works of the Structuralist Fries (1957), followed by the Transformationalist Harris (1970) who provided a linguistic account of connected "discourse". However, it is only recently that interest in cohesion theory has been noted among researchers in second and foreign language learning and instructing, and the desire to apply some of its principles to concrete situations has ever been increasing.

Systematic studies of textual cohesion have appeared in the last decade only, most of them following the publication of Halliday and Hasan's book on cohesion in English. However, although Halliday and Hasan's book is the most influential on the subject, as it gives an

exhaustive account of devices of cohesion in English, their theoretical model of cohesion does not account for the interpretation of cohesive devices in discourse as it does not involve the reader, and their treatment of lexical cohesion in particular suffers from this point of view. Since the advent of Halliday and Hasan's study which provided a competence model of cohesion, some empirical studies have emerged which have analysed the performance aspects of this phenomenon in reading and writing in the native language, most of them being developmental studies of children's acquisition of devices of cohesion in English as a native language (thereafter referred to as NL. Also foreign language will be subsequently referred to as FL). Those studies involving the native speakers were primarily concerned with grammatical factors of cohesion and little or no interest was shown for lexical factors of cohesion. This seems to follow the general tendency towards neglect of lexical analyses. Also, no study concerned with lexical factors of cohesion specifically can be noted which involved non-native speakers' performance. We believe that lexical factors of cohesion are an important aspect of textual cohesion and coherence which should be studied in their own right. They play a crucial role in text/discourse interpretation, whether the language user is a native or a non-native speaker of English, and with respect to the non-native speaker, lexical cohesion represents a valuable area of his (or her) acquisition of foreign language competence (in order

to avoid prolixity, I henceforth use the pronouns his, him, he in their generic senses to refer to any learner of either sex).

This chapter deals with the description of lexical cohesion as a competence, or langue phenomenon. It seeks to determine the lexical factors which contribute to its achievement. The treatment of lexical cohesion as a performance or (parole) phenomenon will be dealt with in the next chapters, where some of the aspects affecting its interpretation in reading as a FL will be studied empirically.

1.1.0 Cohesion, coherence, text and discourse

This section examines the concepts of text and discourse in relation to cohesion and coherence, and assesses their place in the analysis of cohesion and coherence. In order to bring light on this double dichotomy, it seems essential to recall two basic distinctions generally accepted among linguists today between langue (competence) and parole (performance).

1.1.1 Langue - competence versus parole - performance

The event of structuralist linguistics dating back from de Saussure has generated the distinction between two levels of linguistic analysis, la langue and la parole, la parole being the actualisation of la langue, an idealised system or code existing in the native speaker. Subsequently, Chomskyan linguistics made a distinction between two systems cognitively different in the native speaker, competence, the internalised system,

and performance, the realisation of this idealised code of language. Although *la langue* cannot be equated with competence, nor can *la parole* be equated with performance, they share some fundamental features and thus are used synonymously in this study. This dichotomy between language competence and language performance provides us with two different though complementary levels of analysis of language, the formal level and the functional, which ultimately enables us to draw a distinction between text and discourse, and cohesion and coherence, as is exemplified in the following section.

1.1.2 Cohesion - text and coherence - discourse

Cohesion belongs in the formal level, the level of text which is an element crucial to its definition. Text is characterised by the way sentences are organised into larger units. It exploits the sense relations between grammatical (grammatical is used with the sense of 'syntactic' unless stated otherwise) and lexical items ('lexical items' and 'vocabulary items' are used indiscriminately in this study). Thus, as a supra-sentential unit of language, it comprises sentences connected by formal devices of cohesion signalling two types of relationship between successive sentences, grammatical cohesion, achieved via syntactic devices, and lexical cohesion, produced via lexical devices. But the interpretation of text cohesion concerns the language user's linguistic competence or his knowledge of the linguistic system which includes "rules of usage" in Widdowson's (1978, 1979) sense, as well as his performance which includes "rules of use"?

Coherence belongs in the functional level of language, the level of discourse. Discourse is characterised by the way connected sentences function as utterances in communication, that is how they combine to produce coherence. Coherence is a precondition for the interpretation of sentences in use. It is a non-formal relation which deals with the user's knowledge of "rules of use" in Widdowson's sense, or his "communicative competence" in Hymes (1972) sense. Thus, coherence links are non-signalled in text. They are inferred by the reader and are a function of his "pragmatic" competence. Coherence is "measured by the extent to which a particular instance of language use corresponds to a shared knowledge of conventions as to how illocutionary acts are related to form larger units of discourse" (Widdowson, 1978: 45). A piece of language is coherent to a language user if he can recognise it as a description, a report, an explanation, that is how the utterances relate to each other and function as rhetorical devices. Once he recognises "the illocutionary significance of the relationship", he can create links between grammatical and lexical meanings of the text as a discourse. This distinction between cohesion as a feature of text, and coherence as a feature of discourse is necessarily generated by the separation between semantics and pragmatics which follows.

1.1.3 Text - semantics and discourse - pragmatics
underlying cohesion and coherence

The further distinction drawn between semantics and

pragmatics reflects the double dichotomy between text cohesion and discourse coherence just mentioned. The present separation of semantics and pragmatics is akin to Leech's (1983) view of these two phenomena as "distinct, though complementary and interrelated fields of study" (Leech, 1983: 6).¹ Semantics is concerned with the representation of the meaning of (text) sentences and is defined purely as a property of expressions in a given language. Pragmatics, also dealing with meaning, is defined relative to a user of the language, producer or receiver, and involves the interpretation of those expressions as discourse utterances. In cohesion, some links are produced by vocabulary items occurring in text. These links are linguistic-semantic. They involve definitional relations of meaning. In coherence, some links are established by the language user between vocabulary items in discourse. These links are non-linguistic. They include pragmatic, non-definitional meaning which may be more or less associated with specific vocabulary items. Appeal to pragmatics appears to be essential to this study for if approached from an entirely semantic point of view, the present account of cohesion would be an incomplete description of the phenomenon.

1.1.4 "Rules of usage" and "rules of use"

It seems essential to refer to "rules of usage" and to "rules of use" when attempting an analysis of cohesion. "Rules of usage" are those semantic (often called "logic" or "logicosemantic") rules which underlie cohesion. They

may be regarded as semantic implications on which one can base conclusions. "Rules of use" are those pragmatic rules which underlie coherence. They are made on the basis of pragmatic principles which mostly derive from Grice's (1975) "Co-operative Principle". (These principles are discussed in Section 1.5. below). For instance, if the initial sentence of text (1) below is analysed from a semantic logical point of view, some implications contained in the semantic system of English can be drawn, which cannot apply to the next sentence.

(1) a. John divorced Mary; b. He is at the Sorbonne
this year.

a. John divorced Mary

$(y, x) (x \text{ divorced } y) \rightarrow (x \text{ was married to } y)$

$(\text{Mary, John}) (\text{John divorced Mary})$ implies that
 $(\text{John was married to Mary})$

A semantic interpretation of sentence (1a) can be accounted for by its semantic deep structure, but semantics cannot account for the interpretation of sentence (1b). Sentence (1b) is warranted not by linguistic semantic knowledge but by pragmatic knowledge and therefore involves "rules of use". Then it is possible to envisage that (1b) may not be accessible to the reader who does not possess knowledge to the extent assumed by the writer, viz. knowledge that the Sorbonne is a university institution in Paris. Given that Sorbonne is a lexical item whose meaning cannot be deduced from its semantic constituent (+ Place) for instance, implied by 'place Adjunct "at", its meaning may not be interpreted at all unless

more contextual information is provided.

Meanwhile, the reader may be able to connect (1b) to (1a) via grammatical cohesion which involves rules of usage: he corefers to John; and also deduce that (1b) is consecutive to (1a) by examining the tense aspects of divorced and is and deictic this.

1.2. On defining lexicosemantic cohesion and lexico-pragmatic coherence

In the introductory notes of this chapter, it was emphasised that the aim of this study was to focus on the lexical aspect of (textual) cohesion and to examine its interpretation by a potential reader. In this section we propose to analyse lexical cohesion within the competence-performance framework, described in Section 1.0. Such a framework, which enabled us to draw a distinction between cohesion and coherence, provides us with a further separation between two lexical categories on which the discussion of the phenomenon of lexical cohesion will be based throughout the present study, viz. lexicosemantic cohesion and lexicopragmatic coherence. But before going into detailed account of these two notions, it seems important to present some of the issues involved in the distinction of a functional level of lexical pragmatics.

1.2.1 Some issues involved in lexical pragmatics

As noted earlier, the concepts of cohesion and coherence

imply the distinction of two levels of analysis, lexical semantics and lexical pragmatics. While 'system' factors of cohesion, including semantic ones, are identifiable by reference to a 'grammar' in Chomsky's sense for their analysis and categorisation, 'non-system', pragmatic factors of coherence are not readily identifiable precisely because they involve two different individuals, the producer and the receiver, each of them having his own "encyclopaedia" (Smith and Wilson, 1979: 173) composing his pragmatic competence. Although linguistic competence may also differ from individual to individual utilising the same language, the amount of similarities in linguistic competence remains more important than the amount of differences. But some differences in pragmatic competence between writer and reader may vary so considerably that they may impair communication. Thus, while the lexical semantics of a text are relatively easy to identify, the lexical pragmatics of a discourse may pose problems. There are essential difficulties in attempting to define lexical pragmatics which may not be easy to solve, and many questions are susceptible to be left unanswered: how can the receiver/reader be expected to share the producer/writer's world-knowledge and to make inferences in lexical matters? How far can the vocabulary be expected to supply linguistic and non-linguistic information to the receiver/reader? Where do defining linguistic relations between vocabulary items end and non-defining pragmatic ones begin? Is, then, what we call "pragmatic coherence" just one aspect of

cohesion? Van Dijk (1973: 72) raises the point:

"Is the fact that 'if something is summer, then something is hot' a fact of the semantic structure of the language for which this statement would be valid, or merely a representation of an empirical fact? Much depends on our conception of the lexicon, without which, apparently, no derivation can be serious."

For the text-linguist Petöfi (1978: 43) the lexicon now does not only contain the vocabulary defined for the text. The dichotomy is between "Text-Structure" and "World-Structure" (cf. his "Text West Theory").

Granted that the interpretation of connected sentences in discourse often poses problems that semantics cannot solve because 'grammar' in Chomskyan sense lacks the dimension of context-dependency, appropriateness, relevance and informativeness, all of which are vital to the interpretation of discourse, it is necessary that some problems relating to the interpretation of connected sentences be solved by pragmatics, and following the distinction between cohesion as a feature of text and coherence as a feature of discourse presented in Section 1.12 above, we propose that two main categories be distinguished to deal with this dichotomy, viz. lexicosemantic cohesion and lexicopragmatic coherence subsequently referred to as LS cohesion and LP coherence. These two categories can be defined as follows:

- a. lexicosemantic cohesion is achieved via the occurrence of linguistic semantic relationships encoded in lexical items and holding between various parts of the

text. These sense relations, as defined in Lyons (1977), include synonymy, hyponymy and antonymy. They are definitional and are always signalled on the surface text. Sense relations constitute the lexical semantics of the text and are decoded by the reader by reference to his linguistic/semantic competence.

- b. lexicopragmatic coherence is achieved via non-linguistic/pragmatic meaning relationships between lexical items. Thus, these non-system relations produce pragmatic, non-definitional meaning associated with some specific vocabulary items. Pragmatic relations of lexical coherence are non-signalled in the discourse and constitute its lexical pragmatics. They are implicated by the language producer and interpreted by the receiver by reference to his non-linguistic/pragmatic competence.

1.2.2 Cohesion links and coherence links from the writer's viewpoint

Coherence links are implicit in the discourse and have to be supplied by the receiver. They give the backbone of what the producer is saying. While these links are obligatory features of writing, unless the writer chooses otherwise, cohesive links are non-obligatory. Whether coherence of underlying content should be signalled on surface text as cohesion depends upon the writer's willingness to be explicit and non-ambiguous,

that is, to comply with Grice's (1975) Principle of Co-operation which translates into a reduction to a minimum of the reader's inferencing of missing propositions.

As a non-obligatory feature of writing, cohesion may be viewed as an option open to the writer, "a special case of coherence" (Szwedek, 1980). From the receiver's point of view, cohesion is an expected state which should reflect the internal coherence of a piece of language. It is the 'marked' aspect of coherence. Thus, when producing text, the writer generates a textual entity by supplying signals of connection on its surface decodable by a potential reader (by reference to his linguistic competence) and a discursal entity whose implicit connections are interpretable more probabilistically by the potential reader (by reference to his pragmatic competence). Thus, identifying links of coherence communicated via lexical pragmatics may be more problematic to the reader because semantic meaning underlying lexical cohesion, the receiver 'knows', but pragmatic meaning underlying lexical coherence, the receiver 'creates' or has to 'work' at it. Semantic and pragmatic meaning are coextensive in the text, and are typically utilised simultaneously in discourse.

1.2.3 Definitional and non-definitional relations of meaning in lexical cohesion/coherence

Definitional relations of meaning as those which involve semantic equivalence, entailment and semantic opposition

are potentially cohesive. They can be defined as "downgraded predications" (Leech, 1981: 144) which are semantic elements equivalent to a "feature" in function but have the structure of a proposition (or predication). Downgraded predications are usually expressed by means of a relative clause or a phrase. For instance:

A man who was wearing a wig (clause) (entered the room)

= A man with a wig (phrase)

= A bewigged man (phrase)

They are included in the definition of most nouns:

A butcher is a man who sells meat

A philatelist is a man who collects stamps

A thief is a man who steals things

A butcher sells meat by definition and a philatelist collects stamps by definition. The cooccurrence of butcher and meat, philatelist and stamps, and thief and steals in text produces a definitional link of semantic lexical cohesion.

These meaning relations may also be described in terms of "cases" (Fillmore, 1968). In the following example,

(2) The Jones had all their jewellery stolen.

The thief got away.

The case of thief is Agentive, as thief is the understood Agent of the act of stealing, and stolen has an implicit

Agent, viz. the thief. Thus, stolen and thief can be linked on these grounds. They are likely to collocate, and are potential sources of lexical cohesion.

On the other hand, non-definitional relations of meaning are these various meaning relations not easy to classify in systematic semantic terms, as, for instance, the relations holding between treat and chocolate, happy and win. Their cooccurrence in text produces a non-definitional link of pragmatic lexical coherence. For example:

(2a) John must sell his car. He needs a new Hifi equipment.

The meaning relationship holding between the lexical items, sell and Hifi equipment, can be described as non-definitional pragmatic owing to the fact that a 'pragmatic' inference has to be made in order to connect one lexical item to the other: By selling his car, John will get some money to buy a new Hifi equipment. Such connection is possible only if writer and reader share prior general knowledge that buying and selling involve money transactions. Only then will the reader deduce that the relationship between sell and Hifi is causal although such relationship is not signalled by any syntactic marker of causality (as, for example, 'because').

Defining and non-defining relations of meaning have been included in the category of "collocation" by Halliday and Hasan (1976: 284ff), a category all-embracing and sufficiently vague to include pragmatically

likely cooccurrence over indefinitely wide areas of text. The term 'collocation' will be used in this study with similar meaning, viz. to refer to semantic and pragmatically likely cooccurrence.

Cohesive and coherence links have a vital role in the interpretative process, but because of their explicitness, cohesive links are more likely to facilitate the reader's interpretation of a text than implicit links of coherence which involve recognition of the propositional and illocutionary development of the text as discourse. The interpretative aspect of lexical cohesion is treated in the next chapter.

The present discussion of cohesion and coherence can be summarised by the following diagram in which A implies B, but B does not necessarily imply A. (Diagram 1):

A	B
(linguistic) cohesion text - sentence	(pragmatic) coherence discourse - utterance
Links: are linguistic including semantic i.e. linguistic, definitional meaning.	Links: are pragmatic, including pragmatic, non-definitional meaning which may be more or less associated with specific vocabulary items.
Category: lexicosemantic cohesion	Category: lexicopragmatic coherence

In this sense the labels used in this study represent categories which are not watertight (therefore the notion of dichotomy, utilised hitherto, may be an over-estimation of

the phenomenon), but it is important to keep the levels separate and to use different terminologies for our purposes.

1.3 Review of literature on cohesion and lexical cohesion

A great deal of literature has been written on the subject of (lexical) cohesion and (lexical) coherence but most studies suffer from lack of consistent theoretical and terminological distinctions. Thus, in the field of "text linguistics", the terms 'cohesion' and 'coherence' have been used rather loosely or in most cases, interchangeably with the same meaning, as will be seen below. The joint publication in 1976 of Cohesion in English by Halliday and Hasan seems to be at present the most influential work on cohesion theory and for the techniques of textual analysis expounded in their book. Cohesion in English is largely an extension of Halliday's intra-clausal analysis beyond the sentence boundary. A fairly exhaustive account of the grammatical features of cohesion in English is given, but certain parts of it duplicate with Quirk and Greenbaum's (1973) description of this phenomenon (see Quirk and Greenbaum's chapter on "Sentence connection"). Cohesion in English is also an extension of Hasan's (1968) analysis of grammatical cohesion in English.

The concept of cohesion as defined by Halliday and Hasan (1976) is a semantic one. Cohesion is a linguistic relation, "part of the system of language" (p5). This concept is subordinated to that of text regarded as a semantic unit. Thus, cohesion is a linguistic property of

text, contributing to 'texture'. 'Texture' is defined as the property of "being a text" (p2). This property can be described as a combination of semantic and pragmatic configurations of two kinds: cohesion (semantic) and register (pragmatic) which appears to include content as a subpart. Thus, Halliday and Hasan's concept of "texture" corresponds to what is accounted for as 'coherence' in this study. (Note that the term 'coherence' does not feature in Halliday and Hasan's account of cohesion).

Halliday and Hasan's concept of 'cohesion', being a 'semantic' feature of 'text', it does not in principle include non-linguistic/pragmatic factors for these are not aspects of text-cohesion but of discourse-coherence (in our terminology). However, they are allowed in practice. The most striking example of conflation of the linguistic and the pragmatic level are found in their analysis of 'lexical cohesion' (1976: 286), and their definition of cohesion collapses, we feel, when they attempt to illustrate how cohesion holds in a whole paragraph. Part of their passage used to illustrate this point is reproduced below:

- (3) "... After spending the whole day within half an hour or so of sundown, I was still several hundred feet below the summit. Then my hopes were reduced to getting up in time to see the sunset ..."

Thus, the cooccurrence of sundown and summit in this passage clearly produces a link that is not definitional but pragmatic. Halliday and Hasan define this link as "collocational cohesion" which is a category allowing

pragmatic features of cohesion, and hence is not exclusively semantic. Their following remark (1976: 287, my emphasis):

"But it should be borne in mind that this (collocation) is simply a cover term for the cohesion that results from the cooccurrence of lexical items that are in some way or other typically associated with one another, because they tend to occur in similar environments"

reflects the ambiguity of a situation where the semantic is being extended over to the pragmatic. The point they make that "the relatedness (of lexical items) is a matter of more or less" (p289) may be restated as "a matter of cohesion or coherence".

Van Dijk (1977) proposes a "linguistic study of discourse" (preface pvii, my emphasis) which is in reality both text linguistic and discourse pragmatic in our theoretical framework. This concept of "coherence" as a "semantic property of discourses, based on the interpretation of each individual sentence relative to the interpretation of other sentences" (p93) seems to characterise both cohesion and coherence because it is semantic and pragmatic. Van Dijk's "coherence analysis" of the example below illustrates his viewpoint (p98):

"Clare Russell came into the 'Clarion' office on the following morning, feeling tired and depressed. She went straight to her room, took off her hat, touched her face with a powder puff and sat down at her desk. Her mail was spread out neatly, her blotter was snowy and her inkwell was filled. But she didn't feel like work (...)"

Semantic and pragmatic relations are included in Van Dijk's description of "inclusion", "membership", "part-whole" and "possession". Thus, Clare Russell and face may be viewed as semantic part-whole, also the relationships

between office and room, office and desk are linguistic. But the relationships between mail and blotter, Clare Russell and hat, and face and powder-puff, are based on pragmatic knowledge. Van Dijk seems to imply the semantic-pragmatic dichotomy in the text by remarking that "The individuals represented by lexical items seem to cluster around two concepts, viz. the 'human (female) individual' and the 'office' concepts" (p98).

Some European scholars, mostly working on the French language, did not attempt to discriminate between a semantic and a pragmatic level in their study of cohesion/coherence. 'Coherence' has often been used as a cover term to include both semantic connectedness between textual elements and pragmatic linkage between textual elements. Bellert (1970), Slakta (1975), Charolles (1978) and Marcus (1980) have considered 'coherence' to be a property of text. De Beaugrande and Dressler (1981) have regarded 'cohesion' and 'coherence' as "text-centred notions designating operations directed at the text materials" (p7, their emphasis). Szwedek (1980) has included pragmatic features in lexical cohesion which we view as part of lexical coherence in this study. Also the notions of 'text' and 'discourse' were not kept terminologically distinct in Hasan's (1968) and Halliday's (1970) account of cohesion. Hasan distinguished between "the internal and the external aspects of 'textuality'" (her emphasis), the first having to do with the devices used to link sentences together to form 'text', that is, 'cohesion',

the second involving the ways these sentences link meaningfully with the situation in which they are used. This second aspect of Hasan's description included features of the 'register' which we define as a discourse (coherent) feature. Similarly, for Halliday (1970: 143), the "textual function" of language is concerned with "making links with itself and with features of the situation in which it is used" and cohesion is one aspect of Halliday's textual function of language. Some discourse analysts, on the other hand, as for instance, Coulthard (1977), seem to draw a distinction between discourse coherence and text cohesion. For example (Coulthard, 1977:10):

(4) A - Can you go to Edinburgh tomorrow?

B - BEA pilots are on strike

This interchange is regarded as discourse because it is coherent, and not as text because it is not cohesive. The lexicopragmatic links of coherence produced between go and pilots and go and strike are inferable via pragmatic coherence. Thus, the reader of (4) can supply the missing proposition (or link), "I can't go to Edinburgh" because he knows that on strike implies that B cannot go to Edinburgh, by virtue of his pragmatic knowledge.

1.4.0 Lexical versus grammatical cohesion: similarities

Lexical cohesion (or lexicosemantic cohesion) has been defined as a relation in text that utilises lexical devices for its achievement. These devices include sense relations

which belong to the system of the language. Grammatical cohesion can be viewed as a relation in text which utilises grammatical devices for its production, namely, the system of proforms which include pronouns and deictics. There are similarities between these two types of cohesion involving coreferentiality and repetition of semantic meaning. In lexical cohesion, coreferentiality is carried on with anaphoric lexical reiteration of semantic meaning. On the other hand, the pronominal system of grammatical cohesion, which represents the reduced form of its nominal system, also retains the coreference of the word for which it substitutes when semantic meaning is being repeated.

(4a) My neighbour is a great cook. The man made pizza the other day.

(4b) My neighbour is a great cook. He made pizza the other day.

In (4a) coreference is carried by a reference item plus a lexical item (a hyperonym). Thus anaphoric reference item the accompanies the lexical item man semantically related to neighbour. NP the man repeats part of the semantic content of neighbour through lexical reiteration. This lexical repetition produces lexical cohesion.

In (4b) coreference is expressed by a grammatical item. The semantic content of neighbour is being repeated in a reduced form in He. This grammatical repetition produces grammatical cohesion.

Thus in grammatical cohesion, the linguistic index He is grammatical but functions in the deep structure as a

repetition of a surface lexical form. In lexical cohesion, semantic meaning may be repeated via various sense relations (see Section 1.7. below), but at the point where the reiterated item is a hyperonym (as for instance man) accompanied by reference item the, the dividing line between lexical and grammatical cohesion becomes less clear and one type of cohesion shades into another.

However, when pragmatic meaning is being repeated, grammatical cohesion no longer bears similarities with lexical cohesion. For example:

(4c) John turned to the ascent of the peak. The air felt pure.

(4d) John turned to the ascent of the peak. It felt pure.

In (4c) the air is the ellipted form of the air of/during the ascent. It is understood pragmatically by reference to ascent. But in (4d) It is not the reduced form of ascent. It does not refer to ascent, nor can it substitute for it. It is pragmatically incongruent with ascent.

It is generally assumed that a text exhibiting grammatical devices of cohesion has a tendency to be more ambiguous to the reader than a lexical cohesive text in the sense that explicitness provided by the repetition of lexical meaning via full lexical forms is lacking in grammatical cohesion. It is quite conceivable that heavy use of pronouns, proforms and ellipted forms are likely to obscure the writer's message and hence the explicitness provided by lexical devices of cohesion is crucial to text understanding. From the point of view of the writer

producing text, the question of whether the amount of cohesive and coherent links he supplies satisfies the reader has to be answered in terms of pragmatic principles being observed or violated when encoding a message. (This point is treated in Section 1.5.3 below within the framework of textual rhetoric.)

1.5.0 Presupposition and implicature in lexical cohesion/coherence

1.5.1 Presupposition

At the basis of lexicosemantic cohesion and lexicopragmatic coherence lie the two notions of presupposition and implicature. Presupposition is a concept much discussed in both philosophy and linguistics. It is a problematic category which can be viewed as logical or as pragmatic, or as both.

There are two rival theoretical approaches to presupposition: the logical theory treats presupposition as a relation between propositions defined in terms of their truth or falsehood. The pragmatic theory explains presupposition in terms of the meaning of a speech act in relation to the speaker's or the reader's beliefs. Following Gazdar (1979) and Leech (1981) we shall view presupposition as a logical as well as pragmatic relation. Logical presupposition underlies the notion of entailment which is a logical or semantic relation. This relation is often called "logical implication". In a logical view of presupposition, a distinction is usually drawn between presupposition and entailment. For example,

- (5) a. John stole a valuable book from the library:
entails John stole a book.
- b. John did not steal a valuable book from the library: does not entail John stole a valuable book.
- c. The book John stole from the library is valuable: presupposes John stole a book.
- d. The book John stole from the library is not valuable: presupposes John stole a book.

The negation test is a criterion of presupposition. It shows that entailment is vulnerable to negation (a. and b.) whereas presupposition is not (c. and d.).

Pragmatic presupposition involves implicature which is a relation of 'pragmatic implication' defined in terms of the speaker's and the reader's assumptions and beliefs (Leech, 1981). In a pragmatic view of presupposition, presupposition is usually distinguished from 'assertion'. Presupposition is that part of the content of an utterance which is treated as if it is familiar and 'assertion' is that part of the content of an utterance which is treated as if it is unfamiliar, new, informative. (This distinction corresponds to Halliday's theme - given and rheme - new: see Section 1.7.8). For instance, the presupposition of:

Prince Charles is a vegetarian,

can be expressed as: a. There is an X such as that X is
Prince Charles

and the assertion can be expressed as:

b. X is a vegetarian

Presupposition seems to be tied to the surface form of expressions and specifically to particular lexical items of different linguistic categories. Levinson (1983) describes presuppositions as non-semantic and context-sensitive which are 'triggered' or generated by certain linguistic items like 'implicative verbs' (see Karttunen, 1971 on implicative verbs). For example,

(5) John managed to stop in time presupposes that John tried to stop and implies that John stops.

But these pragmatic presuppositions which "presuppose" or "imply" certain beliefs of the producer are viewed in this study as "implicatures" rather than "presuppositions".

1.5.2 Implicature

Implicature, or specifically "conversational implicature" (Grice, 1975; see also Leech, 1983: 9, on "pragmatic implications") is one of the most important ideas in pragmatics. One of its assets is that it can offer powerful pragmatic explanations to linguistic phenomena. Unlike presupposition which is a kind of semantic inference, implicature is a kind of "pragmatic inference" (Levinson, 1983: 97) which seems to lie outside the organisation of language and at the same time affect its use via the general principles for co-operative interaction which underlie this type of inference. For instance, implicature provides some explicit account of how it is possible to 'mean' more than what is literally expressed by the conventional sense of an utterance. For example,

(6) A - How old is Sheila's little girl?

B - Well, she started school last term

One reading of B's answer might be paraphrased as follows: "No, I don't know her exact age but I can provide some information from which you may be able to deduce her approximate age, namely, she started school last term (i.e. she is between 5 and 6 years old according to British school age norms). Implicature is explained by "conversational" principles" which include Grice's Co-operative Principle (CP) and other principles such as Leech's (1983: 9-30, 79-116; 1983: chapters 4-7) Politeness Principle (PP). The CP: "Make your contribution such as required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged" consists of four basic maxims of conversation or general principles underlying the efficient co-operative use of language. These are expressed as follows (adapted from Grice, 1975: 45-6):

1. Maxim of Quantity: give the right amount of information i.e.
 - (a) Make your contribution as informative as required.
 - (b) Do not make your contribution more informative than required.
2. Maxim of Quality: try to make your contribution one that is true i.e.
 - (a) Do not say what you believe to be false.
 - (b) Do not say that for which you lack adequate evidence.

3. Maxim of Relation: be relevant
4. Maxim of Manner: be perspicuous i.e.
 - (a) Avoid obscurity of expression.
 - (b) Avoid ambiguity.
 - (c) Be brief. Avoid unnecessary prolixity.
 - (d) Be orderly.

Grice's conversational principles have been reinterpreted by Leech (1983) in the light of his "Textual Pragmatics" and include principles of "textual rhetoric" that may illuminate our analysis of lexical cohesion (see Section 1.5.3.) We now propose to examine how presuppositional semantic meaning and implicated pragmatic meaning are accounted for by cohesive and coherent conversational discourse.

Consider the following examples:

- (7) A - Would you like some fish?
B - Yes, I'd love some of this haddock

In presupposition terms, speaker A tends to presuppose that there is fish generally available, otherwise he wouldn't have offered any (the possibility that A wants to be malicious by proposing something that does not exist is therefore discarded). Speaker B presupposes that there is some haddock immediately available. Thus there is lexicosemantic cohesion in the presupposed fish-haddock relationship. Furthermore, speaker B uses love which is not presupposed by like in A's utterance but which is implied by it. Note that like asserted in A's utterance could have been implicated, as in the following exchange:

(7a) A - How about some fish?

B - Yes, I'd love some of this haddock

In this example, like in A's question is not asserted but implicated. It features in the pragmatics of A's utterance. The reader must infer that A's question implicitly meant: would you like some fish? Even if B's answer was negative: "No, I hate haddock", the same type of inference would have been involved.

Coming back to example (7) above, we may note that even if B's answer was: "No, I hate haddock", hate would still be implied by asserted like.

Thus there is lexicosemantic cohesion in the presupposed fish-haddock and also in the asserted/implicated part, like-love. It therefore appears that lexical cohesion can involve any aspect of meaning, presupposed (fish-haddock) or not presupposed (like-love), implied (like-love) or not implied. Not implied meaning is implicated meaning and features in lexical coherence, as for example, the relationship how about-love in (7a) above.

In 'conversational' terms, B's utterance reflects the fact that A's contribution was intended to be informative, true, relevant and non-ambiguous. That is, B's utterance means: B wants (implied by like) some fish (presupposed in haddock). Compare the example below which deals with "a special kind of informativeness" (Smith and Wilson, 1979: 177), i.e. oblique informativeness:

(7b) A - Would you like some fish?

B - I've lost my socks

B's utterance may be interpreted in at least two ways: have lost has a pragmatic overtone that one might be looking for something and hence not be free to do anything else, i.e. eat or buy fish. Here the pragmatic lexical links come in to make a connection where the CP suggests there should be one but semantically there is no connection. So one may consider that B was being informative and relevant to the extent that B's response is an explanation of why B cannot answer A's question, that is, B's response is pragmatically relevant to A's question and could be an example of coherent discourse. The other interpretation of B's utterance in (7b) may be by considering that B ignores A's question: B's utterance implicates that he tries to escape A's question by giving a non-co-operative reply. B's contribution is rather negative as it does not observe politeness principles and hence violates 'social goals' by not answering A's question. But it satisfies B's personal goals (i.e. B is looking for her socks).

Thus, while (7) was an instance of LS cohesion, (7b) is a case of "weak" LP coherence in comparison to (8) below (Smith and Wilson, 1979: 175):

(8) A - Where's my box of chocolates?

B - The children were in your room this morning

Although B's utterance can be regarded as an oblique response to A's question, the link between both utterances is more direct than in (7b) because there is less reconstruction (or inference making) involved. Assuming

that A and B are observing the CP and share the background knowledge that children are fond of chocolates generally, this enables the implicature to be worked out. Thus chocolates in A's question implicates children in B's response and hence produces a lexicopragmatic link of coherence which was not as obvious in (7b) above. Neither (7b) nor (8) contain cohesion of either type (viz. lexical or grammatical). There is clearly much to be said about the interrelationship of Grice's maxims, implicatures, presuppositions and cohesion, but since this is not the main theme of our thesis, this subject is not pursued.

1.5.3 Leech's principles of 'Textual Rhetoric': how they can explain LS cohesion and LP coherence

Reinterpreting Halliday's (1970, 1973) distinctions of the three functions of language, viz. the 'ideational', which enables the language user to convey and interpret experience of the world through the language, the 'interpersonal' which deals with the expression of the language user's attitudes and judgement and his influence upon the attitudes and behaviour of his interlocutor, and the 'textual' which enables the construction of text in spoken or written mode, Leech (1983: 64) distinguishes the ideational function of language which belongs to 'grammar' in a general sense, and the interpersonal and textual functions which belong to pragmatics. From the receiver's point of view, which is our concern in this study, the 'Textual Rhetoric' constrains the input and the

'Interpersonal Rhetoric' constrains the output of the decoding process. From the producer's point of view, these constraints are reversed. Leech's 'Interpersonal Rhetoric' includes principles which are akin to Grice's (1975) conversational principles. The 'Textual Rhetoric' is like the 'Interpersonal Rhetoric' in the sense that it is based on speaker/writer - hearer/reader co-operation, a "textually well-behaved utterance" being defined as one which anticipates and facilitates the hearer/reader's task in decoding, or in making sense of the text (Leech, 1983: 60). The textual rhetoric consists of four principles which are pragmatic factors constraining the form of texts:

1. The Processibility Principle: "Be humanly processible in on-going time".
2. The Clarity Principle: "Be clear".
3. The Economy Principle: "Be quick and easy".
4. The Expressibility Principle: "Be expressive".

How does the achievement of LS cohesion and LP coherence in text/discourse accommodate these principles which underlie the 'interpersonal' and 'textual' functions of language?

A language user utilising relationships of LS cohesion and LP coherence in text/discourse seems to comply with some of the principles of 'textual rhetoric' and to violate others. Let us examine the extent to which some text rhetorical principles are respected in lexical cohesive and coherent text/discourse.

1 The Clarity Principle: In order to observe this principle, the language producer must utilise explicit devices of cohesion such as repetition of lexical meaning together with reference items so that the likeliness for the reader to draw inferences to understand the writer's message is reduced to a minimum. Thus, explicit signals of cohesion, which have been earlier defined as linguistic semantic and definitional, contribute to the text clarity and reflect the writer's willingness to co-operate with the reader.

On the other hand, grammatical devices of cohesion, as noted earlier, are less likely to fulfil this function. Consider, for example, the following passage (quoted in Williams, 1983: 42) where the grammatical anaphoric item that lacks clarity from the point of view of its reference:

"The construction of Caborra Bassa has already meant that 25,000 Mozambicans have had to move their villages to make way for the 240 km long lake. And - like other countries struggling to stand on their own feet in the middle of a world recession - that may well prove simply too expensive for Mozambique"

What may well prove simply too expensive? The construction of Caborra Bassa, the 240 km long lake or the middle of a world recession? Lexical reiteration would have made this point clearer. This text violates the clarity principle and compels the reader to rely on his world-knowledge to establish coreferential linkage.

2 The Processibility Principle: While the time factor is important in conversational exchanges and underlies this principle ("Be humanly processible in on-going time"),

it seems to be less crucial in written discourse. Thus, a reader presented with written discourse is not subjected to the same time constraints as a hearer. However, another function relating to the principle of processibility, involving written discourse, relates to information distribution (which was part of Halliday's textual function), that is, to its separation into "given" and "new", definable as information interpretable in relation to what has gone before (or old information) versus information that has not occurred before. (This is discussed in Section 1.8 of this chapter). The explicitness of textual information is subordinated to the occurrence of anaphoric signals (or reference items) as for instance definite articles (the) and deictics (this, that) which indicate to the reader what should be treated as 'given' information in the text/discourse. Thus, definite NPs which consist of repeated lexical items accompanied by anaphoric reference items must be decoded as given. The occurrence in text of such cohesive items reflects the writer's observance of the Processibility Principle, which subsumes the principle of clarity defined earlier. However, lack of lexical signals of cohesion, as for instance, those expressed by anaphora are likely to violate the principle of processibility and as a result may be more difficult to process by a reader. (The processing of lexical relations of cohesion and coherence in comprehension is the object of the next chapter). Moreover, heavy use of proforms may render a text unintelligible and

drive the reader into much inferencing.

3 The Economy Principle: This principle is continually in conflict with the Clarity Principle in matters involving lexical cohesion, because the Maxim of Reduction which governs it seems to plead against the explicitness of linguistic links of cohesion and for implicit links of coherence. Thus, by the fact that repetition of linguistic definitional meaning via lexical form is at the basis of LS cohesion, the Economy Principle seems to be bound to be violated. However, in grammatical cohesion, this principle is fully observed for the replacement of a lexical form by its syntactic equivalent is doubtless more economical. For instance, consider the example below where the replacement of The vehicle by it complies with the Economy Principle (as well as protects the intelligibility of the text):

(9) John sold his car. The vehicle was in poor condition
(9a) It

Both Principles of Economy and Clarity are thus respected in (9a). But the Maxim of Reduction which underlies the Principle of Economy and which is usually complied with in grammatical cohesion is not recommended where it leads to ambiguity. Thus pronominalisation in example (10) below complies with the Economy Principle but violates the Clarity Principle, and in order to recover the intelligibility of the message, economy has to be sacrificed in example (10a) by identifying the 'beneficiary' in the initial sentence, that is, Mr. Mishra Dayal.

(10) Unintelligible, ambiguous text:

The Governor announced another attack yesterday. He said he is the first member of the South African Indian Community to have been violently attacked in Durban.

(10a) Intelligible text:

The Governor announced Mr Mishra Dayal's attack yesterday. He said he is the first member of the South African Indian Community to have been violently attacked in Durban.

4 The Expressivity Principle: This principle is concerned with the effectiveness (including expressive and aesthetic aspects of communication) rather than simply the efficiency of a message. It is being observed typically when lexicosemantic cohesion is based on the device of expressive repetition, as in the example below:

(11) Sarah Bilton has lost the locker's key.
Sarah Bilton will have to look for them.

This type of expressive repetition where the emphasis of repetition has the rhetorical value of rousing the interest of the addressee, or impressing, is a device often used in children's fairy tales, and in sporting events radio broadcast (as for example horse races or car races) and is treated in Sections 1.7.1., 1.7.2., and 1.7.3. below.

Thus a text exhibiting signals of lexicosemantic cohesion is a demonstration that at least two principles were being observed by the writer in his attempt to convey

his message to the reader, viz the Clarity Principle and the Processibility Principle. But the reader, specifically the non-native, is likely to be unable to acknowledge with these principles. For instance, he may not be able to draw some linguistic inferences of the type needed to connect the vehicle (superordinate) to car (hyponym) in the example below:

(12) John sold his car. The vehicle was in poor condition.

Thus, he may not treat The vehicle as given information because already mentioned in car, that is, he may not comply with both principles underlying this utterance.

Relationships of lexicopragmatic coherence communicated by the writer via the text are discourse relationships. Hence they cannot be analysed in terms of logical presuppositions and have to be treated within the framework of pragmatic implicature. In the same way as relationships of lexicosemantic cohesion encoded in the text by the writer are likely to comply with some of the principles of textual rhetoric, lexical pragmatic relationships of coherence can also comply with some of these principles and violate others. For instance:

1 The Clarity Principle may be violated when discourse markers (for example, but, so, however) are missing which would help uncover pragmatic links between lexical items. For example:

(13) The Jones didn't have a good picnic. The knives were missing and it started to rain.

The causal relation underlying the cooccurrence of the items, picnic, knives and rain could either be signalled with markers (because) or be deduced from the lexical content of the juxtaposed sentences. Signalling a "logico-rhetorical" relation by using syntactic markers of connectivity as in grammatical cohesion is a demonstration by the writer that he is willing to observe the Clarity Principle in order to co-operate. But using implicit links of pragmatic coherence as in (13) above shows that this principle was not observed and may imply that the writer is not willing to co-operate with the reader. But is the cause of the problematic interpretation of a discourse necessarily attributable to the writer who did not comply with the Clarity Principle? It is often the case that discourses are coherent without being cohesive. They involve implicit pragmatic links of coherence implicated by the writer which have to be deduced by the reader. Does the Clarity Principle necessarily involve the marking of conjunctive relations, for example? And to what extent is interpretation a function of conjunctive markers essentially? We shall attempt to answer these questions in Chapter three.

2 The Economy Principle, however, seems to be observed when conjunctive meaning is expressed via vocabulary items rather than via overt "discourse" markers. For instance, the use of therefore in (14) below, is uneconomical in

comparison to (14a):

(14) John failed his driving test. Therefore he was sad.

(14a) John failed his driving test. He was sad.

But the observance of this principle in (14a) would imply that the Processibility Principle is necessarily violated, that is, conjunctive meaning not overtly signalled but implicit is less easy to process, and that the Clarity Principle is also violated since it is not clear from (14a) whether John's sadness was the effect or the cause of his failing the driving test.

1.6.0 An analysis of the relations underlying lexical cohesion/coherence

Before going into a formal analysis of the various lexical devices of cohesion and coherence available in the English language, it seems necessary to introduce three types of relation that are involved in the achievement of lexical cohesion and lexical coherence, viz. the referential relation, the substitutional relation and the conjunctive relation.

1.6.1 The referential relation of lexical cohesion/coherence

Lexical cohesion may be viewed as a referential relation in text. Lexicoreferential cohesion (henceforth referred to), to which Halliday and Hasan's (1976) lexical 'reiteration' belongs, is a relation which holds between lexical items having common referent in text. In traditional semantics reference holds between linguistic expressions (or signs) in a text and entities in the world. Thus, it is used with 'sense' to discuss lexical meaning. For instance, the meaning of the lexical item bird is in part determined

by its 'sense' which are the component properties of meaning (or 'semantic features') of 'animate', 'feathered', 'flying', etc. and also by its 'reference', that is, the set of objects in the world to which the lexical item can be correctly applied. But reference is also that function whereby a speaker/writer indicates, in the use of a linguistic expression, the entities he is talking/writing about.

The referential relation underlying lexical cohesion (and lexical coherence) is achieved through the presence of reference items in text/discourse. There exist some items in every language that have the property of making reference to something else for their interpretation. They are usually used with vocabulary items, though they may also be used on their own. These reference items are directives which indicate to the reader that information is retrieved elsewhere, that is, that the meaning of the vocabulary items with which they cooccur has to be retrieved in the preceding or the following co-text, or in the larger context. Thus, reference items are potentially cohesive: they occur with vocabulary items which serve as the source of the interpretation. They function as reminders in text as they instruct the reader that the same thing enters into the text/discourse a second time. The referential relation underlying lexical cohesion and lexical coherence is achieved via anaphora or cataphora.

1.6.1.1 Anaphora

Anaphora (or anaphoric reference) is backward pointing reference. It is the most common directionality for coreference. The example below illustrates this point:

(15) A group of women who had set up a peace camp in Buckinghamshire were trialed yesterday. The women were nervous. They were fined £300.

This example instantiates anaphoric cohesion achieved via lexical and syntactic devices. Anaphoric cohesion achieved lexically (or lexicoreferential cohesion) relies on the cooccurrence of women and the women: the women in S2 is anaphoric to women in S1 and coreferential to it. Therefore the complex consisting of the and repeated lexical item women is cohesive by reference. However, the repetition of a lexical item being itself cohesive, reference does not have to be repeated to produce cohesion (This point is discussed in Section 1.6.2). On the other hand, grammatical cohesion achieved via the cooccurrence of the women and they has also a referential function since they (personal pronoun) points back to the women and is coreferential to it.

The next example of lexicoreferential linkage involves pragmatic knowledge:

(15a) Picasso died five years ago. The author of Demoiselles d'Avignon bequeathed his personal collection to the museum of Barcelona.

A referential link of lexical coherence is produced between the author of DA and Picasso. The author is coreferential

to Picasso, but whether or not it is perceived as such by the reader depends upon his world-knowledge. If he knows that Demoiselles d'Avignon is a painting by Picasso, he is being reminded only of the coreferentiality of the link, and this is a case of LP coherence. But if the reader is ignorant of arts, no pragmatic link can be established between the two lexical items. Rather Picasso and the author of DA will be treated as if they had different referents: for example, the author of DA could be a writer (a friend of Picasso's) who handed in Picasso's paintings to the Barcelona museum. If it were the case that (15a) could not be interpreted appropriately, this could be explained in terms of a breach of the maxim of manner by the writer. Thus his would be ambiguous in (15a) because it wouldn't be clear whether his would relate to Picasso, or to the author. As we remarked earlier in Section 1.5.3, pronominal reference in grammatical cohesion can sometimes violate the clarity and the processibility principles all the more if prior knowledge is not available.

1.6.1.2 Cataphora

Cataphora (or cataphoric reference) is forward pointing reference. It is less frequently used as a lexical cohesive relation than anaphora. Cataphora typically occurs with deictics this, that, here, which, by pointing forward, derive their interpretation from something that follows, for example:

(16) This is what you will be missing with ratecapping.

You will lose hundreds of local firms and businesses who depend on the Council for work, you will have to

pay for community services which are now free ...
All these are worth defending (Sheffield City
Council leaflet on Rate Capping).

This connects forwardly with the rest of the text, that is, whatever information is provided by the following sentences. This example features grammatical cohesion.

Cataphoric reference is not found with lexicoreferential cohesion because with the same or related lexical item occurring twice over, then obviously the second occurrence must take its interpretation from the first, as in (17) below:

- (17) Mrs Thatcher arrived late at the press conference.
The Prime Minister didn't apologise. She was wearing a navy blue dress. Her press secretary was accompanying her. The first question to Mrs Thatcher was about EEC policy.

Mrs Thatcher being "thematic" in the text, this lexical item reduces the expectations of the reader while this in (16) above projects him forward into the text to seek for more information. (This may be viewed as "rhematic"). On this point it seems interesting to note that there seems to be no way of predicting, after the 'theme' has been given, whether repetition of it will take a lexical or a grammatical form. Does the reader have 'discourse' reasons for expecting 'full' reference, (The Prime Minister) or 'partial' reference (she, her)? If it is entirely a matter of choice on the part of the writer, this choice may be determined by time factors: the name has not been

mentioned for a while and the writer may have lost track of which person is being referred to. It may be determined by a need to spotlight the hero rather than his actions. In English, the writer's decision to repeat a noun or change to the pronoun seems to depend, in part, on how necessary it is to reidentify the referent, for example, "as a gentle reminder" (Hatch, 1983).

1.6.2 The substitutional relation in lexical cohesion/coherence

Relationships of lexical cohesion/coherence may be produced via the relation of 'substitution'.

Substitution is used in this study in a surface sense only, although in the 'cohesion' context, and is therefore not as restrictive as Halliday and Hasan's (1976) use of substitution viewed as a cohesive semantic relation which involves the replacement of one lexical item by a proform (pronoun or proverb). Lexicosubstitutional cohesion is a relation between linguistic items which involves the repetition of form and/or semantic or pragmatic meaning without repeating reference. Thus, while Halliday and Hasan (1976: 88) view substitution as "a relation in the wording rather than in the meaning", we regard substitution as a relation on the level of form (lexical) as well as the level of meaning (semantic or pragmatic) which does not involve coreferentiality. The substitutional link of cohesion/coherence can be produced via lexical items semantically related, for example, synonyms, as it can also be made on the basis of the repetition of part

of the semantic/pragmatic content of an utterance in another utterance without repeating reference. This phenomenon has been described by Leech (1981: 190) in relation to semantics as "co-semy" or "correspondence of meaning" without coreference. Such correspondence may be total (that is, verbatim repetition of the lexical items) as in example (18) below or partial where part of the meaning of a lexical item is being repeated (and also different parts of speech are involved) as examples (19) and (20) show.

- (18) Some big American cities have ten or a dozen TV channels. Where two towns are close together such as Washington and Baltimore, the choice may lie between twenty or more channels (New Society, 1984).
- (19) To Nehru the end and the means were inseparable. He was firmly committed to socialism. This was a goal which he believed was worth pursuing only if it could be achieved peacefully (New Society, 1984).
- (20) The man began the terrifying ascent of the cliff. Slowly but surely he climbed higher and higher until he had nearly reached the top.

1.6.3 The 'conjunctive' relation in lexical cohesion/coherence

The cooccurrence of vocabulary items in text/discourse may convey a 'conjunctive' relation which may be paralleled to 'conjunctive' grammatical cohesion. Conjunctive meaning can thus be produced by such lexical cooccurrence and result in conjunctive-type lexical cohesion.

Conjunctive meaning may result from the juxtaposition of propositions expressing the conceptual categories of time, causation, contrast, or quantity. Such meaning may be embodied in conjunctive relations of temporality, causality, adversativity or additivity as categorised by Halliday and Hasan (1976) and signalled via syntactic markers. However, syntactic choice to express propositional meaning that is conjunctive is not the sole agent of conjunctive meaning in the discourse as lexical choice can also express this type of meaning. In this sense it may be argued that the presence of syntactic markers of 'conjunction' (in Halliday and Hasan's sense), or lexical equivalents of syntactic markers (paralexical markers) of 'conjunction' is somewhat redundant in text/discourse as conjunctive meaning is also embodied in cohesive/coherent lexical items. For instance, the conjunctive relation of adversativity may be syntactically signalled by conjuncts 'but', 'however', 'yet', 'on the other hand' (see Halliday and Hasan's 1976: 242-43 fairly exhaustive list of "conjunctive adjuncts") and equally be expressed by antonyms or opposites cooccurring in the text/discourse (see Section 1.7.7 on this point. Note also that Halliday and Hasan do not account for "asyndeton", that is, the lack of formal signs of 'conjunction' (Quirk and Greenbaum, 1973)). Sequentially expressed propositions are expected to be relevant to one another and when the link between them cannot be made through syntactic means (via overt linkers explicit on the surface of the discourse) this has to be made via lexical means.

As recipients of an utterance like:

(21) I hated that man. He looked strange.

we use our knowledge of the world and our expectations concerning the sequencing of oral/written production to relate the two propositions. We treat the second proposition as relevant in a particular sense, for instance, as providing an explanation for the initial assessment on the basis of lexical content which is an alternative to syntactic explicitness in this case. Thus, the relationship of lexicopragmatic coherence produced by the cooccurrence of hated with strange conveys the same conjunctive meaning as:

(21a) I hated that man because he looked strange.

in which causality is overtly signalled by because.

Non-syntactic type of linkage has been said to characterise "unplanned adult discourse" as well as child language, that is, "discourse that lacks forethought and organisation preparation" (Ochs, 1979: 55). Conversely explicit syntactic links have been said to be heavily relied upon in "planned" discourse. It has been suggested that it may take more planning to express a specific semantic relation using a syntactic marker than to imply only that some semantic relation obtains. In other words, the speaker/writer encoding task is greater and may demand greater planning. If it were the case, and as remarked earlier in Section 1.5.3, conjunctive meaning marked syntactically would violate the Economy Principle,

but on the other hand, would comply with the Clarity and the Processibility Principles. Compare the examples below (Ochs, 1979: 67):

(22) Two girls (unplanned version)

A - I'm so ... tired, I played basket ball today
(pause) the first time since I was a freshman
in high school.

B - Bask(hh)et b(hhall) heh heh/heh.

(22a) Two girls (planned version)

I am so tired, because I played basket ball for the
first time since I was a freshman in high school.

But Ochs (1979) did not make it clear that the so-called "semantic" relation implied by the juxtaposition of the two propositions expressed by the sentences is in effect a pragmatic relation and also failed to mention that semantic and pragmatic relations obtain between vocabulary items primarily. However, her very interesting study of the organisation of planned and unplanned conversational discourse provides insights into the way some devices of lexical cohesion/coherence such as referential and non-referential repetition of lexical items are embodied in, for example, hyponymy, antonymy and synonymy.

Thus, a conjunctive relation underlies lexical cohesion, whether or not this relation is signalled. Such relation often has to be accounted for by pragmatics.

In (22) the propositions expressed perform the "speech acts" of explanation via the lexical items, tired and basket ball.

Paralexical markers: Conjuncts (or 'connectors') like so, because, therefore, have lexical paraphrase equivalents.

In the lexical connection by paraphrase equivalent (or paralexical connection) the item 'says' what the conjunctive relation is. For instance:

(22b) A. John didn't wear his seat belt. B. This caused him bad injuries.

C. So he had bad injuries.

D. He had bad injuries.

The pragmatic link between the two utterances A and B can be overtly signalled by a syntactic marker (Adjunct so) or by a paralexical marker (This caused) and expresses conjunctive meaning of a causal type. Besides, this meaning can be deduced from the lexical pragmatics of the text: thus, the fact of not wearing the seat belt causes bad injuries is a causal relation which can be made on a pragmatic world-knowledge basis but it would be mostly probabilistic in the absence of discourse markers as John's bad injuries may either be the cause or the consequence of not wearing his seat belt. Wider context is required to disambiguate these utterances. Whether a text containing overt markers of 'discourse' (syntactic or paralexical) or no marker is easier to interpret by a potential reader is a question that is investigated empirically in Chapter 3.

It seems evident that absence of a larger co-text (and context) renders interpretation of some texts/discourses difficult specifically when discourse markers are absent and the lexical content of the text is of little help. For example,

(22c)1. Bert left home at midnight. He missed the train.

This text is ambiguous because it lends itself to several readings. More than one connection is possible.

- (22c) 2. Bert left home at midnight. Therefore he missed the train.
3. " " " However, he missed the train.
4. " " " Nevertheless he missed the train.
5. " " " because he missed the train.

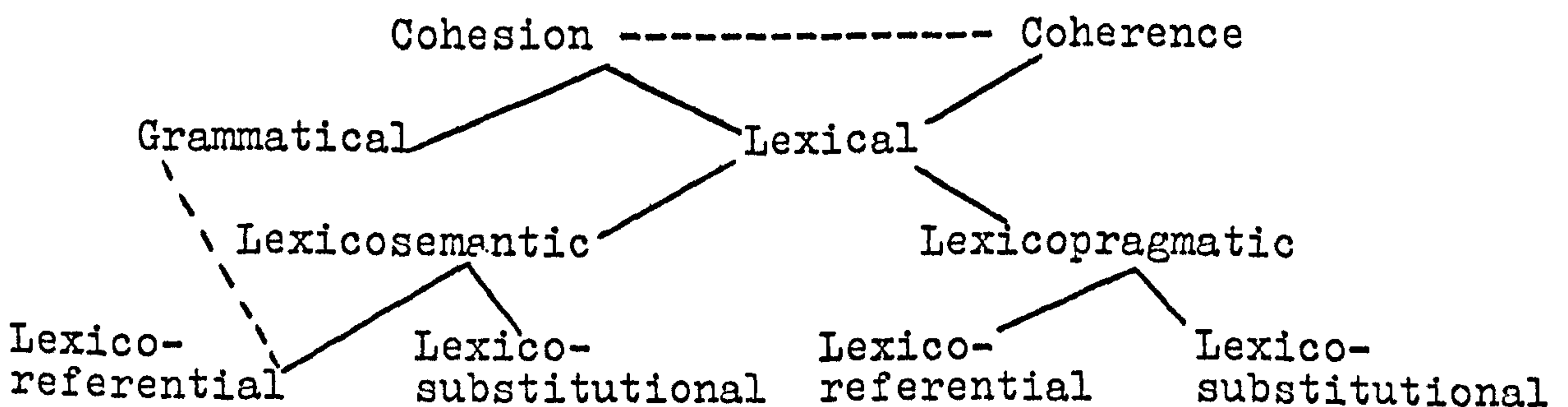
(22c) in uninterpretable. Ultimately it is incoherent, (although cohesive grammatically), because obligatory discourse markers are missing.

To summarise, three relations seem to underlie the achievement of lexical cohesion, viz. the referential, the substitutional and the conjunctive. These relations may coexist in text. Of course lexicoreferential and lexicosubstitutional cohesion are mutually exclusive that is, they cannot simultaneously account for the connectedness of utterances but conjunctive-type lexical cohesion may be coextensive with referential or substitutional cohesion, as can be seen in the text below:

(23) Unfortunately since the New Year Miss Morton has been unfit for school and will be absent until Easter. Fortunately we have Mrs Bramall covering the class until Miss Morton returns.

The cooccurrence of unfortunately and fortunately produces lexicosemantic cohesion. Substitution rather than reference is involved in this linkage. The underlying conjunctive relation implies conjunctive meaning of an adversative type (But/however is the missing connector). Contrastive meaning is also embodied in Miss Morton - unfit - absent and Mrs Bramall (understood as 'fit for school': this proposition is ellipted in sentence 2).

Diagram 2 below is a comprehensive account of the categories of lexical cohesion dealt with in this study.



1.7 Semantic and pragmatic resources of lexical cohesion/coherence

This section deals with the actual semantic and pragmatic means that constitute the potential of lexical cohesion.

The semantic resources of lexical cohesion are finite, but its pragmatic resources seem to be illimited in number. The sine qua non condition for the obtention of lexical cohesion in text/discourse is reiteration or repetition of lexical meaning, which may be semantic or pragmatic. Reiterative devices of cohesion are multiple and may be viewed on a continuum, with at one end complete repetition of the form and content of a lexical item and at the other partial repetition of lexical meaning in the shape of proforms (pronouns, and proverbs). Proforms are devices of grammatical cohesion. Some of them have been regarded as borderline cases of lexical cohesion by Halliday and Hasan (1976: 88ff). Between the two poles of lexical repetition there is a range of lexicosemantic devices which include full repetition of semantic content only, not form, as in synonymy, partial repetition of semantic content only, not form, as in 'cosemy' (Leech, 1981) (for example, nouns deriving from verbs and verbs from nouns), hyponymy, paronymy, antonymy.

1.7.1 Coreferential repetition of lexical meaning

Examples (24) and (25) illustrate coreferential verbatim repetition:

(24) There was a young princess who lived in a land far away. The princess was not very happy.

(25) Mix the butter, sugar, eggs and flour. Pour the mixture into a large baking tray.

(24) instantiates lexicoreferential cohesion based on the verbatim repetition of princess together with referential item the. (25) is an example of lexicoreferential cohesion based on the repetition of mixture, a morphologically (and also semantically) related item of mix, which could be described as a "coseme" to stretch a little Leech's (1981) category of 'cosemy', that is, correspondence of meaning, not form. It may be noted that repetition of morphologically related items does not appear in Halliday and Hasan's categories of lexical cohesion. Their "repetition of same lexical item" (p279) clearly refers to verbatim repetition and hence seems to exclude morphological repetition.² However, morphological repetition is not necessarily cohesive if different senses are involved (including idioms). For example:

(25a) John put in a request for leave but he wasn't given any. He had to put off his trip.

The repetition of put - put does not produce cohesion. It may also be noted that our use of 'reiteration' is broader than Halliday and Hasan's which involves the repetition of a lexical item (for example, a synonym, or a hyponym) referring back to another item related to it by having a common referent. 'Reiteration' and 'repetition' are used synonymously in this study.

The verbatim repetition of a lexical item, although it produces lexical cohesion, may result in lexical incoherence as in (26) below:

(26) Paula met John in Liverpool. Liverpool is a large city in Britain. Britain was bombed during World War 2. John is the son of a Yorkshire miner. Yorkshire beer is famous.

The repeated lexical items, Liverpool, Britain, John, and Yorkshire, link up with their previous occurrence, that is, they are cohesive, but this mere repetition does not guarantee coherence. This text is uninterpretable as discourse because it lacks a specific "topic of discourse" (Van Dijk, 1977 :131ff): Is the story about Paula meeting John, about Liverpool, about Britain's bombing or about Yorkshire beer? Only through the selection of one specific topic will this partly cohesive text (that is, some features of it only are cohesive, not the whole text) will become coherent discourse, as in (26a) below:

(26a) Paula met John in Liverpool. They got married at St.Thomas' Church.

Here the link produced is pragmatic: met - got married - Church , and provides coherence to the text/discourse.

There is also a grammatical link: Paula - John - They, which is cohesive (They corefers to Paula and John). However, in co-operative terms, we may consider that example (26), although incoherent, was nonetheless informative, and believed by the reader to be true. In this sense, neither Grice's Maxim of Quantity nor his Maxim of Quality seem to have been flouted. But the producer has breached the Maxim of Manner by supplying too many surface connections without ensuring underlying coherence. Also the Maxim of

Relation was violated because the message was irrelevant to the reader's expectations of finding out about one topic only, as for example, Paula and John meeting. Thus, the cooccurrence of lexical items in text related through lexical meaning (semantic or pragmatic) does not necessarily produce cohesion when these items are topically unrelated. On this point we may mention that Halliday and Hasan (1976: 319) wrongly note a "continuity of lexical meaning" between opened, key, and door, in their example:

(26b) "Soon her eyes fell on a little glass box that was lying under the table; she opened it, and found in it a very small cake on which the words 'EAT ME' were beautifully marked in currants. 'Well, I'll eat it', said Alice, 'and if it makes me larger, I can reach the key; and if it makes me smaller, I can creep under the door; so either way I'll get into the garden, and I don't care which happens!"
(my emphasis)

Clearly, the cohesion is between key and door, not between opened, key, and door, because Alice opened the box, not the door. The occurrence of opened in this context is irrelevant to the key-door relationship and is purely coincidental.

1.7.2 Non-coreferential repetition of lexical meaning

Repetition of lexical meaning does not have to be coreferential to be cohesive. Reiterated cohesive lexical items may be non-coreferential, that is, they may not

share a common referent. For example:

- (27) Jackson's supermarket is very clean inside.
Supermarkets should always be clean to please
customers.

(27) instantiates cohesive repetition of form and content but not reference via the cooccurrence of supermarket(s) in S1 and S2. Supermarkets in S2 has a generic sense: it refers to all supermarkets and this one included, and supposes a special stress configuration in its production in oral discourse. This is a case of lexicosubstitutional (non-coreferential) cohesion. Note also the non-definitional link produced by the cooccurrence in S1 and S2 of supermarkets and customers, and of clean and please.

These pragmatic links are crucial to the interpretation of (27) as coherent discourse and their absence would disturb its intelligibility. Compare the following example:

- (27a) * Jackson's supermarket is very clean inside.
Supermarkets should always be clean to repel
customers.

The non-coreferential repetition of supermarket in S1 and S2 and of clean in S1 and S2 does not contribute to the coherence of this discourse. Its oddity comes from the fact that its lexical pragmatics are somehow inadequate. There is a mismatch between the reader's association of cleanness with (customer's) happiness in (27), by reference to his world-knowledge, and the text's implication (in 27a) that cleanness is not appreciated by customers,

that is, it is a customer repellent.

Thus, it appears that non-coreferential repetition in (27) produces links that may be semantic-cohesive as supermarket - supermarket, clean - clean, which together with pragmatic coherent links, supermarket - customers, clean - please, contribute to coherence. But pragmatically deviant co-occurrences are likely to yield incoherent discourse. While S2 provides an "explanation" of S1 in both (27) and (27a), that is, they share the "speech act" of explanation, the interpretation of these examples is bound to be different owing to principles of general sociocultural knowledge, that is, that clean supermarkets attract rather than repel customers.

1.7.2 The expression of 'special' linguistic acts/speech acts through lexical reiteration

The choice of repetition of form and content, that is, verbatim repetition, rather than repetition of content only, often has a cohesive function that is essentially stylistic or aesthetic. That is, it expresses 'special' speech acts. Expressive repetition as defined by de Beaugrande and Dressler (1981) gives some "rhetorical value" to the message and hence expresses speech acts such as surprise or arousal of the interest of the addressee. This does not allow the Economy Principle to operate, but suggests that some other principle is in play. For instance, the repetition of country roads in (28) below seems to be prominently used to assert or reaffirm one's point of view: it is used for insistence. In (29), the repetition of

Sarah Bilton seems to carry the implicature: "Sarah Bilton and no-one other than her will look for the keys".

(28) I told you several times I don't like country roads.
I hate country roads. They're bumpy, bendy and
smelly, your country roads.

(29) Sarah Bilton's lost the locker's keys Sarah Bilton
will look for them, said the teacher.

Other speech acts, as for instance, denial, can be expressed by verbatim repetition. Denial involves "a rejection of the background of the utterance, and will amount to denial of the relevance of the utterance itself" (Smith and Wilson, 1979: 187). For example:

(30) A - Your little boy's really grown.
B - He's not my little boy.

B, by repeating little boy does not interpret A's utterance as an intended compliment to her because she is not the mother of the little boy. She denies this. Sometimes the "principle of least effort" or Economy Principle has to be sacrificed when the interpretation of the discourse is endangered. Then verbatim repetition with anaphoric reference is needed to help the reader identify referents appropriately and comply with the Expressibility Principle (see Section 1.5.3 above). Consider Leech's (1983) example:

(30a) If the baby won't drink cold milk, it should be
boiled.

in which the repetition of milk rather than the injudicious use of the proform it, is necessary:

(30b) If the baby won't drink cold milk, the milk should be boiled.

The Maxim of Reduction can definitely not be applied in this case.

1.7.4 Synonymy as a cohesive category in lexical cohesion

Synonymy or 'semantic equivalence' is one of the linguistic relations that is potentially cohesive. Synonyms are lexical items related through meaning in a particular way. Synonyms are mostly "context-dependent".

For example:

(31) Engineers were installing telephone cables this morning. They laid twenty-two outside the post-office.

Installing and laid are in a paradigmatic relationship in the semantic system, that is, the relationship that a linguistic element has with elements with which it may be replaced or substituted (Palmer, 1976). Also, their syntagmatic cooccurrence produces lexical cohesion. Thus, 'signification' synonyms always produce lexical cohesion in text.

On the other hand, 'value' synonymy is a syntagmatic relationship, that is, the relationship that a linguistic element has with other elements in the stretch of language in which it occurs (Palmer, 1976). For example, non-synonymous lexical items (as for instance those which stand in a paradigmatic relation of hyponymy in the language system) may take on the particular 'value' of synonyms usually when they are used anaphorically. The

notions of 'signification' and 'value' were put forward by Widdowson (1978: 11; 1979: 118) to discriminate between lexical items of the code or system of language and their actual use in communication. Thus, 'signification' hyponyms may become 'value' synonyms in discourse and achieve lexical cohesion. The following examples illustrate this principle:

(31) A new manual has been published on how to be a good interviewee. { The book is intended for students mainly.
 { It " " "

(32) The prince reached a dark cliff. He knew that it was probably the danger his dream had predicted.

In (31) book is the hyperonym of manual but as it has taken on text-determined semantic features such as "being a handbook", it now functions as a synonym of manual in this text. The lexical items manual and book are therefore cohesive 'value' synonyms, not 'signification' synonyms. By the same token, pronoun it will have to take on some text-determined semantic features in order to substitute for manual. However, the definiteness of the NP is not an obligatory feature of 'value' synonymy. Non-anaphoric hyponyms may also take on the particular value of synonyms in communication. For instance, book in (31a) below is a non-anaphoric NP which behaves like a synonym of manual:

(31a) A new manual has been published on how to be a good interviewee. Another book on how to know your boss is due to come out next month.

It is the presence of semantically equivalent and semantically related items that produces lexical cohesion.

In (32) the lexical items, cliff and danger, are unrelated in the semantic system of English but they take on the value of synonyms in this text: the danger is anaphoric to cliff. The interpretation of such relationships often causes considerable difficulty to the language receiver. Sometimes deictics may have as referent not merely a noun but any identifiable matter which can extend over a sentence or a whole paragraph. Consider Hasan's (1968: 58) example:

(32) Most alloys are prepared by mixing metals in the molten state; then the mixture is poured into moulds and allowed to solidify. In this process, the major ingredient is usually melted first.

The process does not form a synonymic link with one particular preceding noun but with several lexical elements occurring in the preceding co-text. However, the interpretation of this type of cohesive relation often poses problems to the reader as it may depend more on his factual knowledge than on his knowledge of his linguistic system. In this sense, such relationship, as exemplified in (32) may be appropriately defined as 'pragmatic synonymy' for it involves 'system' as well as 'non-system' meaning relations between lexical items which appeal to the reader's linguistic and pragmatic competences. Pragmatic general sociocultural knowledge is also a condition of the interpretation of the two following utterances. Unless

the reader has prior knowledge that "lunch" and "1 pm" are events which coincide pragmatically, that is, are equivalent in factual meaning, he will not be able to connect 1 pm to lunch as value, not signification, synonyms:

(33) The playscheme does not provide lunch. Please collect your child at 1 pm each day.

Synonymy, as a potential device of lexicosemantic cohesion, can fulfil some communicative functions, that is, express "speech acts". Let us examine the example below involving conversational discourse.

(34) A - That was a lovely meal.

B - Delicious.

(34a) A - That was a lovely meal.

B - Yes, it was.

B's lexical response in (34) fulfils the function 'agreement'. But is this lexical form "merely phatic", and is the function realised different from the proform in (34a) as McCarthy (1984: 18) remarks? It seems that the use of a synonym rather than a proform by an interlocutor in an oral exchange, reveals more commitment on his part than a proform. But such speculations on the receiver's subjectivity involved in an exchange need

empirical investigation. McCarthy also points out that some 'value' synonymy cannot be easily reversed because of the "coreness" of certain lexical items which in conversational discourse as well as in written discourse is a determinant factor of comprehension. For example:

(35) A - Were you angry?

B - Yes, I was absolutely furious.

(35a) A - Were you furious?

B - * Yes, I was absolutely angry.

In order to perceive (35) as cohesive and coherent and (35a) as pragmatically deviant, the reader is required to be able to appreciate relations of scale and intensity between lexical items, intonation, the marked nature of questions containing non-core items, all of which belong to the use of lexical items in communication. For instance, whether in oral or written discourse, it seems unlikely that angry should follow furious if it is accompanied by qualifier "absolutely" or "extremely", for example:

(35b) * I was furious. I was absolutely/extremely angry.

If it does follow furious as in (35b) it violates Grice's Maxim of Manner and renders the discourse unintelligible. These examples reveal that lexical items have a potential for creating synonymy in text that ought to be exploited. This point bears on some pedagogical aspects which are discussed in Chapter 4.

1.7.5 Hyponymy as a carrier of lexical cohesion

1.7.5.1 Hyponymy in the system

An important component of lexical cohesion involves hyponymy. This is a semantic relation where the more general term, with inclusive meaning, is the hyperonym (or superordinate term), and the more specific, the hyponym. Members of a hyponymic set are cohyponyms.

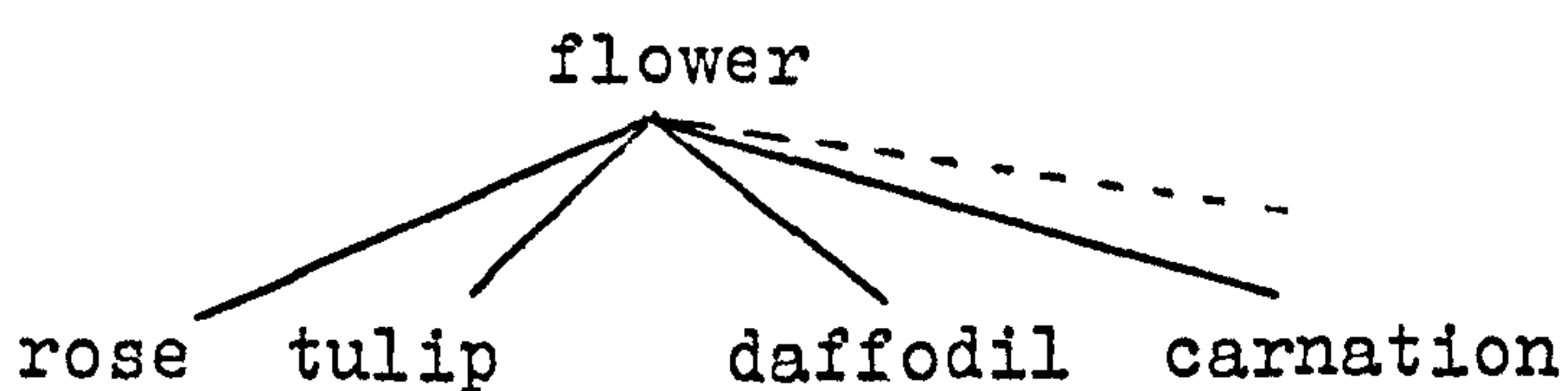
For example:

Rose (hyponym): flower (hyperonym)

Honesty (hyponym): virtue (hyperonym)

Buy (hyponym): get (hyperonym)

A diagram representation of flower would be: (Diagram 3):



Rose, tulip, daffodil and carnation are cohyponyms, that is, members of a hyponymic set.

Hyponyms are defined in terms of 'entailment' (Leech, 1974, 1981). For instance, when we say "I saw a boy" where boy is a "specific" noun, this entails "I saw a child" in which child is a "general" noun. But "I saw a child" does not entail "I saw a boy". The child could be a girl. Likewise specific verbs entail "general" verbs but the opposite is not true: "Harry stole a horse" entails "Harry took a horse", but "Harry took a horse" does not entail "Harry stole a horse".

Harry may have been given it. Therefore, while any hyponym would entail or imply a specific hyperonym, a hyperonym does not imply any specific hyponym.

7.5.2 Hyponymy in use

The occurrence of hyponyms in text/discourse may produce lexical cohesion. For instance:

(36) John bought a budgie. A month later the bird died.

Lexicosemantic cohesion is produced via the cooccurrence of budgie and bird, respectively hyponym and hyperonym. These two lexical items stand in a definitional relationship, and the occurrence of anaphoric reference item the preceding bird marks coreferentiality. Thus, the two lexical items have achieved lexicosemantic cohesion through sense relation (hyponymy) and reference.

Cohyponyms have a strong tendency to cooccur and produce lexical cohesion whether or not they are related through reference, for like synonyms, "continuity of lexical meaning" (Halliday and Hasan, 1976: 320) is a necessary condition for their cohesiveness. Consider the following example:

(37) I've brought you some books. I couldn't find any magazines.

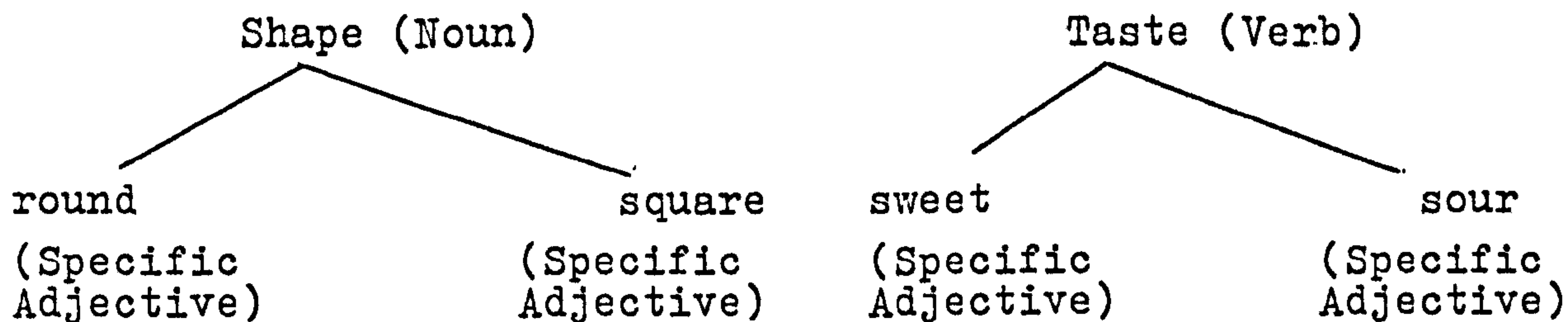
Books and magazines are cohyponyms and their cooccurrence in this text results in (non-referential) cohesion. In effect, the link between them meets requirements to be called 'substitution'. This is a case of lexicosubstitutional cohesion. Furthermore, conjunctive meaning of a causal

type can be deduced from the lexical content of the proposition expressed by text (37): the reader will infer that "buying books" results from "not finding magazines". But this meaning cannot be recovered in the absence of cohyponyms. Compare (37a) below:

(37a) I've brought you some books. I couldn't find any cabbages.

Books and cabbages, not being semantically related, the reader may attempt to relate them pragmatically. But he will not be able to establish a link between them because books and cabbages do not share the requisite semantic feature of 'readability' or any other pragmatic feature. In co-operative terms, Grice's Maxim of Relation has been infringed since the occurrence of cabbages is irrelevant to the reader's purpose: he cannot anticipate cabbages from books on a selectional semantic and pragmatic basis. However, the Quantity Maxim has been observed to the extent that the occurrence of cabbages in S2 is informative, more so than magazines since unexpected.

Lyons (1977: 299) points out that some general "abstract" nouns and some verbs stand in a "quasi hyponymic" relation, that is, a relation which is quasi paradigmatic because it involves different word classes. For instance:



However, Shape (N) and Taste (V) are viewed in this study as true hyperonyms with round and square as cohyponyms of shape, and sweet and sour as cohyponyms of taste because they stand in a definitional relation of meaning with their respective hyponyms. Thus, the state of being round or square is by definition a shape, as the state of being sweet or sour is by definition a kind of tasting. The cooccurrence of such items in text/discourse produces lexicosemantic cohesion, as in (38):

(38) Mary bought a saucepan with an unusual shape. It is square at the bottom and it has a pyramid-like lid.

Square and pyramid-like can be both viewed as cohyponyms of shape. They are shapes by definition and are lexical cohesive items

Halliday and Hasan (1976: 274) have distinguished a class of "general words" which they describe as being on the borderline between lexical cohesion and grammatical cohesion because a "general word" can be a lexical item, that is, a member of an open set, or a grammatical item, that is, a member of a closed system. "General nouns" include people, man, boy, woman, child, creature, business, affair, matter, move, place, thing, question, idea.

"General nouns", viewed by these authors as a certain class of hyperonyms, is said to be more susceptible of introducing an interpersonal element into

the meaning³, like the expression of a particular attitude on the part of the speaker, than would a personal pronoun. But would this personal dimension not occur through the use of hyponymy ipso facto like it would do through the use of personal pronouns? Consider the following examples:

- (39) I've been to York. York is very pretty indeed.
(39a) " " " The town is very pretty indeed.
(39b) " " " The place is very pretty indeed.
(39c) " " " It is very pretty indeed.

It seems to be the case that whether the text producer will use York, the town, the place or it will depend on his degree of personal involvement with the meaning he wants to convey. It seems that the lower down one moves along the scale, from mere repetition of the same lexical item to the use of a grammatical item, the more personally involved the meaning can be. Nevertheless this assumption needs experimental investigation as other factors may be involved in the production of a cohesive lexical item (for example a hyperonym) rather than a cohesive grammatical item (for example a personal (or impersonal) pronoun). This study does not distinguish between "superordinate words" and "general words" in Halliday and Hasan's sense and will refer to them as hyperonyms (or superordinates) indiscriminately. It seems that any hyperonym used coreferentially, that is, accompanied by an anaphoric reference item (the, that, this) is potentially cohesive although non-coreferential hyperonyms may also be cohesive. Coreferential hyponyms:

may take on the particular value of synonyms in text as seen in the previous section. For instance:

- (40) Mr Chugh has opened an Indian restaurant on
The Moor. { The business seems to be doing well.
 { It " " " "

The business, coreferential to Indian restaurant (coreferentiality is signalled by the anaphoric item the), functions as a value/text synonym rather than a signification synonym. This relationship was described earlier as often appealing to the reader's pragmatic knowledge because it is non-definitional: an Indian restaurant is not a business by definition. The combination of hyperonyms plus specific determiner is indeed very similar to a reference item. Thus, substituting it for The business would produce an equivalent cohesive relation with Indian restaurant.

However, lexical items standing in a hyponymic relation do not have to be anaphoric and coreferential to achieve lexical cohesion. For instance:

- (41a) I saw John. The other men had gone to the pits.
Men is not anaphoric to John in a strict sense but a link can be established between men and John pragmatically. John, included in men, is being "repudiated", as it were, by the determiner other which shifts the reference from John to other men.

Thus, the relationship between men and John is exclusive and non-coreferential.

Hyperonyms have a communicative function. They may be used as summarisers of previous stretches of text, as for example, the lexical items, question, point, assertion. For example:

(41b) Without an element of secrecy, British secret services cannot function. The question is how much protection should they receive. This point was on the agenda in the House of Commons yesterday (The Guardian, 1984).

In connection with the use of cohesive hyperonyms and hyponyms in communication, Leech (1983: 91) provides the example of an exchange in which the use of a hyponym can be quite misleading to an interlocutor. In "co-operative"

terms, the infelicitous use of certain hyponyms may violate some of Grice's maxims. Consider the following example:

- (42) Steven: Wilfrid is meeting a woman for dinner tonight.
Suzan: Does his wife know about it?
Steven: Of course she does. The woman he is meeting is his wife.

Following Grice's Co-operative Principle, Suzan is justified in assuming that Wilfrid is not meeting his wife. By using the hyperonym, woman, Steven (whether deliberately or not) has broken the Maxim of Quantity, but he has not violated the Maxim of Quality. Hence, his proposition is true from a logical point of view, that is, wife entails and presupposes woman, but misleading from a pragmatic point of view: the woman asserted in Steven's opening utterance does not implicate wife, hence Suzan's belief that the woman referred to is not Wilfrid's wife. Although this exchange is cohesive (there is a lexico-semantic link between woman - wife and woman), its semantics cannot explain why it is ambiguous. Reference to Grice's "conversational implicature" is necessary.

It is also interesting to note that the choice of a hyperonym in oral discourse restricts the syntactic frame and certain combinations become impossible and others "pragmatically limited" (Cruse, 1977). The semantics of an exchange may be correct but its pragmatics inadequate. For instance:

(43) A - Paperbacks are badly printed nowadays.

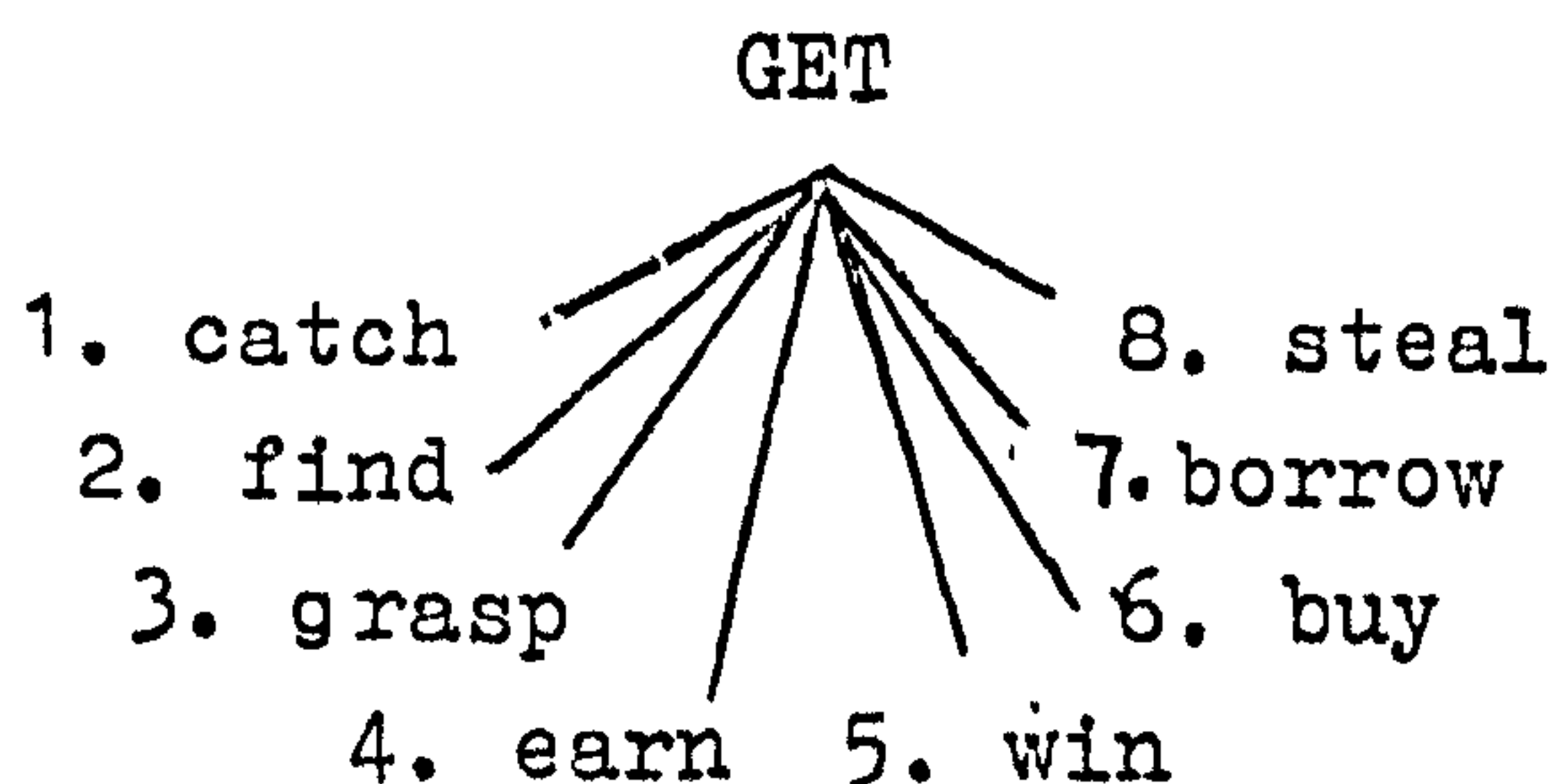
B - * Especially books.

This exchange is uninterpretable because the occurrence of hyponym Paperbacks in A's utterance restricts the syntactic structure of the exchange by determining the type of modifier to be used with the hyperonym books. Thus, modifier all preceding books would be more appropriate to B's utterance (B = All books are) because it would implicate that paperbacks (A's utterance) is included in books (B's utterance) and at the same time, would express the specific discourse function of agreement. That a hyponym cannot anticipate a hyperonym modified by a "restrictive" modifier seems to be a language-universal.⁴

1.7.5.3 Hyperonym verbs in discourse

Some general verbs, as for example, get, move, become, make, act, be, have a large number of hyponyms. For instance:

Diagram 4:



The repetition of hyperonym verbs may produce lexical cohesion, but may also pose problems of interpretation. In written, as in oral discourse, this depends on the type of presupposition shared by producer and receiver, and the conditions of interpretation may be different, as can be seen in the examples below:

(44) Mary got some bread from Leaper's. She got some
for John too.

(44a) A - I'll get some bread from Leaper's. Can I get
some for you too?

B - Yes, please.

For the receiver of (44) there is no way of knowing with precision from the text whether Mary was buying, stealing or borrowing the bread (although meanings 1, 2, 3, 4, and 5, shown in the diagram above, can be discarded on selectional restriction grounds).

The reading of get as buy, borrow or steal is mostly probabilistic. The reader of (44) will tend to read get as buy thereby relying on general pragmatic (socio-cultural) knowledge that people normally buy rather than borrow or steal bread from shops. Although the repetition of got is cohesive, this text is quite ambiguous and to a certain extent, uninterpretable. However, the receiver of A's utterance in (44a) by supplying a positive answer, indicates to A that he shares the presupposition that A will buy or borrow or steal bread. This inference is based on specific rather than general knowledge. We may compare (44) and (44a) with (45a), (45b) and (45c) below in which pragmatic presuppositions are shared by both producer and receiver.

(45a) A - I'll get some bread from Leaper's.

B - Can you buy some for me too?

(45b) A - I'll get some bread from Leaper's

B - Can you borrow some for me too?

(45c) A - I'll buy some bread from Leaper's.

B - Can you get some for me too?

These three examples can be explained in semantic terms by entailment and presupposition and in pragmatic terms by implicature or pragmatic presupposition (see Sections 1.5.1 and 1.5.2 on these notions).

In (45a) get in A's utterance entails and presupposes not only buy (in B's utterance) but also borrow and steal (to name only a few relevant hyponyms). On the other hand, buy in B's utterance entails and presupposes get only (in A's utterance). It does not entail or presuppose borrow and steal. So get and buy form a cohesive link based on the sharing of one entailment (get is entailed by buy) and one presupposition (get is presupposed by buy) only. The same analysis can apply to (45b) in which get and borrow form a cohesive link on the basis of one entailment (get is entailed by borrow) and one presupposition (get is presupposed by borrow).

In (45c), buy in A's utterance entails and presupposes get only, not borrow or steal. On the other hand, get in B's utterance entails and presupposes buy (in A's utterance) as well as borrow and steal. A cohesive link between buy and get can therefore be established by the sharing of one entailment and one presupposition (get is entailed and presupposed by buy)

In pragmatic terms, B uses buy because this was implicated by get in A's utterance in (45a) on the basis of pragmatic knowledge. Likewise, B uses borrow implicated by

get in A's utterance in (45b), also on the grounds of specific pragmatic knowledge. Thus, in (45a) B knows that A will buy bread. In (45b) B knows that A will borrow bread. As a result the pair of lexical items get - buy and get - borrow have taken on the particular value of text synonyms. On the other hand, in (45c) that A will buy bread rather than borrow it or steal it is explicit in his utterance, and implicates get in B's utterance. The cohesive link between buy and get is linguistic semantic and is based on knowledge of the linguistic system of English. Exchanges (45a) and (45b) depend on pragmatic knowledge for their interpretation whereas (45c) involves the receiver's linguistic knowledge.

Cohyponyms, members of conventional sets like those describing seasons, months and days, have a tendency to occur cohesively when one member of the set occurs in one sentence and the other member in another as January - April - September. If they occur in text in non-sequential time order, the receiver's common sense knowledge will help him regard the member of the set occurring first as anterior in time to the second. For instance:

(46) Miss Newman attended a conference in Brussels in July. In January, she was in Frankfurt for an international colloquium on kinesics.

By applying Grice's Co-operative Principle, one can infer that July was last year, and that January is this year and

thus connect both sentences pragmatically. The use of syntactic markers of temporality would then be redundant unless it is to indicate that July is this year and that January should precede in the time sequence as for instance:

(46a) Miss Newman attended a conference in Brussels in July. Before then, in January, she was in Frankfurt for an international colloquium on kinesics.

1.7.5.4 Cohyponyms in discourse

Cohyponyms are members of a hyponymic set. Their juxtaposition in text has a cohesive effect and expresses some important functions, as for example, additivity or adversativity. Thus, relationships of contrast holding in text seems to be primarily a factor of the cooccurrence of some cohyponyms in juxtaposed utterances rather than the result of "parallel structure" (Quirk et al, 1972; James, 1980). Consider the following examples:

(47) It was midsummer festival in the village. Mary wore blue, Jane wore white and Emily red.

(48) Have you ever seen a pig fly? Have you ever seen a fish walk? (Quirk et al, 1972: 716).

The receiver of (47) is likely to assume additivity from the juxtaposition of the cohyponyms of colour, blue, white and red, despite the fact that no signal of additivity (such as and) appears on the surface text. Likewise, contrast is implicit in the juxtaposition of pig and fish,

both cohyponyms of "animal", and fly and walk, both verbs of motion, cohyponyms of "go".

1.7.6 Paronymy as a device of lexical cohesion

There are many lexical items in English and probably in most languages whose meaning cannot be specified independently of some whole - part or part-whole relations of meaning. Some of these relations can be defined in semantic terms. Others are more complex and call for pragmatics for their interpretation.

Lexical items standing in a partitive semantic relation with other items typically show inclusive reference and in this sense they are similar to hyponyms. Thus, body includes arm in the same way as flower includes rose, and animal includes cat. However, by applying the "X is a kind of Y" test, it is easy to see that paronymy (as it is subsequently referred to in this study) implies a different type of inclusiveness of meaning. Thus, rose is a kind of flower, but sleeve is not a kind of garment, or page is not a kind of book and this is where the test can no longer apply. Moreover, the meaning of some lexical items such as second, minute and hour, cannot be explained without specifying the relation holding in the sentence, as for instance, "one hour is equal to sixty minutes" and "one minute is equal to sixty seconds". Like hyponyms, lexical items standing in partitive or paronymic relation in text/discourse may produce definitional as well as non-definitional links as is seen in (49) and (49a) below.

Paronymic relations are powerful devices of lexical cohesion (and coherence). They may involve lexico-referential, lexicosubstitution or conjunctive linkage. The examples below illustrate some of these features:

(49) John couldn't open the door. The handle was missing.

The lexical item handle is a paronym of door and is anaphoric to it. Anaphoric reference is signalled by the reference item the accompanying handle. The cooccurrence of these two lexical items produces semantic cohesion that is lexico-referential: the cohesive link between handle and door is achieved via reference. However, in (49a) below, the link existing between handle and door is not coreferential, but substitutional.

(49a) John couldn't open the door. He needed a handle.

The link between door and handle is not definitional but simply pragmatic: doors usually have handles but not necessarily so (revolving doors, electronically triggered doors, swing doors have no handle). The type of linkage hence produced is non-coreferential/lexicosubstitutional and features lexicopragmatic coherence.

Conjunctive meaning can be inferred from the propositions underlying S1 and S2. Effect-cause meaning ("therefore" could be inserted between S1 and S2) is expressed through the lexical relationship open - door handle. Likewise, conjunctive meaning is also expressed through the lexical pragmatics of (49b) below:

(49b) John couldn't open the door. The letter-box was jammed.

Letterbox is a 'paronym' of door but the link between these two lexical items must be established via pragmatics: the reader has to infer that the key on the string hanging from the letterbox inside the house was obstructed by letters and newspapers in the letterbox and therefore could not be reached by John. The link between these items is not semantic but pragmatic.

1.7.7 Antonymy and converseness as devices of lexical cohesion

1.7.7.1 Antonymy

Antonymy is oppositeness of meaning. It is a factor of cohesion in text. Antonymy can be expressed through binary contrasts which are manifested in antonymic pairs. These contrasts lead to the further distinction between gradable and non-gradable antonyms. Both types of antonyms are distinguished on the grounds of incompatibility and complementarity. The relation of incompatibility is to a certain extent the reverse of hyponymy as it is one of meaning exclusion. In complementarity, to predicate one term is to contradict another.

Gradable antonyms are incompatible and not complementary. The test of negation can show the gradability of lexical pairs such as hot and cold, wet and dry. For instance:

an implicit relationship of contrast which is not overtly signalled by a marker of contrast (for instance, 'but') in this text. Such cooccurrence produces lexicosemantic cohesion.

1.7.7.2 Converseness

Converseness is a lexicosemantic relation where the predication of one term inevitably implies the other. Thus, lexical pairs such as parent and child are converses (Leech, 1981). For instance, "Larry is the parent of Thomas" implies or is synonymous with "Thomas is the child of Larry".

The occurrence of converses in text is potentially cohesive. Links between converses may be established via semantics or via pragmatics. Consider examples (51) and (51a) below:

(51) The doctor was called urgently. The patient was having another heart attack.

(51a) The Washington Post reports another successful heart transplant. The patient is a 50 year old bus conductor and comes from Missouri.

In (51) the connection achieved lexically is semantic. Doctor and patient are converses. The patient in S2 is the patient of the doctor mentioned in S1. It is anaphoric to doctor. The link between doctor and patient is definitional and instantiates lexicosemantic cohesion. In (51a), however, the connection achieved via heart transplant and patient cannot be explained in semantic terms. Heart transplant and patient are not converses.

Rather they share an element of meaning which is pragmatic. There is a relation of coreference as well as a relation of correspondence of meaning or "co-semy" (Leech, 1981: 190) that can be restated as the repetition of part of the pragmatic content of an utterance in another utterance. Thus, part of the pragmatic content of heart transplant (that it involves a doctor and a patient) is being repeated in patient. The co-occurrence of these lexical items in text produces a link or tie of a pragmatic nature and thus instantiates lexicopragmatic coherence.

1.8 Functional dynamism in lexical cohesion

After having examined the semantic and pragmatic resources of lexical cohesion in text, we are now in a position to see how functional dynamism operates in the discourse using these resources. Unlike sentence-based grammars which seem to have neglected the relation of cohesion, the Functionalists of the Prague School, such as Firbas (1964) and Vachek (1964) have concerned themselves with this type of relation and have viewed anaphora as the linking device between cohesion and Functional Sentence Perspective. In Britain, Halliday's textual macro-function, to which cohesion belongs, owes much to the Prague School's Functional Sentence Perspective (FSP) which describes the flow of information through sentences. Thus, thematic progression is the core idea of FSP analysis. It was developed by Daneš (1974) and claims that sentences are held together through thematic links in text. These links are made up of a "theme", which

carries "old" or "given" information and of a "rheme" which carries "new" or "unknown" information. The theme is the most important part of a clause from the point of view of its presentation of a message in a sequence. The process of thematic progression involves the shift of the theme or of the rheme of one sentence into the theme of a later sentence. In this way the themes or rhemes of sentences help the message move forward giving the discourse its functional dynamism. Rhematic elements conveying new information will show higher degrees of "communicative dynamism" (or "the extent to which the sentence element contributes to the development of the communication, to which it "pushes the communication forward", as it were" (Firbas, 1964: 270), than those thematic elements which convey old information. In English and French, two SVO languages, the Subject is normally theme and is in initial position, and the Object is rheme. The Verb is the Transition. In Arabic, a VSO language, the Subject is theme and is either in medial or initial position and the Object is rheme. The Verb is in initial position. For instance:

(53) John likes apples.

(53a) Jean aime les pommes.

S	V	O
Theme	trans- ition	Rheme

(53b) (juḥibbu)(muḥ ammādu)(ʔattuffa:ḥa) = Mohamed likes apples.

(medial position) V S (theme) O (rheme)

(muḥ ammādu)(juḥibbu) (ʔattuffa:ḥa) = Mohamed likes apples.

(initial position) S (theme) V O (rheme)

FSP is marked differently in different languages. In Czech, the language on the basis of which FSP analysis was elaborated, word order is the essential marker of FSP.⁵ Word order in English is less obvious in FSP as other syntactic/stylistic devices such as inversions, passivisation, clefting and pseudoclefting, and marking of definiteness in written discourse (higher pitch and stronger stress in oral discourse) are usually exploited to organise information flow in text and produce cohesion. However, in all these "movement transformations", word order is also involved and although they are largely intraclausal, these movements are still determined by interclausal forces of intersentential cohesion. For example:

- (54) Inversion or the fronting of the object in the sentence placing focus on it:
- (a) I can't stand mushrooms, and I simply hate gherkins.
 - (b) Mushrooms, I can't stand and gherkins, I simply hate.
- (55) Subject inversion involving lexical substitution:
- (a) The rain came after the sun, and everybody left the beach.
 - (b) The sun preceded the rain (but only for a short while).
- (56) Passivisation:
- (a) John bought these flowers, and Julia this cactus.
 - (b) These flowers were bought by John and this cactus by Julia.

(57) Clefting and pseudoclefting:

- (a) Chas made a ragdoll for his daughter. He called it Jemma.
- (b) It was a ragdoll that Chas made for his daughter and called Jemma. He didn't make a glove puppet .
- (c) What Chas made for his daughter and called Jemma was a ragdoll (he didn't make a glove puppet).
- (d) A ragdoll is what Chas made for his daughter and called Jemma (he didn't make a glove puppet).

(58) Definiteness:

- (a) A man entered a hotel and asked for a single room. The manager was deaf.
- (b) Two policemen were following a man with a dark suit. The man entered a hotel.

In example (54), mushrooms, Object in (a), has been fronted in (b) hence placing focus on it: from rheme in (a), mushrooms has become theme in (b). In (55a), rain is theme and the end-focus is on the sun because it occurs after after. Sun is rheme. By fronting sun in (b) which involves lexical substitution (after is replaced by preceded) special focus is placed on it and makes it thematic. Rain is rheme.

In (56), the passivisation of John in (a) shifts the focus to flowers. In (57), the cleft sentence (b) highlights ragdoll which has the full implication of contrastive focus: the rest of the clause is taken as given and a contrast is inferred with other items which

might have filled the focal position in this sentence. Thus, sentence 1 in (a) has an implied negative which is made explicit in the following sentence in the presence of glove puppet which indeed contrasts with ragdoll. So from (a), it has been possible to derive (b) hence highlighting the element ragdoll but of course other elements of the clause could have had the same treatment, that is, daughter or Jemma.

In (58), although both occurring in initial position and therefore thematic, a man in (a) does not need to be recovered from previous text/discourse because it is indefinite. But antecedent context is an essential condition for the interpretation of the man in (b).

The dichotomy between theme and rheme may be paralleled with the semanticopragmatic distinction between presupposition and assertion. Any utterance belonging to a discourse tends to contain elements of meaning which are presupposed in the sense that they are already part of the "pragmatic universe of discourse" (Kempson, 1975: 167) and which correspond to the "theme" of the discourse, and elements which are asserted in that they are not part of that context, that are new, and therefore from the "rheme" of the discourse. For instance:

(52) Tea prices in the shops are set to rise by another
4 pence a packet.

Tea prices (in the shops) are presupposed elements in this utterance. They are already part of the 'context' and therefore 'thematic or given.' Are set to rise ...

are asserted elements, not part of that 'context.' They are rhematic or new. In connected discourse the theme contains presupposed information and is recoverable from the preceding sentence: it is anaphoric. In the rheme the writer asserts information that is new. It is therefore not recoverable from previous discourse.

How does functional dynamism operate in text to produce lexical cohesion? Consider this example:

- | | | | | | | |
|------|-------------------|----------------|------------------|------------------------|---------------|---------------------|
| | T1 | R1 | | R2 | | R3 |
| (59) | Arthur/ | set out | in his best | <u>suit</u> / | on the | <u>road</u> /to his |
| | | R4 | | | | |
| | <u>sister's</u> / | early | in the | <u>evening</u> / | | |
| | T1 | R1 | | R2 | | |
| (a) | It was/ | impossible/ | to get the | <u>sports-jacket</u> / | | |
| | R3 | | | | | |
| | clean | in time/ | | | | |
| | T1 | R1 | R2 | | | |
| (b) | He/ | aimed/ | to arrive | by | <u>six</u> / | |
| | T1 | R1 | | | R2 | |
| (c) | She/ | was expecting/ | the whole | <u>family</u> / | | |
| | T1 | | R1 | | | T2 |
| (d) | In some | places/ | the | <u>drains</u> | were blocked/ | and there |
| | | R2 | | | | |
| | were/ | huge | <u>puddles</u> . | | | |

By looking at theme and rheme parsing in this text as well as focal prominence, the reader may be able to connect (a), (b), (c), and (d), to (59). (59) has one theme and four rhemes. (a) can follow rheme 1 because it has one linguistic clue, that is, the cohyponym sports-jacket which produces a cohesive link with suit. (b) can follow rheme 4 because a link can be inferred between six and evening via pragmatic principles: the reference of evening implies six pm rather than six am. There is also a

grammatical link between He and Arthur which is anaphoric and coreferential.⁶ (c) relates to rheme 3 via grammatical cohesion as She corefers to sister's (thematic in (c)), as well as lexical cohesion on account of the lexical link between sister and family (partitive relation). (d) can follow rheme 2 because of the linguistic link produced by the partitive relationship between road and drains and puddles, and also by the relationship between road (specific) and places (general) which is of inclusive/hyponymic type. Thus, most of these lexical ties are two-place ties, but the occurrence of Arthur - sister - family produces a three-place lexical cohesion tie.

Example (59) thus gives evidence that the production of lexical or grammatical cohesion is not restricted to a particular thematic or rhematic pattern: themes may be anaphoric to preceding themes or rhemes, or can substitute for themes or rhemes:

Arthur	(T1)	He	(T1)	grammatical cohesion
Suit	(R1)	Sports-jacket	(R2)	lexical cohesion
Road	(R2)	Drains ... puddles	(R1) (R2)	lexical cohesion
Sister	(R3)	Family	(R2)	lexical cohesion
Evening	(R4)	Six	(R2)	lexical coherence
Road	(R2)	Places	(T1)	lexical cohesion
Sister	(R3)	She	(T1)	grammatical cohesion

Lexical cohesion and lexical coherence were produced via repetition of themes and rhemes. Also via shifting of rhemes into themes.

While it is difficult to draw a general conclusion from these examples of thematic progression, it is nonetheless interesting to note the tendencies of themes and rhemes to generate one type of cohesion rather than another. Thus, grammatical links tend to hold across themes and between one theme and one rheme; lexical links seem to hold mostly across rhemes, and between one rheme and one theme. But this again depends on the type of text involved. Some themes may be recovered from the information stated in the preceding cotext which includes thematic and rhematic elements of information. Thus, in the example below, theme this fact refers to the whole sentence preceding its occurrence.

(60) Mr Torrino has adapted his delicatessen shop into a shop restaurant where you can buy food and wine throughout restaurant hours and consume it at shop prices. This fact was revealed to us by a respectable couple who at the end of the meal simply corked their bottle of wine and took it home.

So it appears that consideration of thematic progression in the definition of lexical cohesion is quite inconclusive to the extent that the achievement of lexical cohesion seems to be independent of the distribution of themes and rhemes in text. However, anaphoric themes tend to be more easily recoverable than rhemes, and this

has pedagogical implications as will be seen in the last chapter. In a cloze passage, an examination of relationships of lexical cohesion holding in text may help the reader recover themes but rhemes often need world-knowledge to be identified. For instance:

(61) The fox passed, all muscles tensed. The _____
could hear the rustling sound of the _____ and
decided to move towards the sound.

The missing word in S2 (creature) has a thematic position, and the presence of the anaphoric reference item the signifies that the missing meaning must be recovered from the preceding sentence. Here, fox, a hyponym of creature, is theme in S1, and is shifted in S2 as a thematic hyperonym. On the other hand, the missing item grass is in rhematic position and is less easily recoverable; its collocation with rustling may contribute to its identification. In other words, while the recovery of the theme creature in S2 was facilitated by intersentential lexicoreferential cohesion, the identification of the rheme of the same sentence (grass) was a function not of intersentential cohesion, but of intrasentential cohesion, specifically collocation (rustling grass).

1.9 Concluding remarks

The concept of cohesion achieved via lexical means has been defined in this chapter in relation to two notions which complement each other, viz. text and discourse. Appeal to these notions was necessary to our discussion which viewed lexical cohesion as a relation within

the lexicosemantic system of English, and the pragmatic features of which have been described as attributes of lexical coherence. In coherence relations there is absence of lexical signals of cohesion on the surface text. It is also worth recalling that the distinction between linguistic-semantic and non-linguistic-pragmatic knowledge underlying lexicosemantic cohesion and lexico-pragmatic coherence does not imply any sharp distinction between them. The next question to be answered is, how can cohesion theory explain processes involved in text/discourse comprehension? Halliday and Hasan's study of cohesion is a competence model of cohesion and was criticised precisely because it did not account for the reader's text processing (Moe, 1979). Halliday and Hasan's system for analysing texts in terms of numbers and types of cohesive ties allows for the quantification and the identification of types of cohesive ties in text, but it does not determine the 'strength' of the tie or the degree of binding which links semantic and pragmatic relationships. Such system was intended to be used only for linguistic analysis, as Gutwinski (1976) rightly remarks, and does not make provision for a psychological analysis. Halliday and Hasan's view of cohesion as an "epiphenomenon of content coherence" (Morgan and Sellner, 1980) has been criticised by psycholinguists (Carrell, 1982) on the grounds that it assumes cohesion as the source of coherence. The analysis of the psychological factors underlying the interpretation of lexical cohesion is the object of the next chapter.

Notes on Chapter 1

- 1 Leech (1983: 17) calls "sense", "meaning as semantically determined" and "force", "meaning as pragmatically, as well as semantically determined".
- 2 On this point, it may be noted that Halliday and Hasan's (1976) matrix of cohesive devices which involves repetition of "same word" or "same item" is rather vague. Does it mean "same" from a semantic view point only? In that case it is redundant with synonymy on their next level below. Does it refer to sameness of form as well? Halliday and Hasan's categories are organised from a semantic point of view but the first entry "same word" is not clear. The example they provide reflects sameness of form and meaning: "I turned to the ascent of the peak. The ascent was perfectly easy", a category referred to as verbatim repetition in this study.
- 3 An attempt to clarify the distinction between superordinate and general words was made by Hottel-Burkhart (1981: 41):

"Superordinates are limited to the lexical items which fit the phrase 'an X is a kind of M', where X is the original lexical item and M is the superordinate. A general word is a kind of filter word (...) Halliday often used in evaluative statements, as in that old thing or creature, in the comment, 'Mop? That creature was the best mouser I ever owned'" (quoted in R C Scarcella (1984: 25)).

- 4 Note that neither French, nor Spanish or Arabic can accept a hyperonym preceded by "especially" as a response to a statement in which a hyponym occurred.

For instance:

French: A - De nos jours, les livres de poche sont très mal imprimés.

B - * Surtout les livres.

Spanish: A - Estos días los libros en rustica estan mal imprimidos.

B - * Sobreto los libros.

Arabic: A - ?inna ?elkutuba ?e SSa?iRat
?el?ad?mi fi ?aja:mina ?a?i?i
ta?Rifu Tab?atən Radi:?atən

B - * XuSu:Sən ?elkutubu .

In all three examples there is lexical cohesion but as the pragmatics of the interchange are inadequate, it is incoherent as French, Spanish or Arabic discourse.

- 5 Note how theme fronting (or "marked theme") differs in English, French and Arabic.

English: Apples, John likes : Theme is in initial position.
Theme

French: Les pommes, { Jean aime ça! : Theme is in initial
Theme { Jean les aime position.

Arabic: ?inna ?attuffa: ?a ju?ibbu?a mu?ammadu.
Theme (introduced by intensifier "?inna"):

Theme is in medial position.

(literal translation: apples he likes them
Mohamed)

6 Native speakers of English tend to find that the retention of a theme produces a more cohesive and pragmatically acceptable text than its shift as rheme. For instance:

(59a) The child ran into the road and he was hit by a car.

(59b) The child ran into the road and a car hit him.

Thus, in (59b) the shifting of the theme the child to rhematic position in S2 (him), that is, making the sentence active thereby making the 'doer' known, seems to disturb the reader's 'feel of coherence' in the text.

CHAPTER 2

LEXICAL COHESION IN READING COMPREHENSION

"No matter how well a linguistic description seems to capture the meaning of an utterance or the structure of a discourse, it does not describe the way in which people understand" (Sanford and Garrod, Understanding written language, 1981, p61).

Introductory notes

By introducing this chapter with a quotation by Sanford and Garrod (1981) our aim was to point out a contrast. The present chapter treats lexical cohesion, not from a purely linguistic (competence) point of view, but from an interpretative angle involving performance phenomena.

The previous chapter, devoted to the linguistic analysis of the concepts of lexical cohesion and lexical coherence in English, looked at what is "there" and "not there" in the text in terms of linking devices (semantic and pragmatic) of cohesion and coherence, often coexistent in text.

Text was defined as a "semantic edifice" (Halliday and Hasan, 1976: 26) and discourse was viewed as a "pragmatic edifice" to adapt Halliday and Hasan's metaphor in which pragmatic meaning was communicated via relationships of coherence. Our analysis was mostly based on Halliday and Hasan's taxonomy of semantic lexical cohesion, but for the interpretation of lexical

cohesion within pragmatics, we have drawn on Leech's influential study of the principles of pragmatics. In this chapter we focus on performance aspects and propose a psycholinguistic view of lexical cohesion within the process of reading (writing). Reading involves three entities, the reader, the writer and the text. Two of these entities, the reader and the text, will be focussed upon in this study but the third element of the reading process, the writer, although important and influential, will receive relatively less attention. The text can only have "potential for meaning" (Widdowson, 1979). It is best viewed as a "set of directions" which indicate to the reader where he must look in his linguistic and experiential world for the producer's meaning. If he understands these directions and is capable of carrying them out, then he will be successful in his comprehension of the writer's message. But, reading as discourse comprehension cannot be viewed as a reaction to text but as an interaction between the reader and the writer mediated via the text. It is creation of a discourse whereby the reader's interpretation of text often involves "a mixture of sense selection and sense creation" (Clarke and Gerrig, 1983: 605). During "sense selection", the reader selects a conventional word meaning from a list of entries in his mental lexicon which represents his lexical-semantic competence and in "sense creation" the reader creates a word meaning by referring to his encyclopaedic or world-knowledge and thus builds up his lexical pragmatic competence.

Unlike conversational discourse, the interaction between the producer writer and the receiver reader cannot always satisfy Grice's Co-operative Principle. In conversational interaction the speaker knows who his interlocutor is. In writer-reader interaction, the writer does not know who his specific reader is. Hence there may be problems for the writer, in for instance, satisfying Grice's Maxim of Manner ('Be perspicuous, clear and unambiguous') which ensures clarity of the writer's message. In terms of a writer producing a text/discourse, he must predict that the clues (lexical and grammatical) he supplies in his text will be accessible to all types of readers. But such prediction can only be probabilistic for the exploitation of clues in text/discourse can only be hoped for by the writer and can by no means be predicted with certainty. Thus, from the point of view of writer-reader interaction, the writer is bound to violate the Maxim of Manner, specifically when his text/discourse addresses an unspecified audience. However, the "specialist" writer who writes for a specialist readership is more likely to satisfy the Manner Maxim by supplying text clues and is almost guaranteed to be interpreted appropriately.

Understanding the writer's "set of directions" in text / discourse is subordinated to the reader's possession of a certain amount of "background knowledge" . This notion has received various treatments depending on researchers' expertise. Cognitive psychologists analyse "background knowledge" in terms of knowledge structures

which include knowledge of an individual's language. Psycholinguists and linguists view it as non-linguistic knowledge essentially ('world knowledge'), which, unlike linguistic knowledge, is often not shared by reader and writer, and this may be mostly problematic to the non-native reader.

Our aim in this chapter is two-fold:

- a. to analyse the phenomenon of lexical cohesion in the light of the psycholinguistic theories of reading and theories of knowledge in order to capture the way this linguistic phenomenon is processed by human readers.
- b. to examine experimental evidence brought in by cognitive psychologists and psycholinguists on the processing of lexical relations of cohesion and coherence from a linguist's viewpoint.

2.1 Reading in NL and reading in FL

2.1.1 Top-down and bottom-up reading

It is generally agreed among psychologists and psycholinguists that reading consists of a combination of two processing modes, top-down processing and bottom-up processing. However, some psycholinguists view it as a top-down activity essentially reflected in most native adult reading. Thus, Goodman (1973: 22) defines reading as a "psycholinguistic process by which the reader (a language user) reconstructs, as best he can, a message which has been encoded by a writer as a graphic display". This definition implies that the reader anticipates what information will occur in the text and as a result reading cannot be an exact process which depends upon

accuracy or precise identification of all elements of the text, but rather, as a process of hypothesising, hypotheses testing, confirming or rejecting and repeating this cycle until the reader is satisfied with his "guesses". This "psycholinguistic guessing game" (Goodman, 1967)¹ is done in ways which make sense to the reader depending on self-defined purposes (for example, reading for a particular piece of information in a magazine, or reading for pleasure where acquisition of information is of ancillary importance to the reader) or externally defined purposes (as for example, reading in class as part of an academic activity). Whatever the reasons underlying the act of reading, this "guessing game" involves "strategies" upon which the language user relies to produce the most reliable prediction with the minimum information that he can extract from the text. Reading in native language has also been viewed as "reasoning" (Thorndike, 1974) and the most fundamental input to reasoning is existing knowledge (in a general sense of the term) and the way this knowledge is used to interpret a text: we use what we know in order to make sense of what we do not know and to increase our total knowledge. That is why native reading has often been described as "externally guided thinking" (Neisser, 1976) whereby confirmation via the text is hardly needed. Thus, top-down processing reading involves the prediction by the reader of what the next and possibly other sentences are likely to mean on the basis of higher order or general "schemata" implied by the major salient parts of a sentence. Top-down reading has often been referred to as "conceptually-driven" reading (Carrell, 1983: 82).

But reading also involves bottom-up processing, that is, the working out of the meaning of lexical and grammatical items and of the structure of the sentence and the building up of a composite meaning of the sentence. This has been described as "data driven" processing (Carrell, 1983: 82). Intensive processing of textual signs, however, seems to characterise children's native reading (when learning how to read) and some non-native reading generally (as for instance in a second or foreign language). Then the text is heavily utilised for building up and confirming hypotheses. In effect this type of processing allows little hypothesising to take place because the reader is more involved in decoding than in hypotheses building.

To summarise, reading as processing of information from text involves bottom-up and top-down processes which should be operating "at all levels of analysis simultaneously" (Rumelhart, 1977): when we read the first line of a text, we attempt to build some composite meaning for the line we read, on the basis of its structure and the meaning of the vocabulary items involved, at bottom-up level. At the same time we operate an interpretation strategy which involves anticipating what is likely to come next, at top-down level. In Goodman's view of NL reading, top-down processing is the essential part of reading. It is one which, we believe, is most

difficult to achieve by non-native readers because it requires a previously formed knowledge structure which already contains the major relationships, as for instance, those which organise a paragraph. These relationships need "higher order schemata." However, "minor" relationships of cohesion which also organise a paragraph need "lower order schemata," as will be seen below. The nature of these schemata and the way they account for the role of lexical cohesion in the reading process are treated in Section 2.2 below.

2.1.2 Reading in NL versus reading in FL

Whether it is done in native language or in foreign language, reading is therefore an active generative process whereby meaning is attributed to the words on the page, although most research on reading comes from studies on native readers (children and adults). The question of whether potential non-native readers process textual signs differently from native readers is an issue which has received much attention on the part of educationalists interested in FL teaching. Cowan's (1976) investigation which included Japanese and Persian subjects suggested that reading in a FL may be impeded by the learner's application of "perceptual strategies" in the NL, that is, "the cognitive principles used in mapping external representations onto internal sequences to achieve comprehension". This implies that reading in NL involves processes that cannot be transferred from one language to the other without impairing the FL reading

process. Expectancies set up by the reader are said to be "language specific" and when applied to reading in a FL, presumably competing syntactic and lexical processes occur and give rise to confusion and comprehension breakdown. Cowan's subjects source languages were structurally different from their target languages and this may have been the cause for confusion: the subjects were Japanese native speakers reading English, Persian native speakers reading English and English native speakers reading Hindi. Hence one is led to believe that languages not so apart in their syntactic and lexical structure, as for example, English and French, or Italian and French, would be less likely to require specific perceptual strategies. It is quite conceivable that the up to down and right to left text processing required for Japanese and Arabic represent potential sources of difficulty to learners because a different directionality is involved when processing English text. But these reasons do not seem sufficient to justify poor performance by FL learners whose source language is structurally different from the target language² Reading in NL and reading in FL may be compared on the following four conditions but the difference between each language and within each condition does not imply a strict separation between them. In effect, it allows some degree of overlapping.

1. In NL the reader reads for pleasure and has no subsequent questions to answer. In FL, texts are typically followed by questions to check comprehension.

2. In FL the reader reads in a classroom situation which makes him more text-conscious than in NL. He can read in class or at home/work in his NL.
3. The type of texts one reads in a FL are in general graded from a language viewpoint and have specific 'genres'. In NL one can read any type of text. There are no limitations on the language or the 'genre' of the text.
4. In NL the reader uses skimming and inferencing in an automatic fashion, and probably more than in FL. Consequently, a reader can compare his performance in NL and in FL and may find that the former is higher. Skimming is a reading strategy not usually utilised in FL where condition 1 above is required, but inferencing is vital when used appropriately because it enables the learner to "approximate" meaning, although this situation is not always appreciated by language instructors. This point is discussed further below.

Thus, although the material conditions in which reading takes place in NL or in FL are different, it cannot follow that the reading process itself is ipso facto different. Whatever the evidence supporting or invalidating the view that reading in a FL is not different from reading in a NL, when dealing with the reading ability or 'skill', FL instructors should not lose sight of the fact that (adult) learners already possess the ability to read in their NL, which provides

them with a 'skill' that children have to learn afresh.

Reading has traditionally been analysed into a series of subskills. Davis (1972) identifies eight subskills involved in reading comprehension which include the ability to understand or recall word meaning and facts from a passage and the ability to get the gist (or main idea) from it. But more crucially, one of the most productive 'skills' required in reading in NL and in FL is the ability to make inferences. Much of the information conveyed by a text is not conveyed directly as the literal meaning but indirectly as inferences which are drawn from the text and which together with information brought into the text by the reader, contribute to the meaning of the message. What characterises this information brought by the reader into the text during the inferential process?

2.2 A discussion of the notion of 'background knowledge' in relation to lexical cohesion

In order to be able to analyse the characteristics of the inferential process in an FL and to see the extent to which it can account for the interpretation of lexical cohesion in discourse, it seems essential to enquire into the nature of "background knowledge" which is one of the elements involved in the process of inference making.

When readers cannot establish links between various textual elements often because they encounter unknown items (which are a lexical index of the difficulty of the subject matter), they instantaneously have a feeling of discontinuity. This is because continuity that was

assured by their adherence to a particular type of information or knowledge has stopped. How is this particular type of information organised in the reader so that it enables him to perceive a discourse as cohesive and coherent and to re-establish continuity when obstacles are encountered in reading, is a question being presently examined. Research on the psychological processes involved in comprehension dating back from Bartlett (1932) and Piaget (1955) have evidenced that understanding something is a function of an individual's past experiences (which include linguistic and worldly experiences) globally referred to as his "background knowledge". An individual's knowledge of the world as well as his theories about it are constantly building up. New information is entered in the system which relates to old information already in the system and this is part of comprehending what one reads. A new fact becomes part of an "organised mass of experience" (Bartlett, 1932: 206). So we understand what we read in a text only when we can relate it to something we know, to an existing knowledge structure or a "schema" and it often seems to be the case that "the question of how people know what is going on in a text is a special case to the question of how people know what is going on in the world at all" (de Beaugrande, 1980: 30).

All individuals possess schemata. These "interacting knowledge structures" (Rumelhart and Ortony, 1977: 100) are said to be stored in hierarchies in long term memory. They have been shown to guide the comprehension of events

and activities, as for instance, going to a restaurant, visiting the doctor's, attending a meeting, celebrating the New Year, and subsequently the interpretation of the linguistic representations of these events and activities. This aspect of "schemata theory", particularly relevant to the processing of English text, could demonstrate how the processing of English as a NL is the same as or different from the processing of English as a FL. Thus, it could demonstrate how certain content and formal schemata, are accounted for by "frames" (Minsky, 1975), "scripts" (Schank and Abelson, 1977), "macro-structures" (van Dijk, 1977), "expectations" (Tannen, 1978) and "scenarios" (Sanford and Garrod, 1981). These terms are not identical but they share some fundamental assumptions and give insights into discourse comprehension in general, and specifically into what goes on in the reader's mind when attempting to inference unknown meaning while reading.

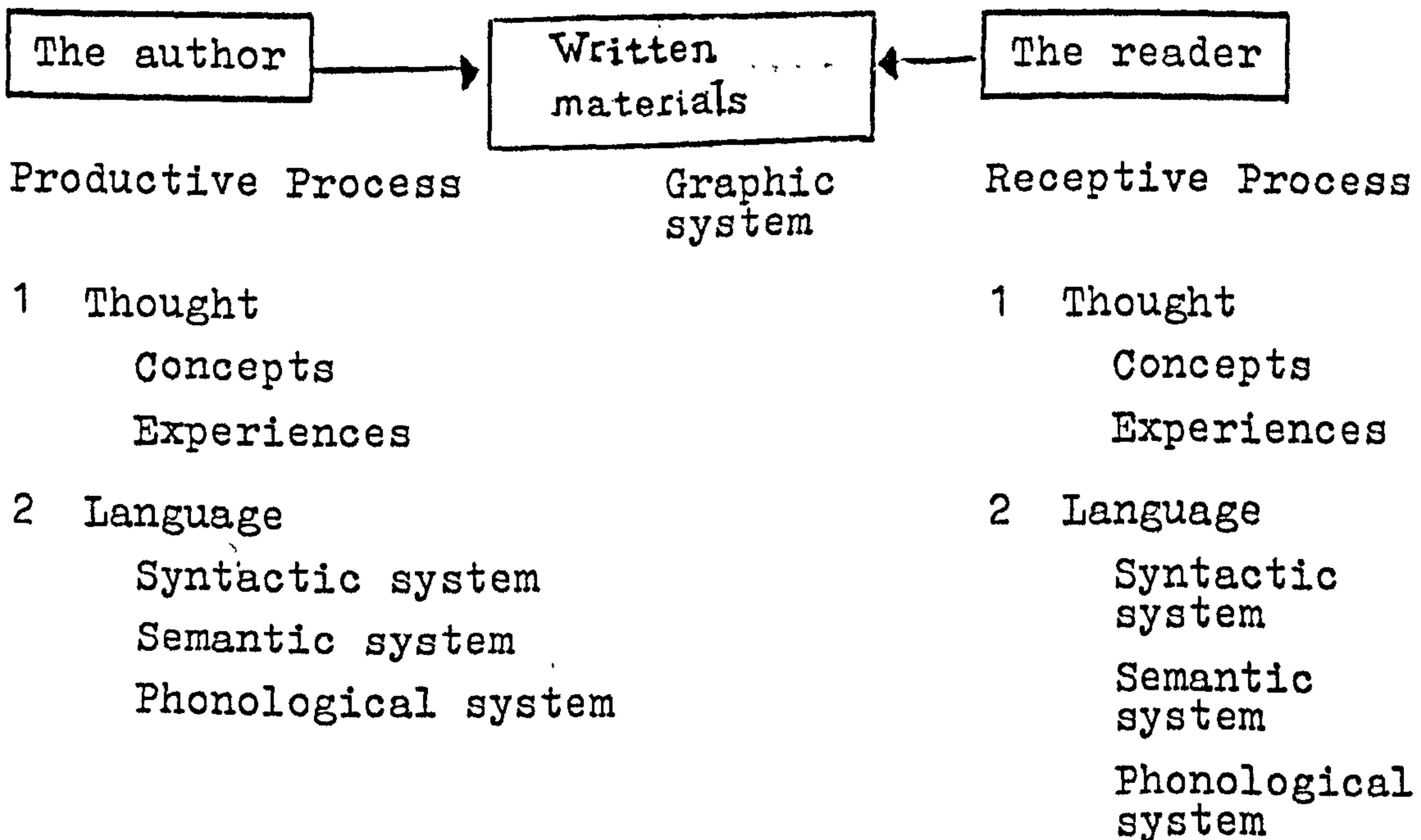
2.3 Content and formal schemata

The distinction between "content" and "formal" schemata was suggested by Carrell (1983) to account for knowledge of "content" and knowledge of "form". Content schemata concern the background knowledge of the content area of a text. They correspond to the specific topic of a text and underlie surface cohesion. They are the "building blocks of cognition" (Rumelhart, 1980).

Formal schemata concern the rhetorical structures of different types of texts (that is, the expected story and text structures). Each type of text, as for instance,

a story, a scientific report, a poem, a newspaper report, has its own convention about structure and knowledge of these conventions and can help the reader understand a text as well as recall it later. These expected story/text structures are structures said to be internalised by the native speaker as generic for different types of texts. For instance, the native speaker reader's schema for an English simple story includes his knowledge that the story will have at minimum a setting/beginning, a development and an ending.

Menosky's (1976: 102) diagram seems mostly appropriate as a summary to this section. It clearly demonstrates the inter-relationship of three elements, the author, the text and the reader, possessing two types of knowledge, pragmatic and linguistic.



The author and the reader contribute varying language structures and experiential backgrounds to the reading process. The reader has to predict the syntactic and semantic structure which the author intended. The author must make these structures explicit. The next section analyses some content schemata and discusses conditions of their activation in the native and foreign reader. It also examines the extent to which lexical relations of cohesion and coherence holding in text/discourse activate specific content and formal schemata.

2.3.1 Frames

The 'frame' is a notion introduced by Minsky (1975). It is a type of content schemata that is activated while reading. Rather than paraphrase Minsky's (1975) description of a frame, we shall supply this long quotation by Minsky which gives a clear and comprehensive account of discourse frames:

"When one encounters a new situation (or makes a substantial change in one's view of the present problem) one selects from memory a substantial structure called a frame. This is a remembered framework to be adapted to fit reality by changing details as necessary.

A frame is a data-structure for representing a stereotyped situation, like being in a certain kind of living-room, or going to a child's birthday party. Attached to each frame are several kinds of information. Some of this information is about how to use the frame. Some is about what one can expect to happen next. Some is about what to do if these expectations are not confirmed. We can think of a frame as a network of nodes and relations. The 'top levels' of a frame are fixed, and represent things that are always true about the supposed situation. The lower levels have many terminals - 'slots' that must be filled by specific instances or assignments must meet. (The assignments

themselves are usually smaller 'subframes'). Simple conditions are specified by markers that might require a terminal assignment to be a person. More complex conditions can specify relations among the things assigned to several terminals" (Minsky, 1975: 212 - his emphasis).

The following is an illustration of Minsky's concept of frame: in a frame representing a 'HOSPITAL', there will be "terminals" or "slots" that will be filled by specific data such as "doctor", "nurse", "medicine", "treatment", "bed", "illness", "operation", and a particular hospital existing in the world or mentioned in the text would be "instantiating" the hospital frame and could be represented by filling the "terminals" with the particular features of that individual hospital. It should be noted that Minsky's discussion of frames is not primarily concerned with linguistic phenomena as it investigates visual perception and visual memory phenomena, but it is centred on a way of representing knowledge, and knowledge of a language is one kind of knowledge that can be represented by frames. Fillmore's case grammar sentence analysis involves structures resembling frames: parts of a sentence are centred mainly around the verb and are therefore used to instantiate a sort of verb frame. But in discourse, as sentences are understood, the resulting sub-structures must be transferred to a growing "scene frame" to build up the larger picture. Brown and Yule (1983) remark that the unfortunate but nevertheless logical outcome of a frame theory of how we use our stored knowledge is that it predicts that a lot less discourse should occur than actually occurs. However, "there are many situations in

which discourse is produced where the intended audience can be expected, but not guaranteed, to have stereotypic knowledge of what is to be communicated" (p240). Indeed, a large part of what a producer communicates is non-stereotypic knowledge and often a discourse becomes unintelligible to a reader when the writer's expectations about his reader's stereotypic knowledge translate into a non-explicit type of discourse. The discourse produced should represent the information in a form which serves as a "reminder" to the reader who already knows and in a form that serves as an "instruction" to the reader who does not know. Thus, in a FL learning situation, learners may be expected to possess the "top levels" of a frame described by a text but may not be able to access its "lower levels" because of problems or gaps relating to their competence of the language. The title of a passage may explicitly activate a frame in the reader but text comprehension occurs at a level more complex than that of knowing what frame is involved. In language teaching terminology, a title calling up a frame will enable the reader to start with a 'general idea' of the passage and sets up expectations in him, but unless he can handle further textual information which will call up sub-frames, the reader's discourse comprehension will not go beyond that initial top level frame.

Sometimes, a text cue initially activates a specific frame in the reader's mind, but a further cue, usually occurring later in the text, activates a different frame

which may seem incompatible with the initial frame activated. The reader may then experience a sense of disruption and disorientation but only temporarily because frames become hierarchically organised and the reader goes to an initially less likely frame and proceeds with new information input. This oft-quoted example by Rumelhart (1977: 265) and Fillmore's (1980)³ suggested ending illustrate this point:

"Mary heard the ice-cream man coming down the street. She remembered her birthday money and rushed into the house ..." (Rumelhart, 1977).

"and locked the door" (Fillmore, 1980).

The frame activated by most readers and against which they will interpret Rumelhart's text is that Mary dashed into the house to get her birthday money and buy an ice cream. However, Fillmore's text activates a different frame in the reader's mind which brings temporary bewilderment as to the interpretation of the whole discourse. The reader's expectations are not met by the last sentence. However, he will attempt to make sense out of it by adjusting his frame in the light of non-stereotypic new information in the text. Frame-activation is expectation-based. These expectations are made on the basis of textual information. That is why if a text sets up several expectations in the reader (often because it is cohesive non-coherent), consequently several frames will be activated that will be incompatible, and the text/discourse will seem incoherent to the reader because no specific frame could be imposed onto the text/discourse.

2.3.2 Scripts

The "script" is a type of content schemata which may be activated while reading. It was developed by Schank and Abelson (1977) and Riesbeck and Schank (1978) as a representation of predictable situational sequences. The script is a variant of Minsky's frame hypothesis, that is, a subclass of Minsky's frame, but it is more involved in linguistic phenomena than a frame.

It is a device for analysing and comprehending texts as stories. The basic principle is that some pieces of text can be understood if they can be related to a situational stereotype. A script is therefore a detailed list of events arranged in a sequence which characterise a given "standard" situation. It typically contains a list of roles played by the characters in the script, with the goals of the person(s) involved in the situation and what to do when things go wrong. This procedure has been applied in Artificial Intelligence for the understanding of stories by computers. The diagram below, adapted from Schank and Abelson (1977), shows how the computer must understand the structure of a restaurant script from the point of view of the customer:

Diagram 5

Script: RESTAURANT

Detail: Coffee shop

Props: Tables

Menu

Food (F)

Bill

Money

Roles: S = customer

W = waiter

C = cook

M = cashier

O = owner

Entry conditions: S is hungry

S has money

Results: S has less money

O has more money

S is not hungry

S is pleased

(optional)

Scene 1: Entering

S into restaurant

S looks at tables

S looks for where to sit

S goes to one table

S is in sitting position

Scene 2: Ordering

(menu on the table) (W brings menu) (S asks for menu)

Menu is for S

S signals to W

W goes to table

S says to W he needs menu

W goes to fetch menu

W goes to table

W gives menu to S

*S chooses F

S signals to W

W goes to table

S says "I want F" to W

W goes to C

W orders F to C

C says "no F" to W

W goes to C

W says "no F" to S

(go back to *) or

go to scene 4 at no pay path

C cooks (that is,
prepares F script)

To scene 3

Scene 3: Eating

C gives F to W

W gives F to S

S eats F

(optionally return to Scene 2 to order more;
otherwise go to Scene 4)

Scene 4: Exiting

W writes bill

W gives bill to S

S tips W

S goes to M

S gives money to M

(no pay path): S goes out of restaurant

S signals to W

W gives bill to S

A language processor accessing a script is provided with a set of expectations about what will happen next and more generally, what the order of events should be. It also gives him expectations about which entities are likely to be involved and in terms of linguistic occurrences, which lexical items and which lexical relations are likely to appear in a text/discourse. The script can prompt sense selection at the lexical level and can help resolve pronominal assignment problems. To quote Leech's (1983) example:

(1) If the baby won't drink the milk, it should be boiled. Supposing that this sentence calls up a milk script, then the antecedent of it will have to be understood as milk because milk can be boiled but babies can't!

The script can also serve to perceive some texts as pragmatically anomalous although linguistically acceptable.

Consider Sanford and Garrod's ⁽¹⁹⁸¹⁾ example below:

(2) John could not get a waiter to take his soup order.

So he contented himself with eating his sweet course.

The authors claim that this example is pragmatically anomalous because the sweet cannot be served to John before the soup (I myself did not find this text pragmatically odd or the situation inconceivable.⁴

The situation described to the reader is unusual yet possible). Scripts may be language and culture-specific but they may also be reader-specific. The non-native reader encountering this text is likely to rely on his own script knowledge to interpret it. Necessarily his processing of the text will be different from a native speaker reader's as he will try to make sense of the text out of a situation familiar to him. Being more programmatic than the frame, the script can be applied in Artificial Intelligence to the 'understanding' of newspaper stories about car accidents, for example. The following text, an actual newspaper story about a car accident was designed by Schank and Abelson (1977: 180) for computer processing. The programme, SAM (Script Applier Mechanism) running at Yale University (USA) was designed to understand stories that rely heavily on scripts and processed the car accident story. To handle it, the authors built a general car travelling (with an accident "track") script as well as emergency hospital and police investigation scripts which are called up by the

accident track. SAM 'understood' the story, that is, he could create a linked causal chain of conceptualisations that represent what took place in the story.⁵ Thus, the computer understood the car accident newspaper story and answered questions about it, using the words in the input text to access routines or "expectations" that predict what conceptual and linguistic structures are likely to occur later in the text. We shall examine the types of linguistic-lexical and pragmatic-lexical information that the computer employed in answering the questions.

Text: "Friday evening a car swerved off Route 69. The vehicle struck a tree. The passenger, a New Jersey man, was killed. David Hall, 27, was pronounced dead at the scene by Dr Dana Blauchard, medical examiner. Frank Miller, 32, of 593 Foxon Road, the driver, was taken to Milford Hospital by Flanagan Ambulance. He was treated and released. The passenger was extricated from the vehicle by the Bramford Fire Department. No charges were made. Patrolman Robert Onofrio investigated the accident"

Question-Answering Output:

Q1 Was anyone killed?

A1 YES DAVID HALL DIED.

Q2 Was anyone hurt?

A2 YES, FRANK MILLER WAS SLIGHTLY INJURED.

Q3 Why was Frank Miller hurt?

A3 BECAUSE THE AUTOMOBILE HIT A TREE.

Q4 Did Frank Miller go to the hospital?

A4 YES, HE WAS IN YALE - NEW HAVEN.

Q5 How did Frank Miller get to the hospital?

A5 AN AMBULANCE TOOK HIM TO YALE - NEW HAVEN.

The responses supplied by the computer indicate that two types of competences were needed: linguistic competence to answer Q1, Q2, Q4 and Q5, and pragmatic competence which was needed to draw a (pragmatic) inference to

answer Q3.

Q1 - lexical cohesive link: (was) killed - (was) dead - died.

Q2 - lexical cohesive link: (was) hurt - treated and released - (slightly) injured.

Q4 - Hospital - (Milford) Hospital - ("in Yale, New Haven" is pragmatic knowledge).

Q5 - (get to) hospital - (was taken by Flannagan) Ambulance.

Q3 - the link hurt - vehicle (struck a tree) needs a pragmatic inference based on effect - cause: the vehicle striking a tree caused the passenger to die. This pragmatic link complements the lexicopragmatic link of coherence between vehicle and passenger.

Thus, answering Q1, Q2, Q4 and Q5 required linguistic knowledge. Following Widdowson (1978: 100), those questions were "usage reference questions" which made appeal to the processor's knowledge of usage. On the other hand, answering Q3, a "use inference question", required the reader to inference meaning from what he knows of the situation, from his pragmatic knowledge.

It thus appears that knowledge of a car accident script was mostly needed to make pragmatic inferences of a cause-effect type, but "usage reference" questions like those which needed to relate text synonyms (was killed- was dead - died, or was hurt, was injured) did not require any specific script knowledge to be answered which implies that script knowledge is mostly needed in the absence of explicitness provided by textual/lexical cohesion.

2.3.3 Scenarios

Scenarios are types of content schemata. Like scripts, they are devices for analysing and comprehending written texts (Sanford and Garrod, 1981). A scenario describes the "extended domains of reference" where "knowledge of settings and situations" are seen "as constituting the interpretative scenario behind a text" (Sanford and Garrod, 1981: 110). Scenarios are said to be automatically activated in the reader's mental representation if there are elements in the text that constitute part of the scenario itself. Sanford and Garrod write (1981: 129) that "in order to elicit a scenario, a piece of text must constitute a partial description of an element of the scenario itself" (their emphasis).

Unlike other types of content schemata which are more general kinds of knowledge representation, scenarios are said to be situation specific (in a restaurant, at the hospital, at the library.) Scenarios have great predictive power in interpreting subsequent text. The scenario-based approach to discourse understanding implies that a text about, for instance, 'Using the library' does not have to mention 'books' explicitly for the reader to know that the text describes a situation where books are involved, because book is implied by definition in library. It is treated as a 'default' element. Thus, when the library scenario is activated by a text about library regulations in the reader's mind this will automatically bring a book 'slot' into the representation. However, consider this example:

(3) Loan periods are displayed at the issue point on the mezzanine floor. All material taken out of the library must be presented at the turnstile, with your library ticket. Tell the attendant if you want any help.

The situation described by this text may evoke any type of lending library, that is, a book library, a video library, a toy library, because neither book nor video or toy are explicitly mentioned in the text: the scenario for a book library which this text describes may not be the dominant scenario in the reader's mind because other similar scenarios are possible candidates. (It would be interesting to investigate experimentally, by comparing subjects' reading times, whether the book library scenario is most typical for this type of text or whether other scenarios, as for instance, video library, toy library, record library, are likely to be activated by such a text. It seems possible that some relatively recent functions attached to a conventional library, (as for example, video library), would be less readily instantiated by this text in some people than in others, probably on account of socio-economical factors that will need to be examined). There are only few text-explicit lexicosemantic links of cohesion in the passage above mentioned: for example, (library) material - library - library ticket, which are definitional links. Other links which do not relate to libraries specifically: issue point - turnstile

mezzanine floor - turnstile, taken out - turnstile, loan - material, have to be made via pragmatics. It is then possible to conclude that "the domain of reference" book library may not emerge explicitly from the text because of the small amount of lexical cohesion that the text contains. Sanford and Garrod's (1981) evidence about specific role slots being activated in scenarios suggests that linguistic and pragmatic links of cohesion and coherence, explicit or implicit in text/discourse, have a psychological reality. The authors recorded substantial differences in the reading times for two target sentences which contained the same lexical item lawyer but which appeared in slightly different texts. For example:

(4) a. Title: In court

Fred was being questioned.

He had been accused of murder.

Target - the lawyer was trying to prove his innocence.

b. Title: Telling a lie

Fred was being questioned.

He couldn't tell the truth.

Target - the lawyer was trying to prove his innocence.

Sanford and Garrod (1981: 112) report that when the "In court" scenario was activated in condition (a), reading times for the target sentence containing the lawyer were substantially faster than in (b) condition. In (a) the "slot", the lawyer, corresponded

more specifically to the "In court" scenario precisely because lawyer connects more explicitly to court than lawyer to tell a lie. The linguistic-semantic type of cohesive link in (a) has prompted an immediate scenario in the reader's mind whereas the connection between lawyer and telling a lie needed an inference via pragmatic knowledge. The title "In court" in (a) has set up expectations about court proceedings in the reader who was looking for more explicit information in the text on this subject. The power of the linguistic-semantic link between court and lawyer seems to have diminished the effect of the pragmatic links produced by the "collocational chain" to use Halliday and Hasan's image, questioned - accused, murder - innocence. In (b) the reader's expectations were also to find some textual information to match linguistically the title: "Telling a lie". Then, the encounter of tell the truth, an antonymic phrase of "Telling a lie", confirmed the reader's hypothesis about an explicit definitional link with the title of (b). But the occurrence of lawyer later in the text invalidated this hypothesis and forced the reader into making fresh inference to relate lawyer to telling a lie. This additional cognitive activity was reported to take time. Explicit linguistic links between lexical items can therefore be regarded as a necessary and sufficient condition for the activation of specific scenarios in the reader. Non-explicit pragmatic links produced by lexical items seem to activate non-specific scenarios because of the probabilistic nature of

the inferences involved in their interpretation. The parsimonious utilisation, by a writer, of explicit devices of cohesion, may activate non-specific scenarios, which may result in extreme cases of obscure and ambiguous texts.

2.3.4 Summary: content and formal schemata: their role in lexical cohesion

Content and formal schemata may be best summarised by the concept of "structures of expectations" (Tannen, 1978, 1979) to describe what is involved when one attempts to understand written discourse. Thus, "frames", "scripts" and "scenarios" are "structures of expectations" based on past experience which help the reader process and comprehend stories and serve to filter and shape perception. These schemata which include formal schemata, are activated simultaneously during processing of texts. They have been found to be culture-specific (Kaplan, 1966, 1972; Steffensen et al, 1979). In written discourse comprehension, this implies that lack of familiarity with culture-specific content schemata underlying certain lexical items and lexical relationships, or with specific formal schemata underlying the rhetoric of certain texts, can lead to inability on the part of the reader to draw adequate pragmatic and linguistic inferences to access intended meaning. Carrell (1983: 89) remarks that "Content schemata may be absent within as well as across cultures".⁶ Indeed, certain content schemata may be field-specific, as for instance, knowledge of the rhetoric

of scientific reports, which may prevent the reader from drawing inferences and comprehending relationships implied by certain lexical items in text/discourse.

Formal schemata can be restated as expected combinations of sequences of speech acts, logical connections, anaphoric links, etc., that is, a sum of linguistic devices which contribute to the cohesion/coherence of a text. From a cohesion point of view, these formal schemata forming devices can be described along a continuum with, at one end, devices of referential, substitutional and conjunctive-type lexical cohesion and at the other, more grammatical type of cohesion including (discourse) marked conjunctive cohesion, use of pro-forms and ellipsis.

How does a reader use these formal schemata when attempting to reconstruct a text, the coherence and cohesion of which have been purposely mangled? Consider the following sentences:

1. Up he jumped right in the middle of the jigsaw puzzle.
2. "Now, we'll have to count them and start again".
3. "Goodness me, just look at the time", said Pat.
4. When Pat called at Greendale Farm, the twins were busy doing a jigsaw puzzle.
5. Jess wanted to see too.
6. "Looks hard", said Pat.
7. The box says there are twenty pieces.
8. "It's a flower picture", said Katy.

9. Pat rushed to the church to see the Reverend Timms.
10. "It will be very pretty when we've finished".
11. scattering pieces all over the kitchen floor.

The original story (see note 7) consists of a setting/beginning, a development and an ending, and any information that is topical - explicitly stated at setting/beginning level, will likely be repeated at development and ending level and give cohesion and coherence to the text. Thus, will the reader attempt to recover the internal logic of the story by matching it to formal schemata of cohesion and coherence or will he not utilise these textual templates to unscramble the sentences?

The scrambling of sentences has been shown to inhibit the inferential connection necessary to determine the scaffolding of the "event chain" (Warren et al, 1979) rather than disrupt the matching of the story to some formal schemata. In order to recover the text meaning the reader has first to spot the "focal" event in the narrative development, that is, the 'key' or topic sentence supporting this event (here it is sentence number 4). Then on the basis of elements of text structure such as 'conjunction' (in Halliday and Hasan's sense) achieved syntactically or lexically aslexicoreferential or as lexicosubstitutional cohesion, he will connect propositions expressed by sentences. For instance, the reader sees that a specified (physical) state or "action" motivates a specified 'goal' when he realises that the scattering of

the pieces all over the kitchen floor is the result of an action by some animate object, the cat, jumping in the middle of the jigsaw puzzle. Then on this basis he will realise that sentences 5, 1, 11 and 2 should be connected as follows:

Goal: Jess wanted to see too.

Action 1: Up he jumped right in the middle of the jigsaw puzzle.

Action 2: Scattering pieces all over the kitchen floor.

Action 3: Now, we'll have to count them and start again.

On this point, Warren et al (1979) remark that in the development of reading and inferencing from text, young native readers may be more likely to connect adjacent actions and goals than more distant ones, even though the semantic and pragmatic bases for the connection are inconsistent. Thus, they may incorrectly supply (4' b) 'action' as an answer because it is adjacent to the 'goal' (4' a) rather than (4' c) which is pragmatically more consistent with (4' a), though distant from the 'goal'.

(4') a - Chris wanted to help his mother (goal).

b - Chris broke all the eggs in the fridge (action).

c - Chris picked some tomatoes from the garden (action).

d - Chris finished in time for supper (action).

The connection of (b) to (a) was described by Warren et al (1979: 49) as "semantically incongruous" as compared to (a) - (c) - (d) because of presuppositional

meaning underlying help and broke: help implies 'not break' and break implies 'not help'. 'Picked tomatoes' contains the pragmatic feature 'help' and 'finished in time' is a consequence of 'help' but 'broke eggs' disagrees with the idea of 'help' pragmatically.⁸

The connection of (b) to (a) may lead to an inference unintended by the author. The reader expecting a helpful action after reading (a) might revise his inference. For example, Chris went to the fridge to take out some milk (to help his mother) and accidentally knocked over a box of eggs. Thus, for young readers, closeness of events in the chain prompts connection even though some connection results in common sense/pragmatic incongruity (for the adult user of language). It seems to be the case that more idiosyncratic meaning resulting from personal egocentric experience is involved in the processing of texts, generally, by children than would be by adults.

The fact that an adult reader can simultaneously determine the various goals and actions in a scrambled text reflects his ability to interpret cohesion and coherence links between sentences and propositions that express them. These links are encoded partly in the vocabulary and partly in the syntax which, in combination with punctuation devices, contribute to effective text/discourse reconstruction. Hence, the reader's determination of text/discourse meaning does not seem to be subordinated to preconceived formal schemata as story grammars/formal schemata theorists would have it,

but to a more flexible model of processing which suggests that "knowing about stories amounts to knowing about the kinds of permissible connections between events rather than particular higher order structures" (Warren et al, 1979: 50); "permissible connections between events" reflect knowledge of content schemata and imply their culture and language specificity.

Thus, it seems important to underline that the present discussion has conflated two views: the linguist's view and the psychologist's view. For the linguist, most narratives, as noted earlier, begin with a setting which mentions time and place, followed by a description of the cast of characters. Then the narrative events are presented in a series of temporally ordered clauses (development) and the final coda or 'moral', finishes off the story. These narrative events are conveyed via lexical cohesion, coherence relations including speech acts, logical connections, anaphoric links. The psychologist, on the other hand, is more interested in discovering the intellectual processes that the reader utilises in understanding narratives which are said to be processed as "problem-solving reports" (Rumelhart, 1975). In every narrative the hero meets a problem, to solve it he faces sub-problems which he solves or avoids. Then all the sub-problems cumulate in meeting the final 'goal' of the story (for example, Chris wanted (goal) to help his mother; Jess wanted (goal) to see the jigsaw puzzle; the

detective must (goal) find the murderer). Once the goal is identified, the reader starts to search the narrative for how Chris/Jess/the detective/the hero solves the problem that is the 'plan'. By discovering the goal and the plan, the reader can solve the problem and understand the text/discourse.

Problem-solving procedures often depend on the amount of lexical relations of cohesion and coherence in the discourse. These relations also often need the drawing of inferences to be identified which seem to be of two kinds: linguistic and pragmatic. They are discussed in the next section.

2.4 The processing of lexical relations of cohesion and coherence by native and non-native speakers readers of English

2.4.1 Linguistic and pragmatic inferences

In general terms, the notion of inference is used to describe "that process which the reader (hearer) must go through to get from the literal meaning of what is written (or said) to what the writer (speaker) intended to convey" (Brown and Yule, 1983: 256). For example, the following utterance:

A It's stuffy in this room.

is generally interpreted by the native speaker hearer as an indirect request for him to open the window. Thus, a hearer will have to work from the literal meaning to the meaning "Please open the window" via inference of

what the speaker intended to convey. Pragmaticists describe this intended meaning conveyed to the hearer indirectly as implicature (see Chapter 1, Section 1.5.2). The question of whether the hearer of A's utterance will understand the implicature underlying it depends on whether he shares the social conventions of the language governing indirect requests.

This general account of inference may be further analysed into two types of inferences depending on the circumstances in which this activity takes place.

- (a) Inferences as additional propositions to what has been stated or missing links.
- (b) Inferences as filling in lexical gaps, that is, when the meaning of a lexical item is unknown.

Inferences of type (a) seem to be typically utilised by native speakers and readers, whereas (b) type inferences are commonly used by non-native readers and often mentioned in vocabulary discussion in relation to "contextual guessing". This second type of inference is that proposed by Carton (1971) in relation to SL learning, called "inferencing". Carton views inferencing as a two-stage process which consists in "identifying unfamiliar stimuli" by utilising "attributes and contexts that are familiar". (p45). Carton distinguishes between "intralingual" (or within-the-text) clues and "extralingual" (or outside-the-text) clues. Both are involved when drawing "linguistic" and "pragmatic" inferences for text/discourse comprehension (see also Couves (1978) on inferencing).

Carton's third type of clues are the "interlingual" clues which are clues functioning across languages and may sometimes be a facilitative factor of comprehension.

These "within-text" and "outside-text" clues that the reader utilises when attempting to supply a missing lexical item enable us to draw a distinction between linguistic and pragmatic inferences for the purpose of this study. Thus, a linguistic inference will be viewed as one which utilises the linguistic-semantic relationships of text, as for instance:

(a) This afternoon a strange man came to my office. His nose was nearly purple (van Dijk, 1977: 118-119).

which contains a cohesive relationship of partonymy between man and nose.

A pragmatic inference is one which utilises the non-linguistic/pragmatic links produced between lexical items, as for instance, picnic and corkscrew in (b) below.

(b) The picnic was ruined. No one remembered to bring a corkscrew (Carrell, 1982: 484).

2.4.2 Linguistic and pragmatic inferences, and how they relate to lexical cohesion: literature review

Some recent theories which have attempted to explain the structures and processes underlying comprehension of connected discourse have implicitly assumed the drawing of inferences as a component of discourse comprehension but rarely addressed it directly (for instance, Kintsch, 1974, Fredericksen, 1975b). Inferences have a

major function in discourse comprehension as they provide a context for the interpretation of incoming information in order to establish coherence in the text/discourse. The reader's ability to extract relevant information and make necessary inferences depends on stored information of various types which include knowledge of the conventions between writer and reader (as formulated by Grice's (1975) conversational postulates), about presupposition and implication (see Just and Clark's 1973 empirical study of the effects of negation on the drawing of inferences from presuppositional and implicative verbs) and knowledge of linguistic referential relations such as described by Halliday and Hasan (1976) and analysed by Haviland and Clark (1974) within the "Given-New Strategy".

Much research has studied the processes involved in performing specific types of inference.⁹ But most studies relate directly or indirectly to memory representation of a discourse and do not attempt to analyse inference in terms of the specific linguistic phenomena involved to the exception of Just and Clark (1973).

Thorndike's (1976) recognition memory experiment carried out to see whether or not people differentiate between types of inferences provides evidence that pragmatic inferences are made on the basis of a frame activated by the text in the reader, as his model below shows:

Thorndike's (1976) model of generation of inferences
(1976, p439)

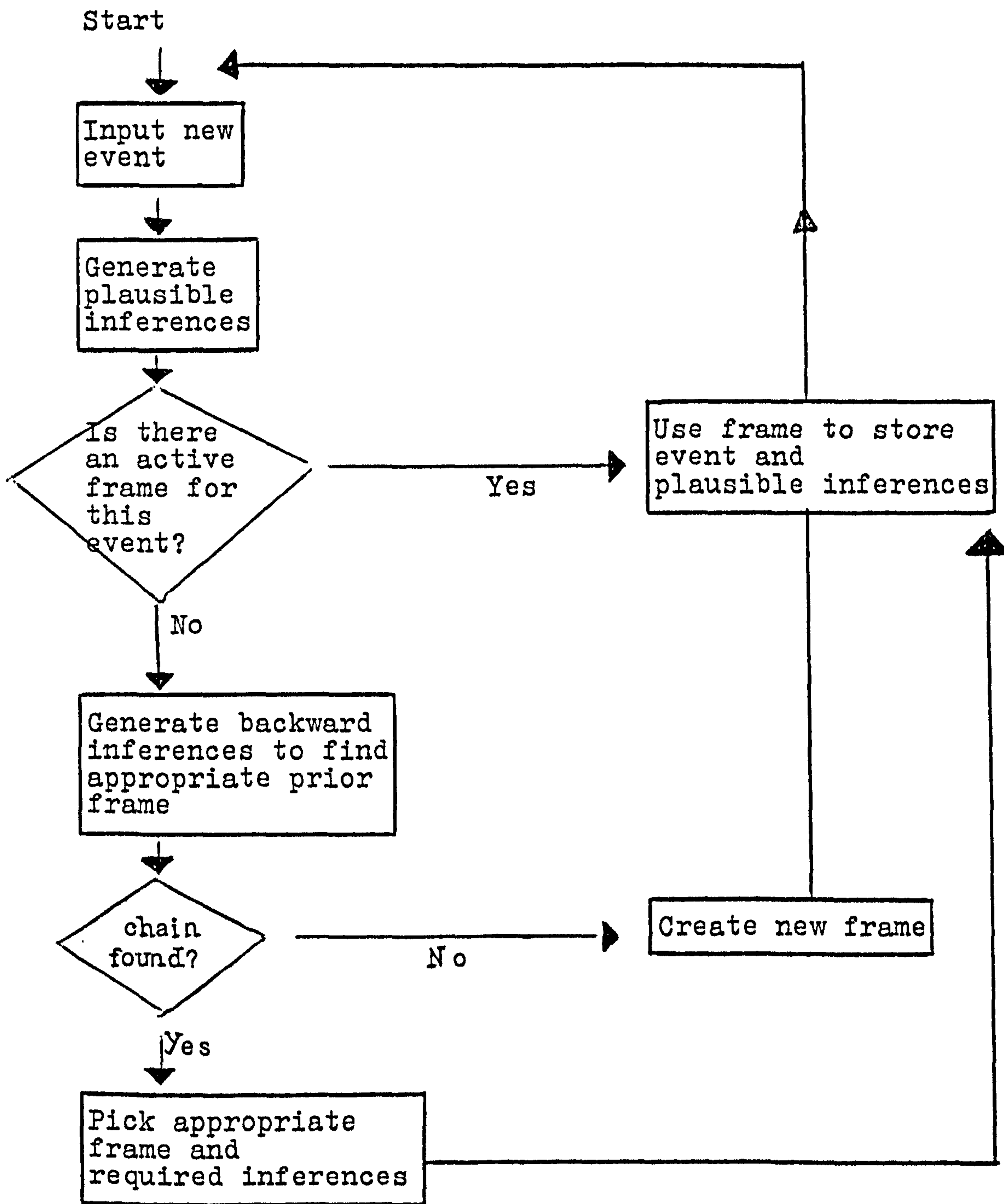


Diagram 6

The results of Thorndike's experiment suggest that a reader will try to identify a general 'contextual frame' when generating pragmatic inferences, that is, he will attempt to establish a bridge, backwardly, to an earlier situation or event, basing his reasoning on textual/lexical information. Consider these examples:

- (5) The hamburger chain owner was afraid his love for French fries would ruin his marriage.

Thorndike suggests that the following inferences (5a) - (5c) might be drawn to understand this sentence, inferences presumably made on the basis of the reader's content and formal schemata:

- (5a) The hamburger chain owner got his French fries free.
(5b) The hamburger chain owner's wife did not like French fries.
(5c) The hamburger chain owner is very fat.

Thorndike points out that these inferences have all been prompted from the text and they are all possible inferences¹⁰, but only (5c) will be retained as the most plausible if sentence (5) is followed by sentence (6) below.

- (6) The hamburger chain owner decided to join weight-watchers in order to save his marriage.

This sentence which occurs later in the text obliges the reader to reduce the number of possible inferences and select (5c) as the most likely inference. Then the whole text should read as:

- (7) The hamburger chain owner was afraid his love for French fries would ruin his marriage. The hamburger chain owner decided to join weight-watchers to save his marriage.

That inference (5) is regarded as the most appropriate . can be explained in terms of its pragmatic congruence, which demonstrates the fact that the reader can establish a link between three causally related events:

- a - eating French fries.
- b - becoming fat.
- c - joining weightwatchers.

from two (surface) coherent lexical items, French fries and weightwatchers. Thus, in order to arrive at the interpretation of (6) as a consequence of (5) he must supply at least two propositions:

- a - French fries eating made him look fat, which worried his wife.
- b - Weightwatching made him lose weight, which pleased his wife.

These propositions are made at different levels of pragmatic background knowledge. General pragmatic knowledge indicates to the reader that over-consumption of some types of food ("his love for French fries") can result in obesity. Specific pragmatic knowledge covers knowledge that chips and hamburgers are always served together. "Hamburger and French fries" seems to be regarded as an idiomatic phrase by American native speakers in the same way as 'fish and chips' would

be regarded as an idiomatic expression by British native speakers.¹¹ But this type of specific knowledge is not crucial to the understanding of the whole text. It represents a subframe which would not hinder comprehension if it were not activated in the reader. Clearly, by making an inference like (5c), the reader has demonstrated that he has perceived the pragmatic meaning which connects weightwatchers to French fries. Had weightwatchers not occurred in (6) above, the causal link between French fries eating and marriage ruin would not have occurred. Consider (8) below:

(8) The hamburger chain owner decided to see a marriage councillor in order to save his marriage.

Sentence (8) does not reinforce the validity of inference (5c) (nor inferences (5a) and (5b)) in the sense that it is pragmatically incongruent. (5a), (5b) and (5c) say nothing which may help relate (8) to (5). The information about French fries in (5) which was the cause of joining weightwatchers in (6) is now irrelevant to the interpretation of (8) because it is less plausible antecedent of marriage councillor. (8) cannot be viewed as the effect of (5). But (8) can relate to (5) via lexical links of cohesion only: repetition of the lexical item, hamburger chain owner in (5) and (8), presence of text antonyms: ruin (his marriage) - save (his marriage), repetition of related items: marriage, marriage councillor, marriage, but not via pragmatic links of lexical coherence.

Pragmatic inferences must belong to the contextual frame being activated while reading. Otherwise they are irrelevant to comprehension. Thus, inference (5a) and (5b) above did not fit into the contextual frame of the text. The lexical semantics and lexical pragmatics of a text/discourse can explain why certain inferences are more appropriate than others. Thorndike (1976) failed to provide a linguistic explanation to the plausibility of inferences in text/discourse comprehension.

Thorndike (1976) carried out further experiments using similar materials as above with longer passages about various topics. Subjects were asked to read through the passage and were later given a memory recognition test. The results suggest that information based on successful pragmatic inferences provides the reader with a false impression that the information is explicit in the text and not inferred (it was falsely recognised as having been presented in the text). These results show that frame activation is a prerequisite to the success of pragmatic inferences. Linguistic and pragmatic links of cohesion/coherence can only be perceived by the reader if the frames being activated by the text are familiar to him, that is, if he has content-schematic knowledge and formal-schematic knowledge which underlie the comprehension of logical connections (as for example, cause-effect in (7)).

2.4.2.1 Linguistic and pragmatic inferences involved
in the processing of lexicoreferential
cohesion and coherence

The process of inference making in connection with anaphoric relations was discussed within the "Given - New" approach by Haviland and Clark (1974), and Clark and Haviland (1977) who pursued Halliday's (1967) idea that anaphora is the most prominent example of signals of Given and New information. Anaphoric elements of text such as pronouns and definite NPs are indexes of the division between Given and New information in text. Thus, Halliday (1967) and Halliday and Hasan (1976) were suggesting that part of the structure of discourse has as its function to indicate to the reader what should be recovered from "co-text" (the preceding text) or "context" (the situation) because it is "given" as opposed to what is being newly introduced in the text. "everything in the text has some status in the 'given-new' framework" (Halliday and Hasan, 1976: 27). Anaphora, discussed at length by these two authors was treated as given information.

Clark and Haviland (1977) have brought psychological evidence on a linguistic phenomenon discussed by Halliday and Hasan (1976), viz. lexicoreferential cohesion. The results of the experimental investigations by Haviland and Clark (1974), Clark and Haviland (1977) and Clark (1977), aimed at showing that native readers employ the "given-new strategy" to identify referents for definite NPs, suggest reader's awareness of linguistic and pragmatic links of cohesion/coherence when drawing inferences.

Linguistic and pragmatic inferences seem to have some psychological reality as the search in memory for a matching antecedent (that is, a matching member of a link of cohesion or coherence) is reported to take more time when the 'context (or preceding) sentence' explicitly posits the existence of some entity referred to in the 'target sentence' (that is, is given information) than when it does not. Thus, comprehension time was increased (by about 200 msec) when the subject had to make a bridging inference, that is, supply a missing link of a pragmatic nature between two propositions. Consider these examples (Haviland and Clark, 1974):

(9) a - We got some beer out of the trunk.

b - The beer was warm.

(10) a - We checked the picnic supplies.

b - The beer was warm.

Haviland and Clark (1974) and Clark and Haviland (1977) explain that the native reader encountering (9b) will realise that the beer is given information because it is being repeated. The givenness of the information seems to have accounted for the reading time advantage. The NP the beer in (9b) refers back to the NP the beer in (9a), its antecedent. But the reader encountering the beer in (10b) will not automatically realise that beer is given information, although the reference item the may provide him with a clue that its antecedent has to be searched in the preceding sentence. This "non-automatic" connection (Brown and

Yule, 1983: 260) by the reader between the beer and picnic supplies is explained by the fact that in the 'context sentence', the beer does not explicitly posit the existence of some beer. This compels the reader to make a bridging inference of a pragmatic type which takes the form of an additional proposition: beer is an example of picnic things (at least in American culture), therefore picnic supplies and beer should be connected via pragmatic knowledge.

Supposing that beer occurs in the following environment:

(11) We checked the picnic supplies in the car. The beer was warm.

and supposing that the meanings of picnic supplies and beer are unknown to the (non-native) reader, he could well establish that beer is either referentially related to car (something belonging to the car was warm) or to picnic supplies since, if such a case occurred, both picnic supplies and car would be competing for antecedence. But if the reader knows the meaning of picnic supplies and has pre-existing schematic knowledge that picnic supplies may include beer, the referential link between (unknown) beer and picnic supplies would ultimately be clear to him.

As Brown and Yule (1983: 263) rightly pointed out, for some people, beer addicts especially, beer is an essential component of picnic supplies and will therefore be "automatically" activated in this particular reader's

mind. For others, it is not, and it has to be included on a particular occasion. (For people with a Moslem background, beer will not be included in picnic supplies at all, and no connection will be possible between the two lexical items, beer and picnic supplies).

Haviland and Clark (1974) carried out a further experiment to rule out the possibility that the repetition of a noun can account for reading time advantage, and provided a pair of items where the antecedent sentence contained a non-definite NP (unlike example (9) above) which was being repeated with an anaphoric item in the next sentence in each case.

(12) Ed wanted an alligator for his birthday.

The alligator was his favourite present.

(13) Ed was given an alligator for his birthday.

The alligator was his favourite present.

The results evidenced longer comprehension time for (12) than for (13), which made it clear for the authors that repetition of the noun did not account for the reading time advantage in (9). Longer reading time may have been caused by the verb want which, as Chafe (1972) suggested, is a type of verb which does not presuppose the existence of its object : if X wants Y, it does not entail that Y exists; but if X is given Y, it entails that Y exists.

The NP an alligator did not necessarily set up a direct antecedent for a subsequent anaphoric referent because it followed want. Sentence (12) needed a

"bridging" linguistic inference to the effect that the alligator which was Ed's favourite present was in fact the one which he had wanted and presumably that he received as a birthday present.

In (13) the reader knows with precision that the alligator that Ed possesses is the one that was his favourite present, although the NP the alligator in the 'context sentence' is also non-definite as in (12).

The authors of this experiment suggest that repeated lexical items are not necessarily easier to process. They may show apparent processing simplicity but the coherence underlying their occurrence is harder to perceive in some cases. However, when assessing these results from a linguistic viewpoint, we may be tempted by the conclusion that reading time measures as an index of text processing difficulty cannot be explained in terms of the amount of lexical cohesion and coherence since both examples contain lexicoreferential cohesion produced via the repetition of alligator. However, in view of the fact that example (13) contained an additional linguistic definitional link between given and present (given is contained in the definition of present. Present implies a giver and a receiver), that was missing in (12) may enable us to conclude that explicitness provided by lexical cohesive links facilitates text processing thereby reducing reading times.

It would have been interesting to compare reading times taken to identify referents of lexical items under

three conditions.

- a - Repetition of the same lexical item (for example, beer - beer).
- b - Cooccurrence of a superordinate with its hyponym (for example, drink (N) - beer).
- c - Cooccurrence of "pragmatic paronyms" (this category extends the semantic category of paronymy described in Chapter 1 Section 1.7.6, for example, picnic supplies - beer).

We may frame the hypothesis that the reading times required to identify referents of (cohesive or coherent) lexical items would be proportional to the amount of explicit ties contained in text.

Sanford and Garrod (1981) have conducted experiments to test the validity of Clark and Haviland's (1977) findings about the use of inferences when identifying lexical referents. Materials similar to Clark and Haviland's were tested in the same conditions. In condition (14), describe below, the link produced by clothes - clothes is linguistic cohesive, but in condition (15) the link dressed - clothes is pragmatic coherent. Consider Sanford and Garrod's examples (1981: 104):

- (14) a - Mary put the baby's clothes on.
b - The clothes were made of pink wool.
- (15) a - Mary dressed the baby.
b - The clothes were made of pink wool.

The authors found no significant difference between the times taken to process (14) and (15). They argued that because dressing activates clothes in the reader's mental representation of sentence (15a), subsequent mention of the clothes in (15b) would be understood as quickly as it would be if the clothes were mentioned explicitly, as in (14a). Thus, it seems that the pragmatic inference needed to relate clothes to dressed in (15) was no more problematic than the linguistic inference required to relate clothes to (baby's) clothes in (14). Sanford and Garrod claimed that the connection of dressed with clothes is conceptually driven (that is, it is a top-down process) determined by a prior frame activated by dressed, but as Brown and Yule (1983: 265) argued, a dressing frame would not necessarily activate clothing, as other lexical items could also be activated, as for instance, bandage, entrails. Consider their examples (p265) below:

- (16) a - Mary dressed the baby's arm.
b - The bandage was made of white cotton.
- (17) a - Mary dressed the turkey.
b - The entrails spilled out of the bowl.

Brown and Yule reject Sanford and Garrod's assumption that the connection between (15a) and (15b) can be described in terms of a decomposition of lexical meaning ("when a verb like 'dress' is encountered this will evoke from memory a representation which contains slots for a variety of entities implied in the meaning of the verb, such as, 'clothing'" (Sanford and Garrod, 1981: 108)),

but provide no linguistic explanation of the phenomenon described by Sanford and Garrod in psychological terms. Brown and Yule's examples clearly demonstrate that dressed will activate bandage in the reader's mind only if it is accompanied by arm to form a pragmatic link with bandage. It is the combination of dress and baby's arm which seems to set up expectations about bandage in (16). Likewise it is the combination of dress and turkey which is likely to set up expectations about entrails in (17). In other words, the selectional constraints of dressed, different in (16) and (17), seem to be mostly responsible for the lexical coherence of these texts: baby's arm, more than dressed (16) and turkey more than dressed (17) seem to determine the occurrence of bandage and entrails in the respective examples.

To summarise this discussion, Haviland and Clark's (1974), Clark's (1977), Clark and Haviland's (1977) and Sanford and Garrod's (1981) experimental results have implications for our analysis of lexical cohesion. They seem to demonstrate the fact that explicit ('stated') links of referential cohesion are easier to identify because they are there in the text, whereas implicit ('implied') links of pragmatic coherence may be more problematic when reading because they are not in the text but have to be "worked out" by the reader, and this additional processing seems to depend on whether the "domain of reference" (Sanford and Garrod, 1981: 109) available to the reader can be extended to include the 'implied' entities.

2.4.2.1.1 Effect of order of referential hyponyms in reading

In an earlier investigation (1977) Garrod and Sanford brought evidence that finding referents for lexical items standing in hyponymic relation takes longer time when the hyponym occurs in the 'context sentence' than when a superordinate does. For example (Garrod and Sanford, 1977: 79):

- (18) a - A robin would sometimes wander into the house.
b - The bird was attracted by the larder.
- (19) a - A bird would sometimes wander into the house.
b - The robin was attracted by the larder.

On the basis of reading times as an index of processing ease or difficulty in those experiments (18) was reported to be easier to comprehend than (19). When processing (18) ("category-last") the reader knows that the robin in (18a) is a bird before reading (18b) because robin presupposes bird. But when reading (19) "instance-last", he has to deduce from the text that the bird in (19a) is the robin in (19b) because bird does not necessarily presuppose robin.

These findings seem to have presuppositional origin. They imply that a linguistic inference seems to be needed to connect robin (19b) to bird (19a) and this was responsible for additional reading time. The authors suggest that the difference between the two conditions reflects the fact that more information has to be incorporated into the representation that the subject has of the text in the "instance-last" case (bird - robin) than

in the "category-last" case (robin - bird). This information integration which reflects the perception of a link can be explained in terms of a linguistic inference which seems to be more required when a hyponym precedes a superordinate in the text than when it follows it. However, Williams (1983: 40) pointed out that superordinate ties may represent major potential sources of difficulty to the FL learner and predicted that the cohesive link formed by a hyponym and a superordinate term (or "general word" in his terminology) would be particularly troublesome because of the difficulty of forming a "mental picture" of the meaning of this category of items (as for example, man, action, people) and of the fact that they are cohesive not with a single word but with a wider stretch of language. However, no evidence was put forward to test this prediction.

2.4.2.1.2 Effect of distance between cohesive hyponyms in reading

Garrod and Sanford (1977, 1978) conducted a further experiment in which they tested integration of information when anaphoric reference was involved. Unlike the experiment described above, the sentences were not presented consecutively but were separated by an 'obstrusive' or 'distracting' sentence. For example (Garrod and Sanford, 1977: 83):

- (20) a - A vehicle came roaring round the corner.
b - The bus nearly flattened a pedestrian.
c - It had had a brake failure.

- (21) a - A vehicle came roaring round the corner.
b - It had had a brake failure.
c - The bus nearly flattened a pedestrian.

The 'It' sentence has a somehow obstrusive function. It adds no further information about the nature of the referent. It was found that the separation of the 'It' anaphoric sentences (20c and 21c) from their antecedent sentences (20a) and (21a) has affected reading times as (21) was read in longer time than (20). This seems to demonstrate the fact that the linguistic inference needed to connect superordinate and hyponym (or hyponym and superordinate) was delayed when the 'It' sentence was added in second position (b). Williams (1983: 42) suggested that the distance between two elements of a cohesive tie may have cognitive consequences in reading but provided no evidence. Apparently, in reading, words and phrases are stored in the reader's short term memory and are recalled for linkage when the anaphoric signal is read. As Garrod and Sanford's (1977) results have implied, the greater the distance between the antecedent item and its corefering item, the more likely it is that the antecedent will have faded from the reader's short term memory, thus reducing the chances of linkage. Williams (1983) proposed that the limits for effective processing should not exceed two to three clauses, and in any case, no more than two sentences. Otherwise memory load causes comprehension difficulty to the reader who loses the thread of the story by searching for an antecedent to the anaphoric item. As for linkage produced

via cataphoric reference it seems possible, as pointed out by the same author, that cataphora may cause more processing problems to the reader who is generally accustomed to searching backward for the beginning of a tie rather than forward for the end of a tie. But the question of the distance between two members of a tie in cataphoric relations seems less acute than in anaphoric relations because the two ends of a cataphoric tie are frequently close in a text.

The effect of an intrusive sentence and of cataphoric clues on inferencing the meaning of unknown words will be investigated empirically in the next chapter.

2.4.2.1.3 Effect of "cultural background knowledge" on the processing of lexicoreferential cohesion/coherence

The role played by background knowledge can be determinant in the identification (or inferencing) of referential links of cohesion (and coherence) in text/discourse comprehension. Johnson (1982) argues that a non-native reader may appropriately identify the topic of a text (that is, recognise its coherence) but interpret it according to his own "culturally experienced background knowledge", that is, fail to identify its cohesion, and as a result may lose its textual cohesion in recall. For example (Johnson, 1982: 512):

"The assembled witches vowed to obey their god, the master witch who was disguised as an animal. They pledged their children to the god and thanked him for food and life. The religious ceremony was followed by feasting and dancing. The witches dressed up like animals."

Subjects showed in their written recalls a mis-identification of the referent of NP The religious ceremony despite the fact that most of them recognised the topic of Hallowe'en in the text. In effect, the referent of this phrase consists in information about religion and god contained in the first two sentences, and misidentification of this cohesive tie has resulted in incorrect inferences on what happened in the witches' meeting. For example:

"They promise god (to make a good job)".

"There is a story about the witch where they promise their master and god, (but they had broken their promises)".

"The witches had a meeting (to discuss how to control the people)".

This suggests that the linking of propositions in discourse which often involves the making of linguistic and pragmatic inferences is subordinated to the reader's system of values. Values and attitudes are often expressed at the lexical level and can be one of the main sources of difficulty in a FL (see Rivers, 1968, on this point).

2.4.2.2 Linguistic and pragmatic inferences involved in the processing of lexicoconjunctive cohesion

The use of discourse markers to produce conjunctive cohesion or "conjunction" in Halliday and Hasan's terminology has been the object of ample discussion in their study of cohesion in English together with coding of the various

functions of conjunctive cohesion in text. But the use of lexical devices to achieve conjunctive-like cohesion (or lexicoconjunctive cohesion) does not appear to have received any treatment either in their chapter on "conjunction" or in the one on lexical cohesion. As argued in Chapter 1 (Section 1.6.3) conjunctive meaning of text/discourse is not solely dependent upon the occurrence of discourse markers for its conveyance. Absence of syntactic markers of "conjunction" or of their equivalent paralexical forms, may achieve a conjunctive type of cohesion/coherence in the discourse as in (24) below. Compare these examples:

(22) Our garden was a disaster this year. However,
the orchard is looking very healthy.

(23) Our garden was a disaster this year. By contrast,
the orchard is looking very healthy.

(24) Our garden was a disaster this year. The
orchard is looking very healthy.

These three texts are semantically equivalent, although discourse markers (syntactic in (22) and paralexical in (23)) give an impression of smoothness to the reader that (24) lacks.

2.4.2.2.1 Evidence from research in NL

Experimental evidence of native and non-native readers' awareness of conjunctive meaning, whether overtly expressed via markers of conjunction and thence requiring no inference, or covert, thus requiring linguistic or pragmatic inferences, is diverse and sometimes contradictory. Some studies reveal that in general discourse markers have no facilitative effect on reading comprehension: "conjunctions" do not always give clues to the propositional development (see Pierce, 1975, on "Interparagraph continuity"). Rather, the meaning of some of them may obscure the overall meaning of the text. Stoodt (1972) in a cloze study with NL fourth grade American children found a significant relationship between reading comprehension and the comprehension of "conjunctions." Some of them were found to be significantly more difficult to understand than others. This demonstrates the fact that subjects were unable to draw linguistic or pragmatic inferences, that is, to deduce conjunctive meaning from the lexical content of the propositions expressed by the juxtaposed sentences. Chapman and Stokes (1980) reporting on an on-going longitudinal study in which they used cloze-type techniques to assess the mastery of cohesive devices in reading by native British children also found that comprehension was hindered by the presence of some discourse markers which caused considerable problems to even the oldest children (13-14 years old). For instance, the processing of the "conjunctions" which consisted of a group

of words (at the same time) or of single words that are infrequent in children's vocabulary (furthermore, nevertheless, finally) posed acute processing problems. These "conjunctions," it may be noted, are likely to pose problems to non-native readers alike, and their lexical paraphrase may appear to be easier to decode and facilitate overall comprehension.

2.4.2.2.2 Evidence from research in SL/FL

Cohen et al's (1979) study of the reading comprehension of specialised English texts by non-native speakers of English¹² also suggests that non-native readers are unaware of the fact that often conjunctive markers of cohesion are complemented in their function by lexical forms / items that express "lexically" the conjunctive meaning intended by the writer. Thus, the non-native subjects involved in the investigation did not know the meaning of certain conjunctive markers and no attempt was made to compensate for this lacuna. But native speaker readers seem to be less sensitive to the absence of overt markers of conjunction and are in general capable of making bridging inferences whether linguistic or pragmatic) (see Hagerup-Neilsen, 1977; Freebody and Anderson, 1981, on these points). These investigations suggest that native readers had only slightly more difficulty processing texts that were not marked for intersentential relationships than they had processing texts that were so marked. Urquhart's (1977) finding was that the signalling of statements with syntactic connectors did not usually affect recall and concluded that implicit

relationships (of lexical coherence) not signalled via conjunctive markers and holding between sentences are important to consider. For example:

(25) The woodpecker is an unusual bird. It bores holes in trees.

The two propositions expressed by the juxtaposed sentences underlie a causal relation. The occurrence of woodpecker in S1 and holes in S2 produces a pragmatic link of 'cause' which substitutes successfully for a conjunctive marker of causality (as for instance, "because"). However, appeal to pragmatic knowledge has to be made to interpret S2 as the cause, the explanation of the unusualness of the bird rather than as the result, the consequence of its unusualness.

The general point that emerges from these studies is that learners may attend too much to overt markers if they are so trained and may not be ready for texts that do not make use of them.

It seems worth recalling that, as noted in Section 2.4.2.1.3. above, linguistic and pragmatic inferences needed in the processing of lexical cohesion are obviously affected by the reader's background knowledge in terms of the cultural "présupposés" that he may bring into the text while reading. Kaplan (1966, 1972) brought light on the role of cultural background in SL/FL learning, vehemently arguing that rhetoric is not universal and is a cultural phenomenon which varies from culture to culture.

Rhetoric is tied to the linguistic system of a particular

language, and hence determines the forming of schemata in the native speaker reader. As a result it can affect the processing of text in an FL in so far as cultural factors intervene in and may interfere with the decoding and interpretation of relations of cohesion and coherence. Thus, Steffensen's (1981) investigation brings evidence that reading comprehension and text recall of conjunctive markers have a correlated effect with cultural background knowledge. Her finding suggests that when there is a mismatch in cultural background knowledge between the reader and that assumed by the text, there is ultimately inability to appropriately identify the 'schema' underlying a text, and a loss of textual cohesion in recall. Steffensen et al (1984) also found that non-native readers "distort meaning as they attempt to accommodate even explicitly stated propositions to their own pre-existing knowledge structures". (pp60-61).

2.5 Conclusion

At the beginning of this chapter, we proposed a framework within which lexical cohesion, as a text and discourse phenomenon, will be discussed. It was emphasised that such framework could not be purely linguistic, exclusively, "free from contamination from knowledge about cultures, belief systems or facts about the world" (Fillmore, 1977: 76). In order to do justice to the interpretative dimension of lexical cohesion, specifically to its role in the reading process, the cognitive component was introduced in the present analysis and focussed on some general principles underlying the reading behaviour of native speakers and

non-native speakers of English.

The concept of background or schematic knowledge and its involvement in the inferential process was extensively discussed. We have characterised the making of inferences - one of the strategies mostly utilised by non-native readers in reading comprehension, as being of two kinds: linguistic and pragmatic. These inferences are often utilised to identify relationships between propositions expressed by vocabulary items. Typically they are utilised by non-native readers in the encounter of unknown lexical item(s) when reading. Adapting Thorndike's (1976) model of generation of inferences, we may represent the inferencing process at work when unknown vocabulary items are encountered in reading as follows:

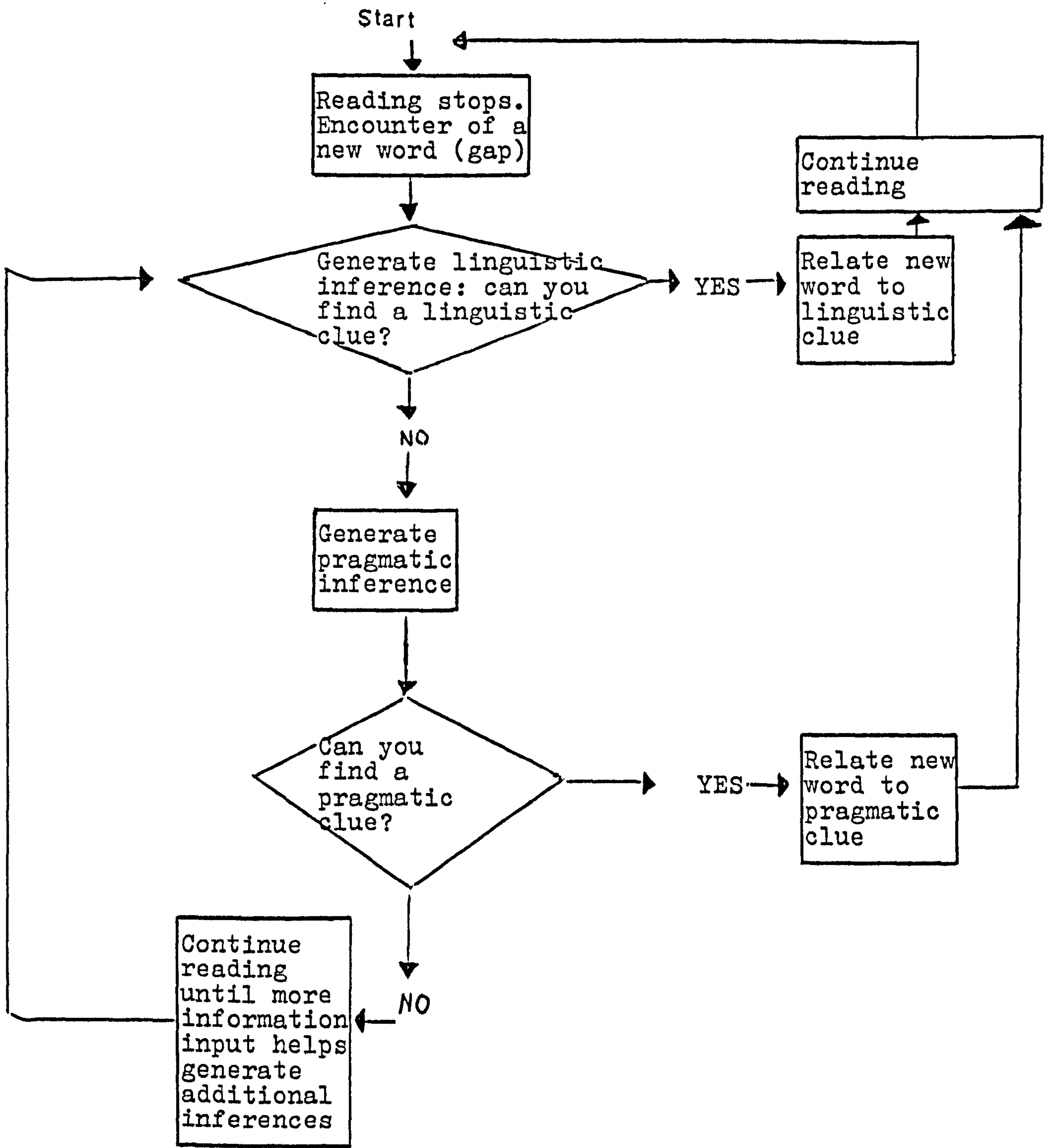


Diagram 7

The processing of referential relations of lexical cohesion/coherence by the native speaker has been experimentally investigated by psychologists and psycholinguists. Some of them based their findings on the assumption that readers employ the "Given - New Strategy" when drawing inferences, that is, when supplying additional propositions to connect two sentences together. Although no explicit distinction was drawn between "linguistic" and "pragmatic" inferences, the outcome of their researches has clearly demonstrated that the "Given - New Contract" between the writer and the reader (an implicit 'contract' which stipulates that the speaker/writer must agree to construct utterances which contain information that he believes the hearer/reader does not know - 'new' information and which also stipulates that the hearer/reader, for his part, tacitly agrees to interpret the sentences following these assumptions), was often violated in matters involving lexicoreferential cohesion/coherence when linguistic and pragmatic inferences were needed to be drawn for text/discourse understanding. By failing to identify lexicoreferential links of cohesion and coherence, the reader revealed his unawareness of reiterated and referential lexical meaning as 'given' information, and therefore his inability to comply with the rules laid down by the writer concerning 'givenness' and 'newness' of information.

The cognitive processes underlying reading and

inferencing do not appear to have been explained in terms of the linguistic phenomena involved, specifically of the lexical relationships of cohesion holding in text and discourse. Furthermore, psychologists' work has offered a cognitive treatment of lexical cohesion in native language essentially. Such shortcomings of the literature give justification of the experimental investigation that we propose in the next chapter. We shall attempt to identify whether linguistic and pragmatic links of cohesion/coherence have a psychological reality in FL learners by examining their performance on a vocabulary inferencing task, that is, we shall attempt to understand the working out of relationships between lexical/textual elements (that is, reliance on text-presented information) and the use of various knowledge structures (or schemata) which enable one to establish links between textual elements and draw inferences.

Notes on Chapter 2

- 1 See more on this point in Goodman (1973), Smith (1971, 1973), Clarke and Silberstein (1977).
- 2 Researchers concerned with teaching reading in a FL and assessing their learners' abilities in it in general had limited knowledge of their subjects' abilities in their native languages. Often subjects' performance in the FL was not measured in relation to their performance in the NL. It has often been the case that proficient FL readers are also proficient readers in their NL.
- 3 Fillmore's (1980) personal communication: quoted in Carrell (1983: 83).
- 4 In some Chinese and Indian restaurants, savouries and sweets are served at the same time and the customer is free to start with any type of dish he pleases.
- 5 What follows is Schank and Abelson's (1977) account of how the programme SAM works:

"SAM works by analysing each sentence into a Conceptual Dependency representation. If this representation fits into a script, that script is brought into memory. Succeeding inputs are analysed and the result is looked for in the script. If the result is found, any necessary conceptualisations that are known to have been skipped between the first input and the second are inferred to have happened. This continues until there are no new inputs or until a new input does not match a part of the current script" (p178)

"Each script possessed by SAM defines a context which consists of:

- a - a list of patterns which predicts what inputs will be seen at a given point in story;
- b - a binding list which links the tokens for objects produced by MEMTOK (a memory module) with script variables;
- c - a record of the script scenes which are currently active;
- d - a list of scriptal interferences - anomalies - which are currently outstanding; and
- e - a 'strength' indicator which SAM uses to flag how strongly it believes in its inferences" (p184).

6 See on this point, Anderson et al (1977) who conducted a research project on content schemata knowledge with monocultural American groups.

7 Original text: When Pat called at Greendale Farm, the twins were busy doing a jigsaw puzzle. "Looks hard" said Pat. "It's a flower picture" said Katy. "It will be very pretty when we've finished". Jess wanted to see too. Up he jumped right in the middle of the jigsaw puzzle, scattering the pieces all over the kitchen floor. "Now we'll have to count them and start again. The box says there are twenty pieces". "Goodness me, look at the time" said Pat. Pat rushed to the church to see the Reverend Timms.

8 In effect, whether these features are to be regarded as pragmatically or semantically presupposed is not clear. One can actually help someone suffering from cancer, for example, or other incurable diseases, by giving an overdose of morphine and stop his life. Then should 'kill' be regarded as semantically or pragmatically presupposed by 'help'.

- 9 Verb-based conceptual inferences have been studied by Schank (1972) and Rieger (1975), inferences concerning the integration of information expressed by many different sentences experienced successively and often non-consecutively in time was studied by Bransford et al (1971), inferences relating to spatial integration were reported by Bransford et al (1972), and inferences relating to memory for narrative discourse by Thorndike (1977).
- 10 Native speakers and some non-native speakers (including myself) did not find (5a) a possible and plausible inference for sentence (5).
- 11 "Fish and chips" would be idiomatic for British speakers. "Hamburger and chips" has also been tested with native British speakers who felt it as more of an idiomatic phrase than, for instance, "hamburger and beans".
- 12 Student informants were instructed to underline all vocabulary and structures that they found difficult to understand. Then they were asked overtly on problematic areas by the researchers, that is, whether some word or structure was a problem and whether it interfered with the comprehension of the sentence, paragraph or passage overall. Interview sessions were usually conducted in the student's native language (Cohen et al, 1979).

CHAPTER 3

AN EXPERIMENTAL INVESTIGATION OF THE AWARENESS OF LEXICAL
RELATIONSHIPS OF COHESION AND COHERENCE BY NON-NATIVE
ADULT READERS OF ENGLISH ON AN INFERENCING TASK

Part One

The previous chapter emphasized that in reading comprehension the recognition and interpretation of linguistic information are functions of 'schemata' or global knowledge structures that enable the reader to reconstruct and interpret messages. Thus top-down processes and the role of background knowledge view the reader as an active participant in the NL/FL process by making predictions and processing information. Top-down operations are also complemented by bottom-up processes which are invoked by the incoming linguistic information. These cognitive processes are determinant in the comprehension of oral or written communication, but as they involve two basic types of knowledge i.e. linguistic and non-linguistic, it seems legitimate to assume that they are also paralleled by two different types of text/discourse or word comprehension procedures i.e. linguistic and non-linguistic. Granted that top-down and bottom-up processes are determined by linguistic and non-linguistic background knowledge, the implications for our analysis are that these processes are also determined by lexical relations of cohesion and coherence in text/discourse comprehension. Ultimately these points raise two questions relative to the utilization by

individuals of lexical relations of cohesion and coherence in reading:

1-How do readers exploit linguistic and non-linguistic information when attempting to comprehend written text/discourse in their NL?

2-How do readers exploit linguistic and non-linguistic information when attempting to comprehend written text/discourse in a non-native language?

While these questions are necessary to pose as preliminaries to our discussion, they nevertheless need to be refined for the purposes of our investigation and should therefore be formulated as follows:

1-How do readers exploit lexical semantic relations of cohesion and lexical pragmatic relations of coherence when reading written text/discourse in their NL?

2-How do readers exploit lexical semantic relations of cohesion and lexical pragmatic relations of coherence when reading written text/discourse in a non-native language?

These last two questions enquire into the nature of the strategies at work when an individual is reading. In order to provide adequate answers to them empirical research on specific reading strategies needs to be carried out.

3.1 The rationale for empirical research

The work of theoretical linguists and psychologists have made applied linguists aware of the need for empirical research to test their claims about many issues in linguistic description. For the past decade, literature on empirical investigations relating to 3L learning has been flourishing. Yet little work could be noted in the field of FL learning(literature relating

to this subject was reported in Section 2.6 of the previous chapter). As emphasized earlier, most researchers on SL/FL learning nowadays are not interested in the formal features of language which stop short at sentence boundaries. They are interested in unity and organization of a text and the effect of the text on the reader (see for example Widdowson 1978 for writing pedagogy, Witte and Faigley, 1981 for preliminary investigation of cohesion and coherence in relation to writing; Scarcella, 1984 for empirical investigation on cohesion devices in writing). However this emphasis did not seem to be paralleled by similar research in reading. Specifically, little empirical research has been undertaken to further our understanding of how cohesion and coherence links are perceived in reading and no investigation seems to have been conducted on the awareness of linguistic and pragmatic links in reading in a FL. Although it is undeniable that these areas of study can be investigated, it is also true that there are difficulties inherent in empirical research relative to the choice of an adequate experimental method and to the interpretation of the results obtained, which cannot be ignored. The question therefore concerns, not the principle of empirical research itself, but the application of this principle to specific situations. Are there optimum methods of investigation which would reduce chance factors for example to a bare minimum? When human individuals are involved in experimental study unpredictability is high however perfect the design of the experiment may be. Our aim in this research is to control a certain number of variables. However other variables will remain inevitably uncontrolled and are likely to affect the results. Some examples in the sections below

indeed show that certain uncontrolled variables have acted favourably, or unfavourably on the subject's performance and as a result have obscured the effect of the variables being investigated. It is therefore essential that the design of an experiment be meticulously done so that contingent factors of chance, probability and risk are reduced to a minimum.

While linguistic (and pragmatic) links of cohesion and coherence can be successfully measured in written discourse (cf. Halliday and Hasan's analysis of cohesive ties in two literary texts), their recognition and exploitation by a potential reader may be more problematic to the experimenter because there will always be doubt as to whether or not the subject has actually utilized the experimental procedure being investigated. Thus it may happen that subjects do not utilize clue-searching strategies at all and nevertheless obtain correct responses. In empirical research, an important point is to know WHAT to select as experimental task (or independent variable). In the present research, our aim was to measure "objectively" the reader's awareness of lexical relations of cohesion and coherence and to formulate generalizations about clue-exploitation in reading in a FL.

3.2 Utilizing inferencing as an experimental task

In chapter two it was stated that all readers naturally utilize the strategy of inference-making when comprehending text/discourse in their NL or in FL. Furthermore when encountering a new or unknown item (and depending on its topical importance), they would exploit a number of explicit textual clues, but also ignore others. They would also make appeal to their general background knowledge to access its meaning. However the

exploitation of this strategy of text/discourse comprehension is presumably different when one infers unknown meaning in a non-native language owing to various factors (linguistic, cognitive, sociocultural) which determine the comprehension of oral and written communication in one's NL or in a FL (see for example the study by Grunlin, Courtney, Langer, Pehrson, Robinson and Sakamoto on the use of the cloze procedure in comprehension in three different languages)*.

As stated earlier, whether the reader is aware of the nature of the inferences he makes in his native language to access meaning is a question that also needs empirical examination. Can the reader account for the quantity or the quality of the clues he exploits when reading in his NL in precise terms? Presumably a reader is aware of his use of the 'context' a notion often defined quantitatively rather than qualitatively, when reading in his native language, but he may be unaware of what to exploit as context in a SL/FL.

Since our concern in this study is with non-native reader's performance exclusively, we shall limit our analysis to a close examination of one type of inferencing procedure in non-native reading and see whether learners are aware of lexical relationships of cohesion and coherence when recovering 'lost' meaning. However a qualitative, empirical study of the performance of native readers on an inferencing task, to measure their awareness of lexical links of cohesion and coherence in text/discourse comprehension will need to be undertaken in future research. It will also be rewarding to compare the performance of the same subjects in an inferencing task, in both NL and FL, using similar types of lexicosemantic and lexicopragmatic clues/links.

It may then be possible to reach some general conclusion about

the exploitation of specific inferencing procedures across languages.

It seems difficult practically to analyse the cognitive processes of individuals reading a prose passage written in a FL (or NL) with the view to discovering whether they establish links between lexical items and how they do it as part of their overall reading (comprehension) activity. Approaches to reading passages are different among individuals often because they have diverse linguistic semantic and non-linguistic pragmatic competences. As a result it would be difficult to provide any systematic account of their approaches to text/discourse understanding and draw conclusions on their use of lexicosemantic and lexiconpragmatic clues if their reading activity is not under careful control. Also by creating "obstacles" in the reading comprehension of the subjects, it seems possible to control their approach to text/discourse understanding and to arrive at some generalizations regarding their reading strategies in the encounter of problems or "obstacles". The creation of physical obstacles in text/discourse such as blanks (or clozes) is likely to have the effect of a "psychological obstacle" in the reader's mind and compel him to direct his attention to particular elements of the text. Ultimately this will enable the experimenter to focus on subjects' problem-solving strategies essentially.

The experimental methods which can serve this purpose include the cloze procedure (of item deletion) which, when complemented by the verbalization of subjects' responses, seems mostly appropriate to account for some of the processes involved in reading (comprehension) in a FL. Our next point deals with the procedure of verbalization and discusses its success and its limitations when utilized in combination with the cloze procedure.

3.3 The verbalization and cloze procedures

3.3.1 The verbalization procedure

The procedure of verbalization as a technique which utilizes readers' verbal protocols (or verbalized responses) seems at present the most promising in matters concerning the analysis of the cognitive processes involved in the reading comprehension of (native and non-native) readers. The design of experiments utilizing this method may vary among experimenters but the rationale behind this technique remains the same and aims at one specific target: to check the comprehension of subjects reading silently in NL or FL. Thus subjects are given reading passages (often followed by one or two comprehension questions on each passage) and are required to verbalize their mental activity while searching for responses, i.e. to "think aloud". They are also told that all their verbal protocols are ^{tape} recorded by the experimenter (who is often the language instructor). The role of comprehension questions following reading passages is normally to give a purpose to the reading subject whose motivation will depend on the type of questions asked. He will either scan the text i.e. go quickly through it in order to find the particular piece of information needed to answer the question,

or skim it i.e. run quickly his eyes over it to get the gist of it. In the technique of verbalized responses, the experimenter may therefore seek to check either discrete points or overall comprehension by following the thinking process of the subjects. Experimental methods based on the analysis of verbal responses seem to operate successfully in conjunction with the cloze procedure as the experimenter can analyse written protocols and check their validity against oral protocols.

3.3.2 The cloze procedure

The cloze procedure is no doubt the technique mostly utilized by researchers interested in empirical studies. It involves item deletion from a passage which may be totally random i.e. every nth word is removed "mechanically" from the text, or rational i.e. the deleted items are carefully selected and their replacement based on certain specific criteria (see Greene, 1975, Clarke 1979, and Bachman, 1982 on this subject). Item deletion is generally considered a reliable index of reading comprehension, and its ease of operation and versatility renders it an appropriate measuring instrument in empirical researches. As noted earlier, by removing an item from a passage, an obstacle interfering with the subject's reading process compels him to utilize problem-solving strategies to overcome this obstacle and proceed with his reading activity. However it may be argued that the choice of an obstacle is often based on considerations which interest the experimenter primarily. What may seem to him

a "good" item to remove because it fits nicely into his theoretical framework may not be felt as such by the subject. The notion of missing/unknown item implied by a cloze in a cloze experiment is an experimenter-defined notion which does not necessarily coincide with that of the experimentee. However in empirical research, some considerations have to be treated as secondary and assumed to have a non-significant effect on the results. Assumptions have to be made despite the fact that they may not always conform to the reality.

In the present investigation, we have utilized the rational procedure of item deletion. It consisted in removing carefully selected items whose replacement was based on two different criteria. In the Pilot Test (preliminary to the experiments), one lexical item with strong lexical (semantic and pragmatic) clues in its environment, was removed from each passage. The types of clues were not controlled systematically. They were generally intersentential although some texts allowed intrasentential clues to operate. In the Experiments, deletion consisted in carefully selecting lexical items whose replacement required comprehension of the relationships of lexical cohesion and coherence holding between elements of the text. Moreover deleted items were not limited to those drawn from the cohesive (text) system but included items drawn from the discourse system. (Deyos, 1937: 129 makes a distinction between "text cloze tests" and "discourse cloze tests", the former testing knowledge of the "language system" and the latter demonstrating learners' understanding of the text as it develops, of "the communication as a whole").

3.3.3 Cloze procedure combined with verbalization in the Pilot Test

The Pilot Test combined two experimental procedures viz cloze procedure and verbalization procedure. The Test was conducted with a sample population of 38 undergraduate students of mixed abilities enrolled at the University of Algiers as first, second and third year students reading for the Licence Degree (in English language and literature). Students were supplied with a panoply of 'choice' items which consisted of semantic features (in a loose sense of the term) arranged in three sets. Granted that lexical items of text project their semantic features onto other text items, the sets of features supplied by the experimenter had the purpose of enabling the subject to match them against the semantic features underlying the lexical items of the text. Only one response was allowed for each text and consisted in selecting one set of features A, B or C. One of these sets was a distractor. Alternatively these sets could have been choice lexical items (a list of 3 or 4 items could have been supplied to the students). But by supplying features and not words, our aim was to make the subjects sensitive to meaning rather than form. This was also to reduce problems that may arise from poor lexical competence. This method was seen as an alternative to wild guesses when choice lexical items are supplied and as an invitation for the subject to look at words not merely as physical entities but as bundles of semantic and pragmatic features that can be matched to other semantic and pragmatic features of text. Consider the example below:

Separating hob and oven gives you several advantages over a freestanding cooker. An oven at a convenient level makes life easier as it puts an end to bending and ----- heavy dishes from a low level. Here we focus on built-in ovens and test their cooking performance and ease of operation (Text 2)

A	B	C
-raising from one level	-preparing food	-place, position
-higher levels	-eating	-limited space
-use of hands	-use of heat	-centre
-bending	-oven, fire	-localize
-heaviness	-dishes	-find room for something

Thus each passage was complemented by three sets of 'features' as shown above(see also Appendix VIII). Testees were required to supply one set A, B or C as well as verbalize their responses or "think aloud". For instance subject n°15 selected the correct set A and her verbalization enabled us to follow her 'reasoning' as she was searching for clues to build up her response:

"We use hands...prepare food without losing time...from low level...higher levels...heaviness...use of heat...of course we use heat...centre...centre of the oven...no... B...dishes...you bend...and then you...you...raising from one level...set A...bending...yes it's A...to bending and raising heavy dishes...les soulever...les plats...set A...yes

Analysis of students' recorded responses and of their written protocols for cross-reference has allowed us to conclude that cloze technique combined with verbalization may be an ideal instrument for measuring reading comprehension strategies. The Pilot Test has provided us with general information and has allowed

the making of several hypotheses for future research. However the "think aloud" technique does involve problems regarding its utilization. If the combination of cloze and verbalization procedures seemed quite workable with a small sample of students, there are difficulties in administering it to a larger population (up to 100 subjects) and the results are likely to be contingent. The Pilot Test was essentially based on students' verbalization of responses and much time was spent on instructing the testees on the procedure itself. Normal reluctance on their part had to be removed gradually by giving them several practices before the actual test. Because they were tested on their mental activity 'from the inside' they tended to regard the experimenter as an intruder and some reacted quite negatively by not verbalizing their responses. Although the insights brought by this method are invaluable, it nevertheless carries some risk. There are personal and sociocultural factors related to it which are inhibitive. Some students found that verbalization interfered with their mental activity of clue and meaning searching during the practice sessions and attributed their wrong responses to this interference. The general conditions in which the Test took place were not entirely satisfactory: students asking for help hence disturbing others, some wishing to give up half-way through, others unhappy with their oral answers and wishing to correct them, a fairly high noise level and two defective tape recorders. The verbalization method was used with a sample population as a diagnostic test rather than an end in itself, to develop hypotheses about clue-searching and design further experiments on the basis of results. It was therefore quite instructive but would be operationally ^{difficult} with a larger population as it requires

considerable means(material and human) , and is time-consuming. For this reason, this research procedure was not utilized for the Experiments described in this study. But it is undeniable that verbal protocols are al ways illuminating when unclear cases are encountered even in the most carefully designed (cloze) experiments.

3.4 The Experiments: Aim and nature and research questions

The present research concerns the awareness and use of lexical devices of cohesion and coherence in reading comprehension by non-native readers of English. It may be better described as an exploratory experimentation with particular subjects. Thus this cross-sectional study involves a large number of subjects (in comparison to the Pilot Test) and seeks to determine difficulty order with a view to improving future FL reading instruction. Our purpose in conducting this investigation was also to try out the experimental procedures themselves since no work of this type has been done before in this connection. The format developed is the result of a pilot study conducted with a sample of undergraduate students reading English as a FL at Algiers University in the Autumn term 1985 (see section 3.3.3 above for description of Pilot Test and appendices VIII and VII for test passages and sample of subjects' verbal protocols). As stated earlier, the Pilot Test was aimed at tapping one of the comprehension strategies utilized in FL/NL reading viz inferencing unknown meaning and at obtaining information concerning the use of linguistic and pragmatic relations of cohesion and coherence in inferencing while reading. The Pilot Test suggested hypotheses for further research into

the exploitation of these devices in reading.

Following the Pilot Test, four experiments using cloze procedure were designed and administered to undergraduate students reading EFL at the University of Algiers in the Autumn Term, 1984. Each of them was aimed at testing a number of hypotheses relating to the use of lexical devices of cohesion and coherence in reading comprehension in an attempt to answer the following research questions:

- 1 How do FL learners utilize lexical resources of cohesion and coherence when inferencing unknown meaning while reading?
- 2 How does the use of lexical resources of cohesion and coherence vary as a function of FL proficiency?
- 3 How does the use of lexical resources of cohesion and coherence vary as a function of language background ?

The testees were expected to utilize linguistic semantic as well as non-linguistic pragmatic information when inferencing acceptable meaning. These two types of information were embodied in "clues". A linguistic clue is defined as a stretch of language standing in a definitional or "dictionary-like" type of (lexicosemantic) relationship with the cloze item (i.e. synonymy, antonymy, hyponymy, etc). For instance the relationship between cholera and disease is definitional. The two lexical items are semantically related via hyponymy. Cholera is a hyponym of the superordinate term (or hyperonym)

disease. A non-linguistic clue, or pragmatic clue, on the other hand describes a stretch of language standing in a non-definitional, pragmatic type of relationship with the missing item, as for instance between cholera and war. These two items are not related semantically essentially but via pragmatics (see chapter 1 for the distinction between linguistic and pragmatic meaning). Linguistic and pragmatic clues appeared in various contexts in the experimental texts, but also occurred with other linguistic and/or pragmatic clues that were difficult to eliminate. For example, in Experiment 1 (see Texts 1-20 in Appendix II), Text 15 (condition 1: clue before gap) and Text 16 (condition 2: clue after gap) both contained additional linguistic and pragmatic clues that could not be eliminated without disturbing the topicality of the text. Examine T15 and T16 below:

T15 The lack of organization in the Crimean war was appalling. Cholera and dysentery were widespread. Many soldiers died of ----- . Mary was disturbed by the dreadful stories about the war which came back to her.

T16 The lack of organization in the Crimean war was appalling. Many soldiers died of ----- . Cholera and dysentery were widespread. Mary was disturbed by the dreadful stories about the war which came back to her.

In T15 the lexicosemantic clues cholera and dysentery precede the gap. In T16, they follow it. However war is a strong pragmatic clue to the gap in both texts and subjects were misled by this thematic clue thereby supplying war instead of disease as response (see the discussion of this point with 'Miscue Analysis' at the end of this chapter). Thus pragmatic

clues are powerful indices of discourse coherence to the extent that they can obscure linguistic clues in the text. Also some intrasentential clues such as died (whether preceding or following the gap) could not be eliminated, although when the sentence 'Many soldiers died of -----' was tried on native speakers to check the degree of collocation between died and the missing item, this sentence produced responses such as thirst, hunger, wounds, boredom, returning home, the war but not diseases. This led us to believe that diseases could only be recovered in connection with the linguistic semantic clues cholera and dysentery. This last point suggests that diagnostic testing with native speakers is not always reliable and in this particular instance did not account for intersentential relationships of lexicosemantic and lexicopragmatic meaning.

Each of the four hypotheses formulated for each experiment implied a different independent variable whose values were being compared. The results of the experiments are presented in section 3.7 below on a two-part format. First a 'descriptive' analysis of the 'main effects' which analyses the results of each individual variable and an assessment of them inferentially ("comparisons"). Then an examination of the interactions between the different variables and an inferential assessment of these interactions for statistical significance (for instance 'YEARTIM BY CONDITIO' refers to the interaction between the two variables year of study and experimental condition).

(i) The hypotheses

Four hypotheses were formulated in this research in order to answer the following three research questions (already noted above):

- 1 How do FL learners utilize lexical resources/links of cohesion and coherence when inferencing unknown meaning while reading ?
- 2 How does the use of lexical resources/links of cohesion and coherence vary as a function of FL proficiency ?
- 3 How does the use of lexical resources/links of cohesion vary as a function of language background ?

These hypotheses concern the exploitation of linguistic lexical and pragmatic lexical clues in inferencing defined qualitatively in terms of their order, distance, linguisticness and non-linguisticness. Four experiments to account for these hypotheses, referred to as null hypotheses in empirical research, have thus been designed.

The purpose of experiment 1 (see Texts 1-20 in Appendix II) was to investigate the effect of the experimental variable 'order of clue' on the success of learners' inferencing a plausible filler for the 'gap' by measuring their performance on two different cloze conditions, or experimental tasks. The null-hypothesis formulated for this experiment was:

"There is no effect of the order--immediately before or immediately after, in which a crucial clue comes relative to a gap on the success of learners' inferencing a plausible filler for the gap".

Two experimental conditions were examined: in condition 1 the clue(item) was before the gap. In condition 2 it was after the gap. The clue was a lexical item standing in one of the two lexical relationships semantic or pragmatic, with the cloze item. For example:

T3 Some shops sell cushions either with natural or with synthetic fillings. However, the oil in feathers dries out, leaving them dry and brittle, and eventually they turn to dust they Fire is also to be thought about.

T4 Some shops sell cushions either with natural or with synthetic fillings. However, feathers can -----: the oil dries out, leaving them dry and brittle, and eventually they turn to dust. Fire is also to be thought about.

The clue to the cloze item disintegrate in T3 above (for condition 1) is embodied not in one lexical item but in the whole sentence preceding the gap viz 'the oil in feathers dries out, leaving them dry and brittle and eventually they turn to dust' which describes the process of (feather) disintegration in pseudodefinitive terms, that is in a non-conventional fashion. Oil dries out, dry, brittle, turn to dust are the most important lexical pragmatic elements of this pseudodefinition. This clue may be viewed as lexical pragmatic since its lexical content is more pragmatic than semantic and connects to the process of feather disintegration specifically rather than to that of disintegration generally. Also the presence of the dash as an index of the discourse function 'definition' is an important discourse clue to the cloze item disintegrate. On the other hand fire cannot function

as a clue although it is thematically prominent.

In T4 the pseudodefinitional lexical clues occurring after the colon (implying explanation, illustration) are now after the gap, i.e. oil, dries out, dry, brittle, turn to dust. Furthermore in both T3 and T4 the occurrence of the syntactic marker 'However' preceding the gap may be confusing for it may be taken as a linguistic clue to imply that feathers cannot sell (In actual fact this marker did confuse some subjects, See the Miscue Analysis at the end of this chapter).

Predictions were made that inferencing a filler when the clue was 'anterior to the gap' would be easier than 'after the gap'. These predictions were made on the basis of the results of the Pilot Test administered to other subjects before the present experiment and also on common knowledge that searching for clues backwardly seems a more natural reading strategy than searching for clues forwardly.

Experiment 2 (see Texts 21-40 in Appendix II) was aimed at testing the effect of the experimental variable 'Distance of clue' in the performance by subjects of a cloze task. The null-hypothesis formulated for this experiment was the following:

"There is no effect of the distance of a preceding clue relative to a gap - whether in the immediately preceding-gap arrangement (that is, clue-gap-possible disclue) or in the distantly-preceding-gap arrangement (that is, clue-possible disclue-gap) - on the success of learners 'inferencing a plausible filler for the gap'".

This experiment is similar to experiment 1 in the sense that it also involves the order of a clue relative to a gap: in condition 1 (Experiment 1) the clue was immediately preceding the gap. In condition 1 (Experiment 2) the clue is immediately preceding the gap but it is distantly preceding it in condition 1.

(Experiment 2) as a sentence was inserted between the clue clause and the gap clause to create a distance between the clue and the gap. This sentence contained a potential distractor or disclue, often equally cohesive with the cloze item* but whose lesser degree of coherence made it an inadequate candidate for the gap. For example:

T21 A debate has raged about the control of the British Police. To the -----, such control should be assured by the superior members of their staff. To the critics, there should be locally elected committees responsible for the local community.

T22 A debate has raged about the control of the British Police. To the critics, there should be locally elected committees responsible to the local community. To the -----, such control should be assured by the superior members of their staff.

The linguistic clue for the cloze item police in T21 (condition 1) is located in the sentence immediately preceding the gap clause i.e. (British)police. A non-linguistic clue may also be noted in the gap sentence, staff whose role as a clue is minor but nonetheless important. However in condition 2 (T22) the clue (British)Police occurs two sentences away from the gap and as a result of its position may be less effective as a clue. Also the presence of the non-linguistic clue staff at the end of the text is somewhat subdued and may be unlikely to fulfil this function. The intrusion of critics in the medial sentence, an item coherent non-cohesive with (British)Police is ⁱⁿ this sense distractive and contributes to weaken the role of (British) Police as the main clue.

Predictions were made that the further ahead the clue would occur the less effective it would be as a clue. (See Text 28 where sauce the cloze item is cohesive with vegetables occurring in the distractor sentence) and also with salad, the normal clue to the cloze item occurring in the initial sentence.

occur relative to a gap the less likely it would be to recover intended meaning.

Experiment 3 (see Texts 41-76 in Appendix II) named "Conjunctive clue" was developed to examine the differential effect of the degree of explicitness of a "conjunctive" clue in the performance of a cloze task by subjects. A null hypothesis was formulated for this experiment:

" There is no effect of the logical connection, whether overtly signalled by a conjunction, or a lexical equivalent to a conjunction, or covert, that is, not signalled, relative to a gap, on the success of learners' inferencing a plausible filler for the gap".

Three experimental conditions were manipulated:

Condition 1 : in this condition, the conjunctive clue was an overt syntactic linker or marker (but, and, furthermore, etc.).

Condition 2 : the conjunctive clue was a covert linker, that is not signalled via syntactic markers but implied. Thus, logical relations of additivity, adversativity, causality, for instance underlying the juxtaposition of propositions were deducible from the lexical content of these propositions. In other words, propositional content expressing conjunctivity could be embodied in the text's lexical semantics (for example, Mary wore a blue dress. Jane wore a white one, where blue and white signal a linguistic relation of cohyponymy (see section 1.7 in chapter 1), or in its lexical pragmatics (for example, John was happy. He won the polls, where happy and won underlie a causal relation produced non-linguistically via pragmatics. These linguistic and pragmatic clues deducible from propositional content were actually there additionally in condition 1, but

their role was singled out and their function became more vital in condition 2 when syntactic markers were no longer present.

Condition 3 : in this condition, the conjunctive clue was an overt lexical linker, that is, a lexical paraphrase of the syntactic marker -that we shall ^{call} "paralexical clue". Thus, it may be added, it is true that, by contrast, as a result, are lexical paraphrases of the syntactic markers, furthermore, really, however, therefore, Note that condition 2 and condition 3 were both 'lexical' but in different ways. For example:

T56 Jose de Molina says that Argentine writers should adhere to the tradition of Spanish literature. But I say that Argentine literature can be defined as a desire to become ----- Spain. The search for European themes is a well-known phenomenon in 20th century literature.

T57 Jose de Molina says that Argentine writers should adhere to the tradition of Spanish literature. I say that Argentine literature can be defined as a desire to become -----Spain. The search for European themes is a well-known phenomenon in 20th century literature.

T58 Jose de Molina says that Argentine writers should adhere to the tradition of Spanish literature. This disagrees with my definition of Argentine literature as a desire to become ----- Spain. The search for European themes is a well-known phenomenon in 20th century literature.

In T56, the clue to the cloze item separated from is obvious. The syntactic marker But signals adversativity and therefore the cloze item must carry the meaning of 'not adhering to something', i.e. 'independent'. Furthermore, this linguistic clue is ^{the} complement of the pragmatic lexical clue I(say) which contrasts with Jose de Molina(says). But because of its

explicitness, the syntactic marker But may be more powerful as a clue in this context than the lexical clues which have to be deduced from propositional content.

In T57 the explicit marker of contrast But has been removed and the text now relies on its lexical semantics and lexical pragmatics entirely. Then the role of the implicit lexico-pragmatic clue Jose de Molina (says) becomes prominent as it is responsible for the meaning of 'not adhering to something' in the gap sentence and thus contrasts with I(say). The coherence of this text is mostly produced by its lexical pragmatics.

In T58 the contrast underlying this text is recoverable through the presence of the paralexical clue This disagrees with. This clue is 'lexical' in the sense that its basic form is overtly lexical (as opposed to but) but in combination with other items it functions as a syntactic marker of adversativity. Paralexical markers may be best described as hybrid forms between syntactic markers and lexical items. (Some syntactic markers may have been full lexical forms originally e.g. furthermore but their synchronic form is no longer regarded as lexical). Thus this disagrees combined with my (definition, contrasting with Jose de Molina's definition), are lexical pragmatic clues to 'not adhering to' i.e. being separated from Spain .

The objective in discriminating between overt and covert clues was to see which type of "conjunctive" device of cohesion/coherence was a successful predictor of a missing item in text, that is, whether deductions regarding its meaning could be made in the absence of explicit markers. Predictions were made that the overtness of the "conjunctive" clue (syntactic or paralexical) would facilitate inferencing, and the result would

be no difference in performance between these two conditions overall. On the other hand it was anticipated that the covertness of the clue would be problematic and would generate lower performance overall.

Experiment 4 (see Texts 77-96 in Appendix II) sought to examine the effect or otherwise of the experimental variable "Linguistic clue" on students' performance of a cloze task. The null-hypothesis formulated was:

"There is no effect of the presence of a linguistic or a non-linguistic clue relative to a gap on learners' success in guessing a plausible filler for the gap".

This was measured by two experimental conditions:

Condition 1: the clue was linguistic and occurred in anterior or posterior position to the gap.

Condition 2: the clue was non-linguistic and occurred in anterior or posterior position to the gap. The purpose of separating linguistic and non-linguistic information in this experiment was to see whether linguistic meaning (explicit) produced by lexical relationships of cohesion was a more successful predictor of missing items than non-linguistic meaning (implicit) produced via lexicopragmatic relationships of coherence. For example:

T87 Some people are capable of vandalizing their country, transforming it into a place without history or beauty. They live in the immediate present and are unaware of historical continuity and without culture. ----- today is a widespread social phenomenon.

T88 Some people are capable of treating their country the way some teenagers today treat buses and phoneboxes. They live in the immediate present and are unaware of historical continuity and without culture. ----- today is a widespread social phenomenon.

The strong clue to the cloze item vandalism in T87 is linguistic (vandalizing) although morphologically different and occurs in the initial sentence. A non-linguistic clue may also be noted: the clause transforming it into a place without history or beauty which functions as a pseudodefinition to the cloze item vandalizing. This is called pseudodefinition because reference to no history or no beauty is more metaphorical, connotative (pragmatic) than denotative, strictly definitional (semantic). Note that the theme of history and culture being relatively redundant in the text, it may be a potential factor of confusion.

In T88, the clue is pragmatic and will only be perceived by the reader if he has background knowledge that teenagers treat buses and phoneboxes badly (as most of them do in many countries!). The pragmatic clues unaware and without culture in both T87 and T88 may be useful indications that some negative statement is being made in the last sentence. There are no linguistic clues to the cloze item in this text, which makes its meaning totally dependent upon pragmatic information.

Predictions were made that inferencing a missing item when the clue depended on lexical pragmatic linkage alone would be problematic to the testees because of its lack of explicitness. Lower performance overall was expected in this condition.

Thus the aim of the four experiments was to test whether subjects were sensitive to the quality of the textual/discoursal clues rather than their quantity. The subjects' ability to recognize and to recover lexical links when inferencing unknown items will be a function of the amount of knowledge they bring into the text whether linguistic semantic, to exploit the different

types of lexical semantic clues, or non-linguistic to exploit lexical pragmatic clues.

(ii) Some considerations on clozefication(or item deletion)

The suitability of an item as a cloze item(hence deletable from a text) was measured in terms of the following criteria:

(i) The cloze item was a noun(accident-T1), a verb(disintegrate T3) or a determiner(usually a predeterminer (synthetic-T6), but not a syntactic marker.

(ii) The cloze item was a simplex or a complex linguistic item with two members only(accident-T1 or successful with-T74).

(iii) The cloze item was theme or rheme in the sentence from which it was ^{to be} deleted^{///}. For example:

The //passengers// amused themselves by reading the difficult names(T25).....THEME.

So we have prepared the perfect //sauce// to meet your needs (T27)RHEME.

(iv) The cloze item was a member of a lexicosemantic or lexico-pragmatic relation(which could additionally be lexico-referential or lexico-substitutional). For example:

lexicosemantic relation: nylon-synthetic (T5) : synthetic was removed.

lexicopragmatic relation: drowned-accident(T1) : accident was removed.

(v) The cloze item was being repeated as same or derived form: fried-fry(T7): fry was removed ; towns-townspeople(T11): townspeople was removed.

(vi) The cloze item enjoyed at least one intersentential clue

(linguistic or pragmatic) in addition to the other member of the lexical relation to which it belonged. For example: However (linguistic syntactic clue)-watchtowers (cloze item)-enemy (pragmatic lexical clue) in T65.

3.4.1 The subjects and the Placement Test

3.4.1.1 The subjects

A sample of 90 subjects whose names were randomly selected from approximately 700 students officially enrolled for the three-year Licence Degree in English in the department of English (Institut des Langues Etrangères) was used for the Placement Test. The names of the students were randomly drawn through a mechanical procedure of shaking a dice to avoid bias in the selection, out of this population. This sample is representative of other universities and colleges in Algeria.

3.4.1.2 The Placement Test

Because this study also targeted specific language dominant groups, a Placement Test was administered to these 90 students in order to discriminate between the Arabic dominant subjects and the French dominant ones. The test lasted fifteen minutes approximately, including instructions and trial tests, but the actual Placement Test lasted five minutes. It consisted in having students "brainstormed" to identify in Arabic or French (that is, in any of the two languages that they speak and/or read and/or write) as many items as the subjects could associate with each of the following ten topics, and in written form:

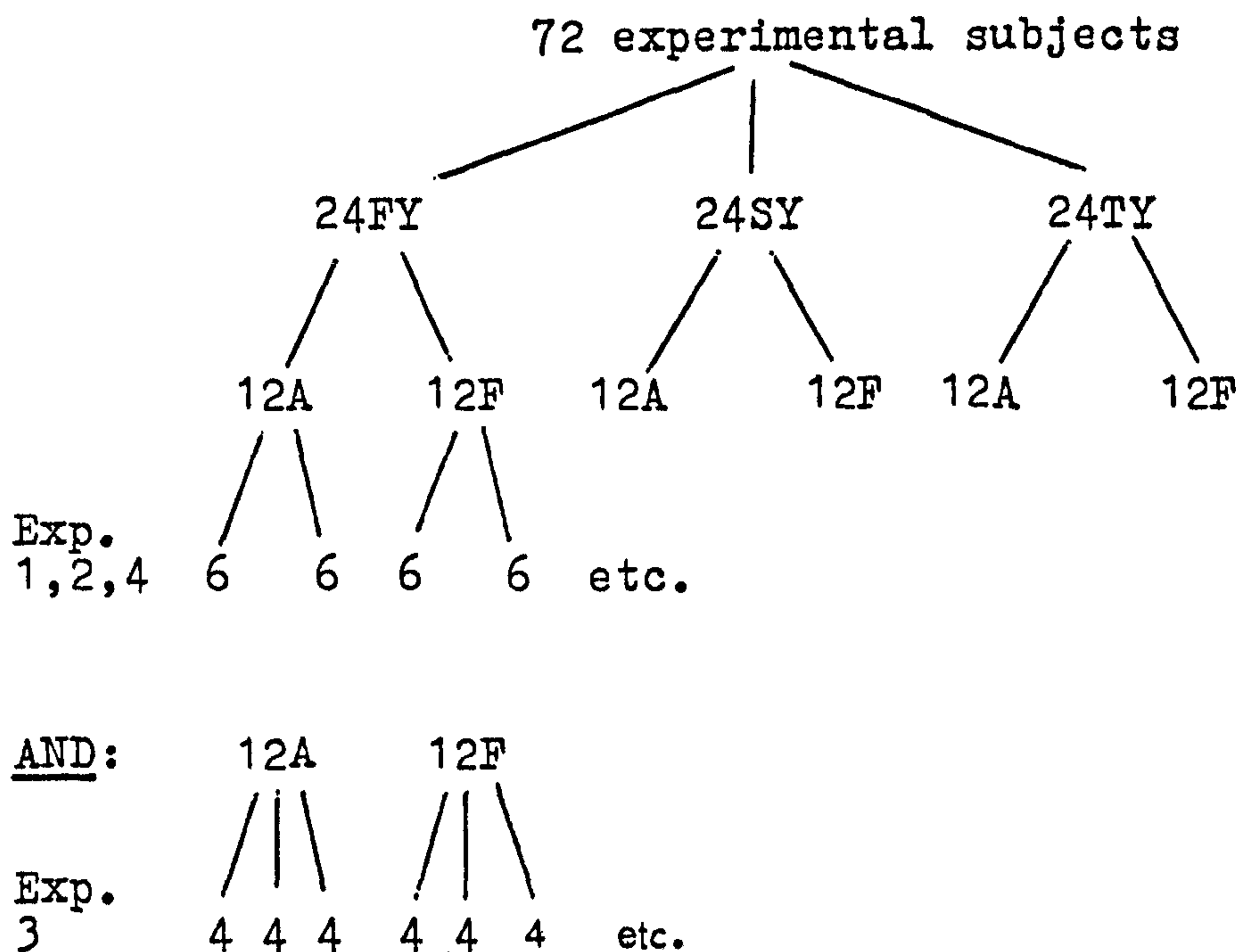
- 1 Football.
- 2 Cooking.
- 3 At the hospital.
- 4 Repairing a car.
- 5 Politics and the government.
- 6 Visiting a TV factory.
- 7 At the University.
- 8 Travelling abroad.
- 9 At the cinema.
- 10 Shopping.

This wide variety of topics had the purpose of obtaining an authentic account of the language that the students would use across various fields.¹ Thereafter, those students who supplied all Arabic responses were categorised as "Arabic-dominant" and entered in the A Group, and those students who provided all French responses were entered in the "French-dominant" group. Unclear cases such as those who answered in Arabic for some topics and in French for others were discarded because they could not fit in either category. (Those individuals were referred to as "bilinguals" in a loose sense). The subjects were given half-a-minute to answer each item, that is, five minutes for all ten items. Only 72 subjects were retained after this preliminary test in order to equalise the number of students in the experimental conditions.

3.4.2 Selection for experiments

The 72 subjects were male and female, between 19 and 21 years of age (there were 8 mature students aged 30-35), of diverse geographical and socioeconomical backgrounds, which did credit to the representativity of this university population. They had different levels of proficiency in English as an FL: all came from the first, second and third year of the undergraduate course in English leading to the "Licence d'Anglais". The 72 subjects were made of 24 subjects for each year group (that is, first, second and third year), and each year group was subdivided into two language dominant groups, A and F. Each language group was then randomly split into subgroups of six subjects for experimental design purposes (see Section 3.5 below). The following diagram shows the general distribution of the population for four experiments:

Diagram 8



FY = First Year

SY = Second Year

TY = Third Year

A = Arabic-dominant Group

F = French-dominant Group

3.5 The design

3.5.1 The variables

Seven experimental variables have been identified in this research. They include four independent variables (IV) in any one experiment, viz. (1) 'Order of clue'; (2) 'Distance of clue'; (3) 'Conjunctive clue'; and (4) 'Linguistic clue', and two moderator variables in the sense defined by Hatch and Farhady (1982), that is,

'subject' variables: yeartime(5) and language dominance (6). The dependent variable (DV) being measured was the success of learners' inferencing a plausible filler for the gap and therefore was represented by the students' scores (7). (Plausible is to be understood as "acceptable" as measured with a sample of native speakers, rather than the exact word omitted). Thus, each score was dependent upon the experimental condition variable, the yeartime variable and the language dominance variable.

5.2 The procedure

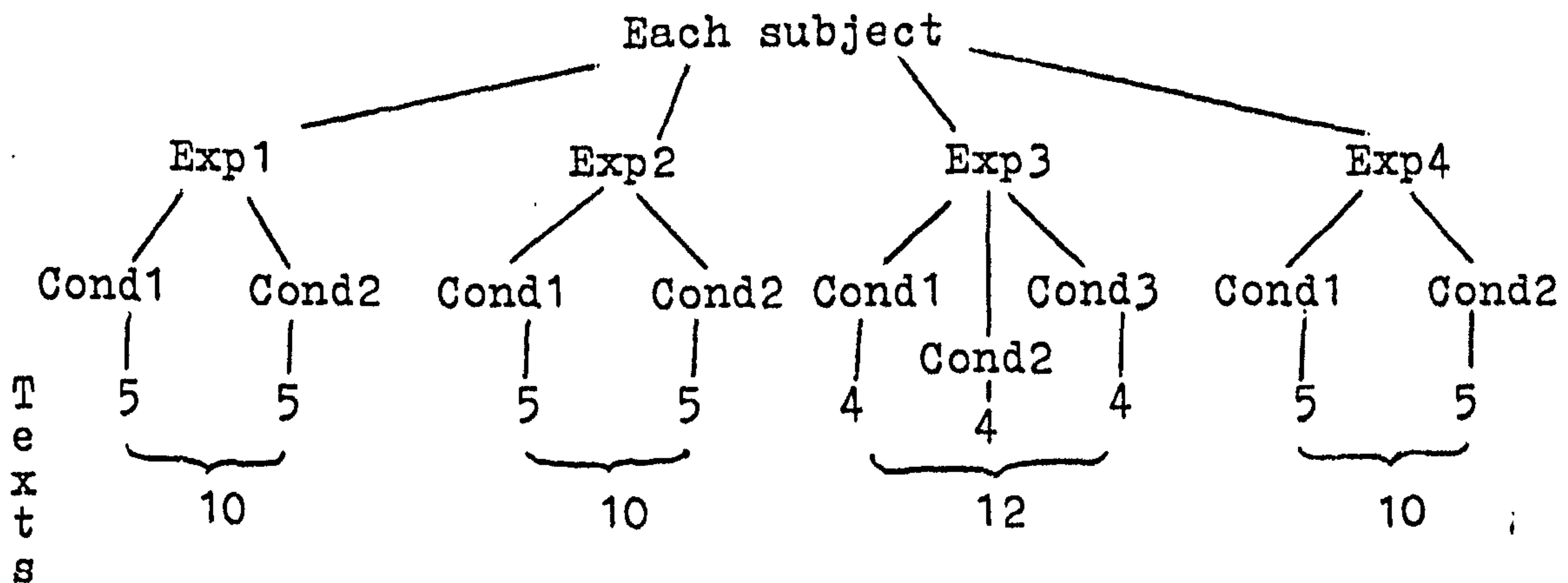
Two (or three) experimental conditions were manipulated in each of the four experiments to observe the effect of the variations on the subjects. Other subject variables such as yeartime and language dominance were obviously not manipulated. The design was repeated measures. Matched sets of texts were used so that each subject did not see the same text in both conditions. Thus, the texts had a different random order for every six subjects. The table below gives a visual account of the distribution of texts and subjects for experiment 1, as an example:

Table 0.1

Experiment	Subjects A	Numbering of texts given to each person for both experimental conditions	Text Total
1. "Order of clue"	S1 to S6	T1, T3, T5, T7, T9 (C1) T12, T14, T16, T18, T20 (C2)	= 10
	S7 to S12	T2, T4, T6, T8, T10 (C2) T11, T13, T15, T17, T19 (C1)	= 10
Pattern repeated for Subjects F			

Thus, condition 1 and condition 2 (and condition 3 for experiment 3) were completely counterbalanced across all twelve subjects in each group A and F. Six subjects chosen at random completed condition 1 first and six subjects completed condition 2 first. For experiment 3, four subjects completed condition 1 first, four other subjects completed condition 2 first and four completed condition 3 first, and so on. Each testee was given a booklet which contained ten texts for each experiment (but twelve texts for experiment 3) as the diagram below shows:

Diagram 9



The booklet had a cover sheet with instructions emphasising that only one word had to be supplied, and blank responses were not accepted. It also stressed that responses could be in English, Arabic or French. Thus, in order to eliminate the effect of the productive command of English on the cloze responses, the students were allowed to use any of their background language. It was assumed that the switching of the students to their background language to supply a meaning (meaning had to be supplied, not words) will not affect overall comprehension. However, how much native language is involved in decoding and interpreting an input in the foreign language, and how much interactions there are is a question not easy to disentangle and certainly needs experimental exploration. The utilisation of one of the background languages by the subjects has also eliminated all sorts of errors (spelling, syntactic inadequacies) resulting from 'learners' strategies' (Corder, 1967, Richards, 1974), as for instance, the use of French phonology to spell an English word, as well as errors defined as manifestations of 'interlanguage' phenomena (Selinker, 1972). Above all, this has avoided the non-production of a response. However, there may be some limitations to this method which lie in the difficulty or perhaps impossibility to replicate these experiments when the researcher lacks knowledge of the subjects' background languages. The experiments took place on different dates to avoid fatigue and boredom generated by recurring experimental

conditions. However, response analysis of the data suggested a psycholinguistic universal. The tendency for a subject to develop expectations or a "psychological set" which sometimes resulted in misinterpretation of subsequent messages. For instance, the fact that Text 14 followed Text 12 in some of the copies of Experiment 1 (Text 12 was about the Arab dynasties in Spain) attracted the wrong filler Muslim rather than Roman for Text 14 in several responses.

Thus, the deletion of semantic cohesive and pragmatic coherent lexical items from passages could be a valuable measure of reading comprehension. Such deletions as implied by Alderson (1979) derive from aspects of the reading process rather than the traditional random deletion

procedure (that is, every nth word). Some attempts seem to have been made at deleting grammatical cohesive items only (Levenston et al, 1982).² As Deyes (1984: 128) suggests, "discourse cloze" tests should be used to "reflect the reader's ability to follow information through the text and use contextual clues as well as co-textual ones" so that we are testing knowledge of the language system that involves decoding of the text's lexical semantics and understanding of the communication as a whole which requires the drawing of inferences, the sharing of similar presuppositions with the writer, the understanding of the writer's intentions, all of which are involved in the interpretation of the lexical semantics and the lexical pragmatics of the text.

Thematic as well as rhematic lexical items have been clozefied. Accordingly, clues relative to the gap could be either theme or rheme.

3.5.4 The scoring method for the dependent variable (DV)

In scoring native speakers' performance on cloze tests, the conventional method has usually been to count the number of times a student produces the exact word used by the author (Taylor, 1953), but this method is often viewed as archaic and instructors seem to no longer require exact word replacement on the part of students. In SL/FL research, the tendency has been to use scoring systems which give credit to contextually acceptable responses, but the question of whether there is an optimum scoring method remains a matter of debate. A

number of studies show very high significant correlations between exact and acceptable scoring (Oller, 1972; Stubbs and Tucker, 1974). However, it seems justifiable to accept an answer which satisfies some of the constraints on a 'blank' or 'gap' than one which ignores all types of constraints. In the present research, the items deleted measured the subjects' sensitivity to specific semantic and pragmatic constraints (see Cziko, 1978, for a comparative assessment of these constraints in NL and SL reading) when inferencing unknown meaning. The emphasis was on the process they were engaging in rather than the right-or-wrong product. The scoring method adopted in this study is an adaptation of Oller et al (1972), Clarke and Burdell's (1977) and Clarke's (1979) scoring methods. Responses are accepted in any of the three languages, Arabic, French or English, and are categorised on six discrete points (or 'subjective categories').

1 Entirely acceptable. Exact response. Word is that deleted by the research, or exact translation of it in Arabic or French.	} Response is 'correct'
	(C)

2 Contextually acceptable.

The expected response and the observed response are synonymous given the context.

3 Response fits the larger (passage) level context but it changes the meaning of the sentence slightly.

4 Response violates passage-level meaning constraints and changes meaning of sentence slightly.

} Response
is
'near-
correct'
(NC)

5 Response violates obligatory meaning constraints and seems unmotivated by any substantial degree of comprehension.

6 Blank (or no response) violates obligatory meaning constraints at sentence and passage-level.

} Response
is
'wrong'
(W)

Each subject supplied forty-two responses for the series of experiments, that is, ten responses for each experiment except experiment 3 which required twelve responses. Responses were rated in the following way:

Correct = 3

Near-correct = 2

Wrong = 0

The decision for 'weighting' the responses 3.2.0 rather than 2.1.0 was based on semantic and pragmatic considerations. Given that near-correct responses were closest to correct responses in semantic acceptability

and pragmatic likeliness than to wrong responses, it therefore appeared that the span of values given to the scores should be greater between near-correct and wrong responses than between near-correct and correct responses. This seemed to reflect semantic and pragmatic constraints more realistically.

Each subject could obtain a maximum score of 15 points in each experimental condition in a given experiment to the exception of experiment 3 where the maximum score was 12 in each condition.

3.6 The experimental variables

3.6.1 The moderator variables

There are two moderator variables for each experiment, viz. the yeartime IV and the language dominance IV.

3.6.1.1 The yeartime independent variable

This is a composite variable with two conflated factors: the year of course and the time allowed. The year of course was determined by the overall academic achievements of the subjects. The time factor was closely related to the year of course in that first year students were likely to need more time for each test item than second and third year students. Thus, the amount of time allowed for each test item/text for each macrogroup of twenty-four subjects was inversely proportional to the year of course of the subjects. It was distributed in the following way:

Year 1 = 2 minutes per text

Year 2 = 1 minute per text

Year 3 = $\frac{1}{2}$ minute per text

Each experiment lasted between 15 minutes and 30 minutes approximately, including instructions to testees.

3.6.1.2 The language dominance independent variable

This moderator variable has been defined on the basis of a Placement Test (see Section 3.4.1 above). The Arabic-dominant group (or A group) was made up of students whose dominant language at home and outside home was Arabic. The French-dominant group (or F group) was composed of students whose dominant language at home and outside home was French. Only the most extreme cases of Arabic and French users were considered for the experiments.

3.6.2 The materials

3.6.2.1 Selection

A total of 96 passages of written discourse was selected for the four experiments and arranged for each one in the following way:

Experiment 1: 20 texts (that is, 10 texts x 2 conditions)

Experiment 2: 20 texts (that is, 10 texts x 2 conditions)

Experiment 3: 30 texts (that is, 10 texts x 3 conditions)

Experiment 4: 20 texts (that is, 10 texts x 2 conditions)

The texts were conceived in pairs (in a loose sense, that is, they were equivalent in content but different in form - see Appendix II). The first member of the pair was designed to represent one experimental condition, and the

second member for the other experimental condition. For experiment 3, 30 texts formed triplets where the first member represented condition 1, the second member, condition 2 and the third, condition 3. Each text contained a blank or "gap" from where one simplex lexical item (sometimes a complex lexical item) had been removed. The overall length of the texts as well as their general lexical difficulty was within the competence of the testees.

The texts were extracted and/or adapted from four main sources: The Guardian Newspaper, The Economist, New Society and Good Housekeeping, all of which were issued in 1984. The texts were of medium size, that is, a minimum of thirty words and a maximum of forty words. Each one contained at least one neutral clause 'n' (in some texts, the 'n' clause was optional) which usually contained no clue and often a potential disclue, one clue-clause 'c' which contained the clue, usually though not necessarily linguistic and one gap-clause 'g' where the gap occurred.

3.6.2.2 Text design for each experiment

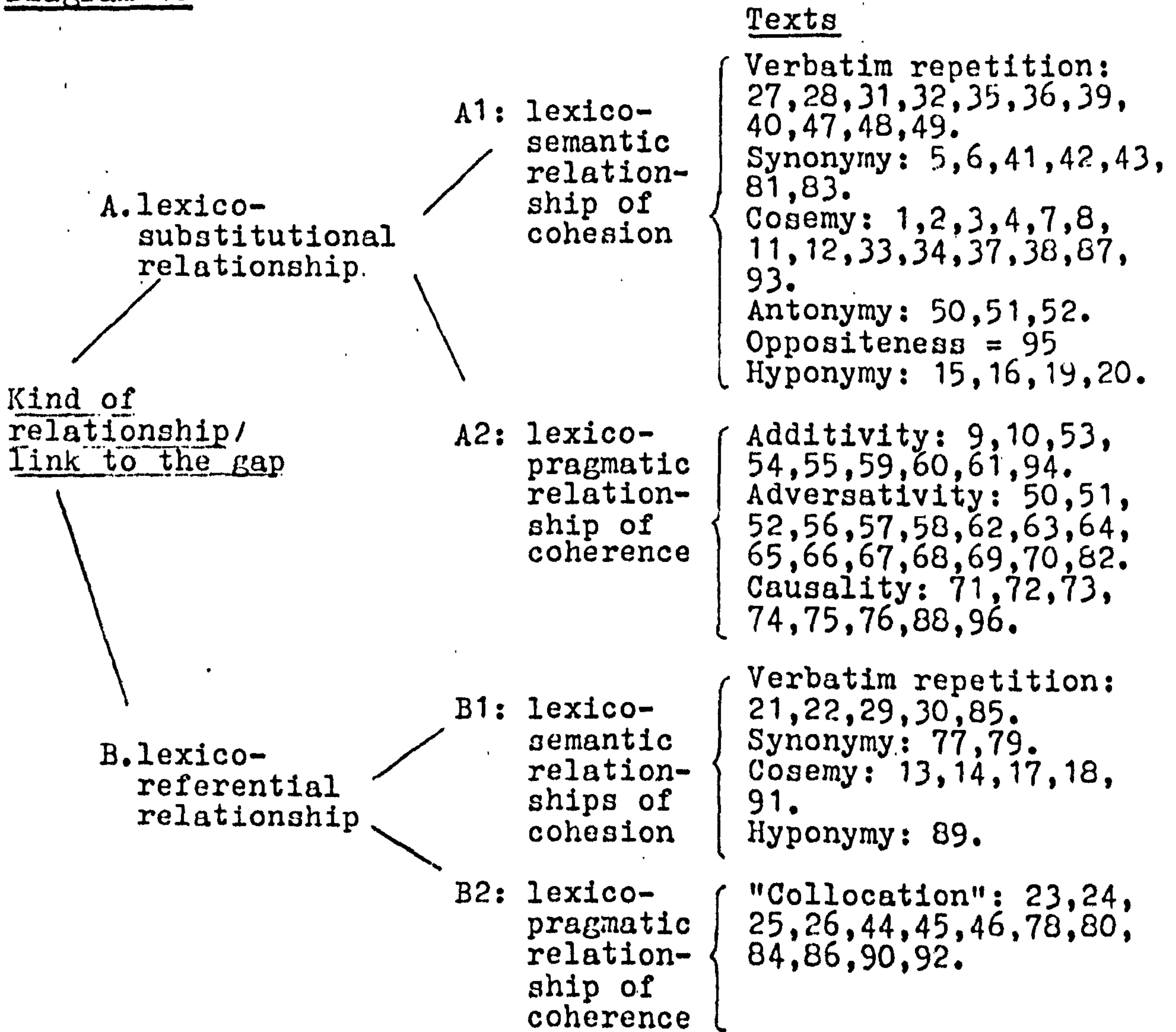
3.6.2.2.1 Control of 'kind' of clue

The texts utilised in the four experiments basically involve two kinds of clues - lexicosubstitutional and lexicoreferential and both 'kinds' of clue may also express conjunctive meaning (see Chapter 1, Section 6.0).

The clue(s) relative to the gap stand(s) in referential

or substitutional relation with the missing item which may be based on semantic criteria, that is, they underlie defining relations of synonymy, antonymy, hyponymy, cosemy, or on pragmatic non-defining criteria. The diagram below illustrates the kinds of clues being utilised in the experimental tests.

Diagram 10



To the exception of experiment 1 which measured the order of the clue relative to the gap, most types of clues, in experiments 2, 3, and 4, were retroactive or anterior to the gap*.

* This may be an indication that writers prefer backward to forward reference, and seems to be paralleled with the general approach to reading and inferencing while reading. Text producers are text receiver too and may be influenced by their reception strategies in production.

3.6.2.2.2 Control of 'type' of text

Experiment 1 measured the effect of the order of the clue on the filler and involved two conditions: 'clue before gap' versus 'clue after gap'. Ten texts were selected to treat each condition. In condition 1 ('clue before gap') the texts had loosely the pattern ncgn (that is, neutral-clause+clueclause+gapclause+neutralclause). In condition 2 (clue after the gap) the texts pattern was ngcn (neutralclause+gapclause+clueclause+neutralclause). The position of c was thus changed in condition 2, but the two n-clauses remained intact. In order to give balance to the experimental conditions and avoid confusion created by different text patterns, two n-clauses rather than one only have been supplied for each text, and their position, initial or final, has been kept constant throughout.³ This has eliminated any possible effect produced from changing the position of the n-clauses. Although these clauses were not essential to the texts structure, their presence was thought to heighten 'naturalism'. The distance in clauses between the clue and the gap was kept constant, that is, the clue always occurred in the c-clause immediately preceding or immediately following the g-clause but the distance in words between the clue and the gap varied between two words and ten words. However, any strong ^{clue} within the clause or the sentence where the gap occurred has been eliminated as far as possible although this was often difficult to do without disturbing the overall balance and intelligibility of the texts. (It is worth noting that 'clue within the same sentence' versus

'clue in another sentence' could have been another point to test). A test to check the degree of awareness of intraclausal and interclausal clues, and intrasentential and intersentential clues in inferencing was carried out with native speakers⁴ prior to these experiments which showed that the distance words between clue and gap was unlikely to affect the filler, and thus allowed intersentential cohesion to operate exclusively.

Experiment 2 measured the effect of the distance of the clue on the filler and involved two conditions: condition 1, 'clue immediately preceding the gap' and condition 2, 'clue distantly preceding the gap'. Experimental condition 1 had the pattern cgn (clueclause-gapclause-neutralclause) and ten texts were used for this condition. Condition 2 had the pattern cng (clueclause-neutralclause-gapclause). Ten texts were selected for this condition. Thus, in this condition the c-clause and the g-clause were being separated by an n-clause which contained a potential disclue functioning as a distractor for the reader. The n-clauses could not be kept constant in this experiment.

Experiment 3 measured the effect of a conjunctive clue on the filler and involved three conditions as follows:

Condition 1: the clue was an overt syntactic linker (sometimes referred to as "discourse marker").

Condition 2: the clue was covert and the linkage was supported lexically essentially.

Condition 3: the clue was an overt lexical paraphrase of a syntactic linker. This is referred to as paralexical clue. Thirty-six texts were selected all together mapped out into twelve texts for each experimental condition. Unlike experiments 1 and 2, four types of text patterns were allowed in this experiment and were used in the three experimental conditions. These patterns are:

cgn = clueclause+gapclause+neutralclause.

gcn = gapclause+clueclause+neutralclause.

ncg = neutralclause+clueclause+gapclause.

ngc = neutralclause+gapclause+clueclause.

This variety of patterns can be explained by the diversity of the types of logical connection being tested. Three main types of logical connection were involved in the experiment: additiveness, adversativeness and causality (as defined in Chapter 1). Note that the n-clause occurred either in final or in initial position relative to the gap but never separated a c-clause from the g-clause, that is, none of the texts had the pattern cng or gnc.

Experiment 4 measured the effect of a linguistic clue on inferencing a filler. It involved two conditions. In condition 1 the clue was linguistic. In condition 2 it was non-linguistic/pragmatic. Ten texts were utilised for each condition and most of them had the pattern cng (clueclause+neutralclause+gapclause).

3.6.2.2.3 Some considerations on text editing

authentic material was used in all experiments. However, it has been necessary to edit the original version of texts in order to accommodate the different variables under investigation i.e. some changes had to be made to their structure. These structural modifications have no doubt had an effect on the semantics and the pragmatics of some of the texts, specifically they have entailed changes in the thematicity of the discourse. Some structures have also appeared slightly unusual after editing as a change in register (from formal to sensibly less formal).

The general distributional pattern of the texts was as follows:

The first member of a pair /triplet was the original version of the text. The other(s) was/were the edited version(s); Thus

Experiments 1 and 2 : Texts with even numbers were 'original'

Experiment 3 : Texts with syntactic markers were 'original'

Experiment 4 : Texts with odd numbers were 'original'.

As a result of text editing, it was felt necessary to draw a taxonomy of correct and near-correct responses for both 'original' and 'edited' versions (See Appendix III).

The second consequence of text editing was apparent in experiments 1 and 2 mainly. The shifting of clue(s) from left of gap to right of gap and the insertion of a distracting sentence created a certain imbalance between the different parts of the textual material.

3.7 The results

This section is an account of the results of the four experiments using cloze procedure described in the previous sections. It provides information on the effect of the four experimental variables on inferencing plausible fillers for the gaps. An analysis of the 'main effects', that is, the effect of the three independent variables separately compares the mean scores obtained for each of the IV and precedes a discussion of the results of the interactions between the different independent variables. The data was analysed with a 3-way Analysis of Variance, with repeated measures of one of the three factors, yeartime, language and condition . (The specific procedure used for this ANOVA was that of the SPSS package of statistical programmes (Statistical Package for the Social Sciences) using Sheffield University Prime 750 computer.⁵ See Appendix VI for statistical results.) The experimental data discussed below attempts to answer three research questions raised at the beginning of this chapter:

- 1 How do FL learners utilise lexical resources/links of cohesion and coherence when inferencing unknown meaning while reading?
- 2 How does the use of lexical resources/links of cohesion and coherence vary as a function of FL proficiency?
- 3 How does the use of lexical resources/links of cohesion and coherence vary as a function of language background?

3.7.1 Experiment 1: 'Order of clue'

The null-hypothesis formulated for this experiment was:

"There is no effect of the order, immediately before or immediately after, in which a crucial clue comes relative to a gap on the success of learners' inferencing a plausible filler for the gap".

3.7.1.1 Comparing scores for the IV "Experimental condition"

This independent variable obtained highest scores overall in condition 2, that is, when the clue was after the gap (mean score m.s. = 7.986) in contradistinction to condition 1 (clue before the gap which obtained m.s. = 6.986. These results do not go along with the expectation of highest scores with condition 1. However, the amount of difference between the two overall scores is one mark on the marking scale adopted in this study (condition 1 and condition 2 were each marked out of 15) and cannot be regarded as important.

The statistical result of the difference in scores between experimental conditions was shown by the ANOVA to be significant at $p = .05$ level of significance: 'CONDITIO' $p = .006$ (99.4%). Although this difference is significant, it is not sufficiently big to be of practical importance.

Why were scores highest in condition 2? The reasons may be related to certain linguistic factors inherent to the design of condition 2. It seems that recovering an item

whose meaning is being repeated in the clueclause/sentence following the gap/ gapclause was comparatively easier because a clue item occurring a posteriori to the gap elaborates on the missing meaning more effectively, benefiting from some discourse markers and punctuation (such as, colons, semi-colons, dashes and commas) which signal "communicative acts" (in Widdowson's sense) of explanation, expansion or confirmation, and thus contribute to the clarity of the missing item.

Shaughnessy (1977: 16-17) defines punctuation as a "code that serves to signal structural, semantic and rhetorical meanings that would otherwise be missed by the reader". For example:

The National Park Committee reported this year another _____; a man drowned while swimming in the underground waters of a cave in Derbyshire (...) (T2).

(...) The soldiers' _____ clothing was inadequate: there was too great a proportion of nylon in their army socks (...) (T6).

(...) These events will endanger the prospects for further growth in a _____ which had not been much on the news before - north west India. (T20).

However, it seems that the general shortness of the texts may have considerably contributed to the feeling of relative easiness experienced by the testees when the clue was after the gap. In real (indefinitely long) texts, the clues would be further ahead of the gap and possibly more

difficult to spot.

One question arises from these results: were the high scores under condition 2 influenced by the year, or the language variable or both? The interactions (see Sections 3.7.1.4 ff below) will bring light on this question.

3.7.1.2 Comparing scores for the 'yeartime' independent variable

Highest scores overall were obtained by year 2 subjects (8.146) as compared to year 1 (6.416) and year 3 (7.896). The difference in scores is important between year 1 and year 2 ($D = 1.73$) and between year 1 and year 3 ($D = 1.48$). These results explain the significant difference between years at $p = .05$ level of significance, showed by the ANOVA:

'YEARTIM' = $p = .016$ (98%)

This implies that the year variable had an effect on the overall level of score in spite of the different time allowances. The middle year was the highest scoring group overall.

This result was unexpected but can be explained by the nature of the curriculum currently taught. There is great emphasis on the 'technical' aspect of language in semesters⁶ one and two (first year) throughout semesters three and four (second year). These semesters are devoted to an exhaustive teaching of the 'tools' for text analysis and

students receive considerable feedback directly in the 'guided' reading class and indirectly in the English and American literature class (see note 6 for details of the course). Besides, the use of cloze exercises to test the comprehension of passages is comparatively higher in year 2 than year 1, although cloze exercises remain virtually non-existent during the three years. Reading comprehension tests usually involves a series of questions about a passage to be answered out in full words which implies that productive language interferes with receptive ability and often, diverts the purpose of the test. Students may also be asked to "summarise" the passage "to show that they have understood it" or to "explain the meaning of some words in their own words", which again focusses on productive rather than receptive skills. Thus, year 2 students have only little more practice with cloze exercises, and year 3 students spend even less time doing cloze exercises. This may be explained by the fact that the cloze technique as an index of reading comprehension has been shunned by FL instructors on the grounds that it allowed too many chance factors to interfere with the correct response (for instance, when supplied with three or four choice items - including vocabulary items, phrases or clauses, the student may opt for the correct choice item only by chance) and some instructors simply discarded it as an inadequate measure of reading comprehension.

The emphasis on the receptive skill through intensive reading (in the reading class) and extensive reading (in the literature class) tends to decrease in year 3 and focus on the productive skill essentially, with written assignments submitted by the students at regular intervals becoming an important feature of the syllabus.⁷

But the contents of the final year reading syllabus is somehow less analytic and text-bound, and at times, the reading class tends to be an oral discussion of ideas that emerged from a passage followed by an assessment of these ideas in written form, rather than a close analysis of the text in terms of its linguistic resources. As a result, final year students tend to become less 'language' oriented and 'text' bound, and in connection with the results of experiment 1, their scores reflect this tendency though not in a very significant fashion.

One puzzling issue was that of the bigger difference between first and second year students' scores in comparison to second and third year students' scores. Given that the syllabus for year 1 and year 2 reading classes contains graded material based on the same theoretical principles of developing an awareness of the linguistic resources of the English language for comprehension purposes, the larger difference in scores between year 1 and year 2 subjects may seem quite inexplicable. However, one of the reasons for such discrepancies may be related to the timing of the experiment itself. The experiment took place in the middle of the first semester and it is possible that the

programmation of the experiment(s) later in the first year would have been more productive as the differences in performance between the years would have been more meaningful.

3.7.1.3 Comparing scores for the 'language dominance' independent variable

The F group students obtained sensibly highest scores overall (7.861) as against the A group (7.1). The difference between these two scores ($D = 0.761$) is interesting and suggestive although not statistically significant on the ANOVA: 'AF' - $p = .148$ NS.

This result of highest scores for the F group was somewhat predictable on the grounds that a European background knowledge was expected to facilitate the performance on the experiment overall. Thus, the English language and the French language, both SVO languages and being similar in various respects - graphic, structural, lexical, as opposed to the Arabic language, this may result in 'strategies of text attack' that readers are likely to use when confronted with languages that share certain physical properties, as English and French do. Presumably Arabic would require strategies different from those needed to 'attack' or handle an English (or French) text, which may justify the lowest scores obtained by the A population on the overall performance of the experiment. However, a bigger difference between the scores of the A

group and the F group was expected, which raises the following question: was the small amount of difference due to the joint effect of the yeartime variable with the experimental variable? (An examination of the interactions at $p = .05$ and $p = .01$ level in the section below will enable us to answer this question) or is it a sign that the discriminative role of the language background in inferencing unknown meaning had simply been overstated? Section 3.7.1.5. attempts to illuminate this issue.

3.7.1.4 Interaction between 'yeartime' and 'language'
(YEARTIM BY AF)

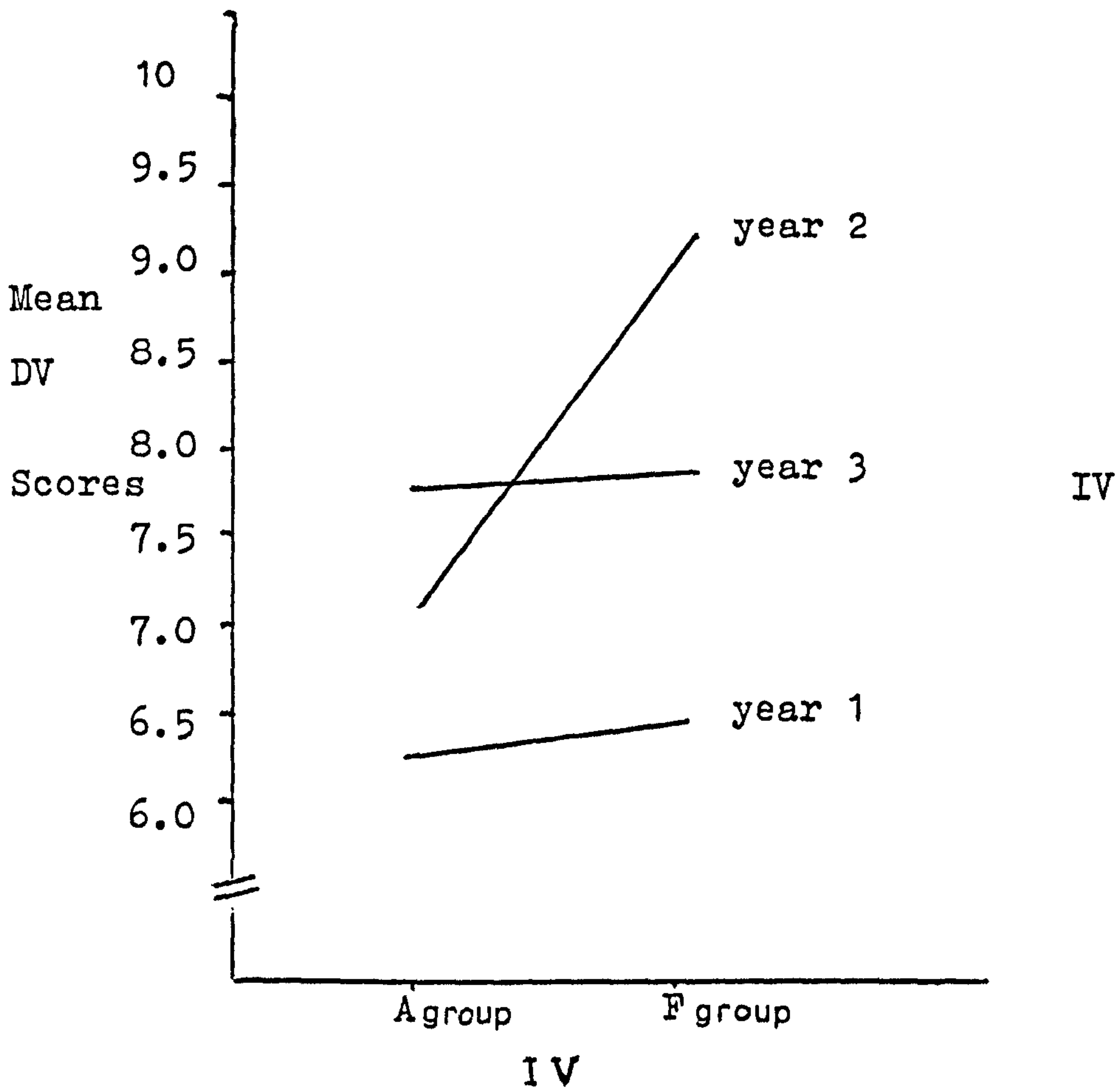
Was the non-significant difference between overall scores of the two language dominant groups much the same in the different years or did the difference of language dominance have a markedly greater effect in some years than others? This difference can be assessed in Table 02 below and visually in Graph 01 below.

Table 02

Mean scores obtained by Y1, Y2, Y3, in A and F					
Y1	A = 6.3	Y2	A = 7.125	Y3	A = 7.875
	F = 6.5		F = 9.16		F = 7.916
	D* = 0.2		D = 2.035		D = 0.041

*D = Difference

Graph 01



Comments

The difference between A scores overall of the two language groups is much the same in year 1 and year 3, but it is markedly greater in year 2 ($D = 2.035$).

Graph 01 shows no important interaction between year and language: the lines of year 1 and year 3 are parallel.

This interaction was shown by the ANOVA to be non-significant:

YEARTIM BY AF: $p = .211$ NS

This means that the year variable combined with the language variable did not affect the overall scores in a

significant fashion. Looking at the overall scores for each year, F students obtained highest scores in year 2, possibly for two reasons. They were considerably better in 'reading work' than year 1 subjects and also better than year 3 subjects, and as a group whose background language was French, this seems to have given them a supplementary advantage over A students. A students, on the other hand, showed steady improvement in their language ability from year 1 through to year 2 and year 3 but it would appear that year 2's superior training in 'reading' (as opposed to year 1 A subjects) was not sufficient to produce equal performance with year 2 F subjects.

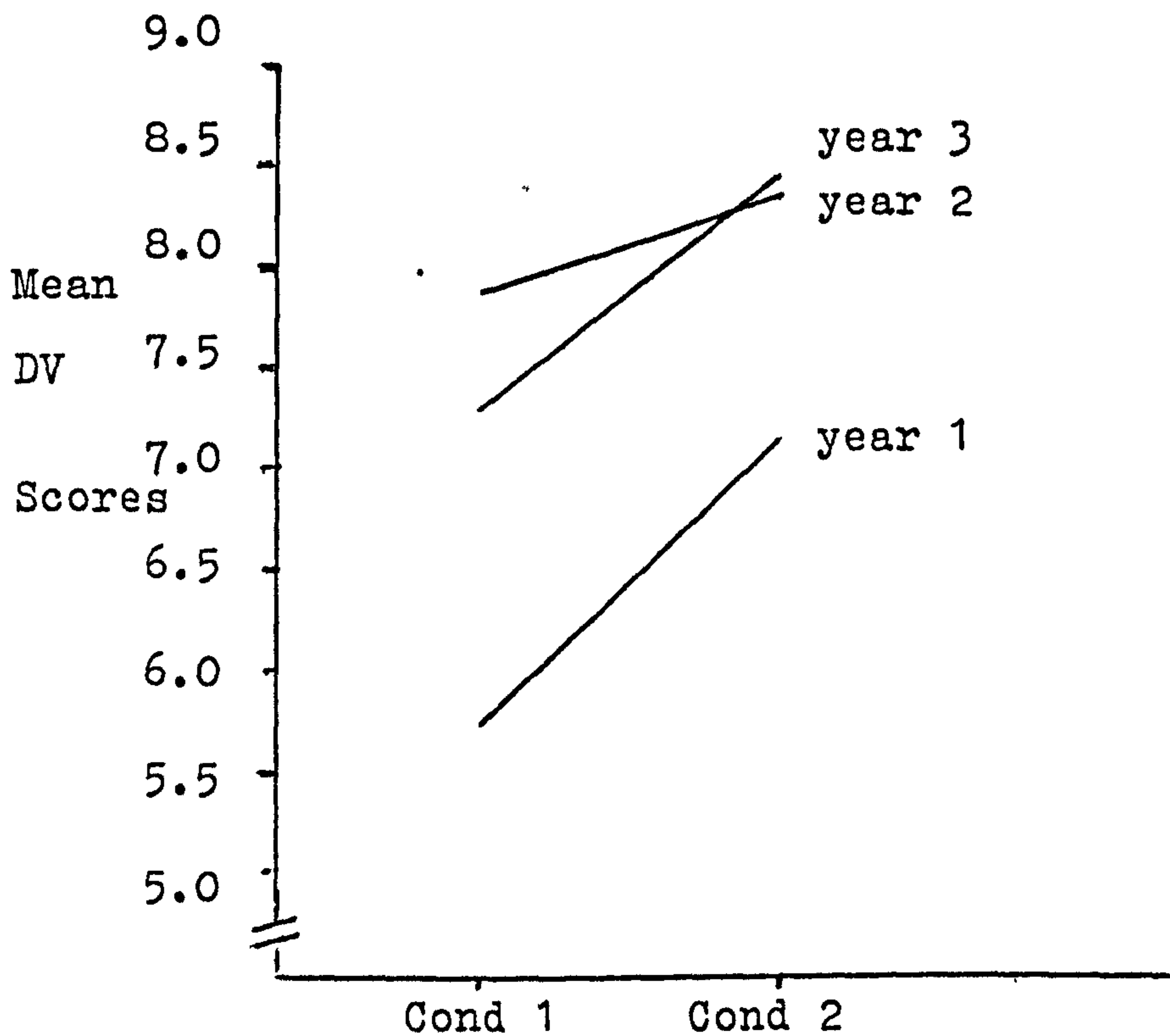
3.7.1.5 Interaction between 'yeartime' and 'condition'
(YEARTIM BY CONDITIO)

Was the significant difference in performance in the two conditions much the same in all years or not? Compare the results in Table 03 below and visually in Graph 02.

Table 03

Mean scores obtained by Y1, Y2, Y3, in C1 and C2					
Y1	C1=5.708	Y2	C1 = 7.916	Y3	C1 = 7.3
	C2=7.125		C2 = 8.375		C2 = 8.458
	D = 1.417		D = 0.459		D = 1.158

Graph 02



Comments

On the graph the difference in scores between conditions for all years is large. It was statistically significant in the ANOVA ($p = .006$). Year 3's performance in condition 2 is the highest. These subjects had no difficulty inferencing meaning when the clue was after the gap. Year 2 performance is also relatively higher and is close to year 3 performance, which of course implies that the highest scores obtained in this condition were mostly due to these two years. There is small to no interaction between year and condition - the lines on the graph are nearly parallel. This is confirmed by the non-significant ANOVA result for this interaction:

'YEARTIM BY CONDITIO': p = .530 NS

This result implies that the year variable had no significant joint effect with the experimental conditions on the performance of the DV overall although the differences between conditions and yeartimes were significant.

3.7.1.6 Interaction between 'language' and 'condition':

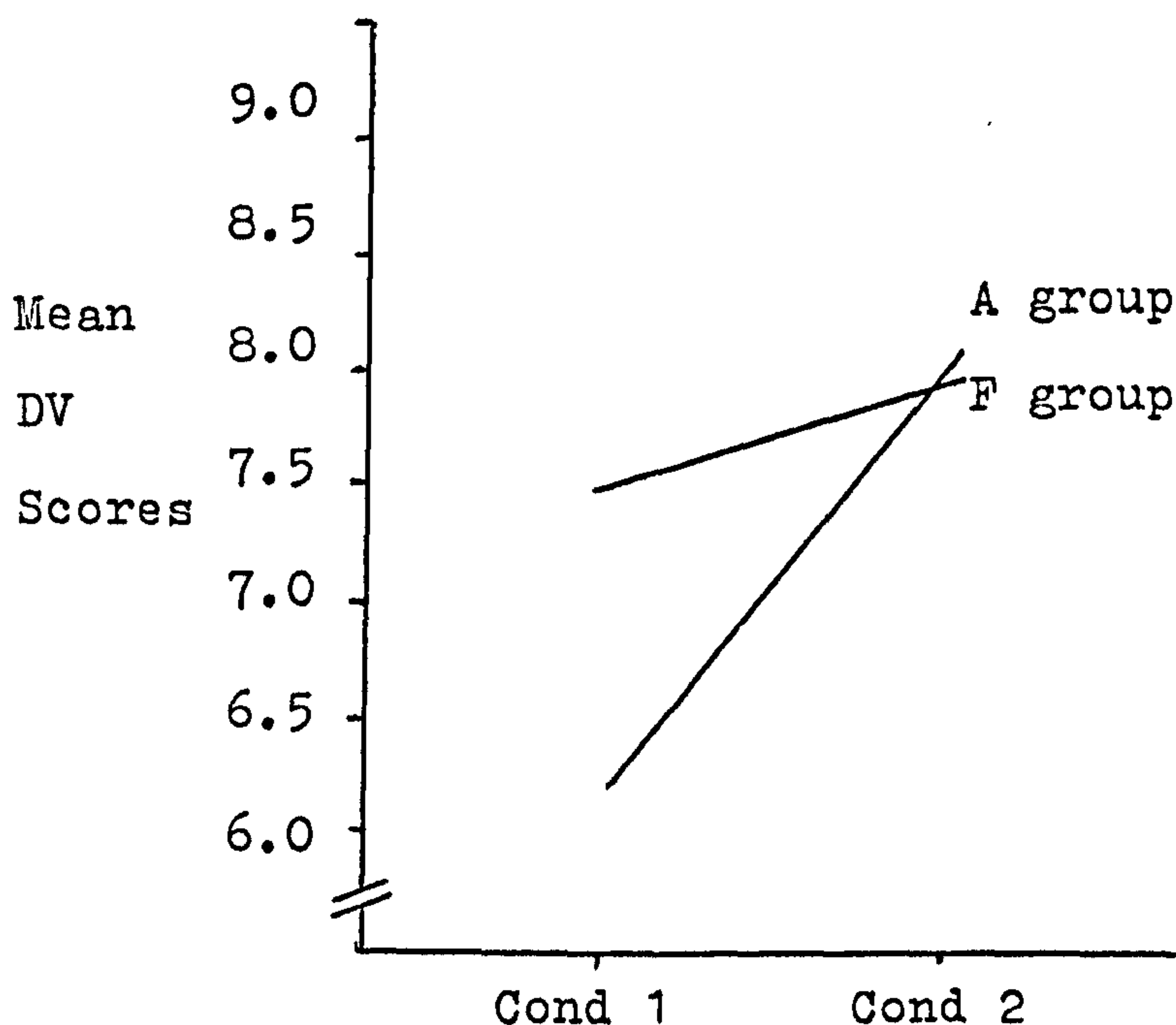
'AF BY CONDITIO'

Examining this interaction enables us to answer the question: was the significant difference in performance in the two conditions different for different language dominant groups? See Table 04 and Graph 03 for a visual comparison of these two variables.

Table 04

Mean scores obtained by A and F in C1 and C2			
A	C1 = 6.2	F	C1 = 7.75
	C2 = 8		C2 = 7.972
	D = 1.8		D = 0.222

Graph 03



Comments

The difference between scores in the two conditions is large in the A group only ($D = 1.8$). This implies that the experimental variable has affected one language group's performance on the DV much more than the other. Looking at the scores in each condition, A students' performance was highest in condition 2. The interaction between the language and the experimental condition was significant at .01 level of significance:

AF BY CONDITIO: $p = .032$ (96.8%)

which implies that the language factor had a significant effect on the effect of the experimental variable. This

significant interaction brings light on the non-significance of the interaction between year and condition and implies that success in the performance of condition 2 especially was dependent upon the background language of the subject rather than his year of course. The greater difference in performance between the two conditions of A students may be due to the fact that inferencing meaning in condition 2 ('clue after gap') required attention to be shifted to the right end of the text, a reading and inferencing strategy which seemed more natural for and accepted by Arabic dominant readers than French dominant readers on account of the reading and writing system to which they were accustomed. This finding goes against the prediction that knowledge of a European language will be facilitative in inferencing meaning in another European language as non-European language A students obtained highest scores overall.

Discussion

What are the implications for language teaching that emerge from this experiment? Given that the differences in performance between conditions overall were significant but not large (only one mark on the marking scale) one may be led to conclude that these differences were not important and that, after all, they do not reflect any particular problem concerning the subjects' awareness and use of lexical relationships of cohesion and coherence in text/discourse. However, there are three findings in this experiment which do not enable us to draw such

conclusions. First, the difference in performance on the experimental conditions across the years reveals that the identification of lexical links of cohesion was indeed problematic for first year students (see Graph 02). Given that the cloze passages utilised in this experiment to test the order of clue hypothesis contained a large amount of lexical links of linguistic, semantic cohesion (based on synonymic, hyponymic and cosemic relations of sense) of a lexicosubstitutional type and in view of the poor performance of first year students in exploiting these devices of cohesion in inferencing, it may be suggested that their teaching could improve students' ability to recognise these ties in text comprehension. For instance, the teacher should encourage the learner to spot links between lexical items involving verbatim repetition, synonymy, hyponymy or cosemy in order to detect a clue(s) in a passage⁸ and in this connection to search for links backwardly as well as forwardly (see some pedagogical suggestions in the next chapter). This leads us to the second finding about the high performance of the subjects overall when the clue was posterior to the gap. An item analysis has revealed that almost all texts contained lexical cohesive ties involving lexicosubstitutional cohesion (and very few texts contained lexicoreferential cohesion) and the discourse function underlying those lexical ties were for the most part an "explanation" of the missing item. In the case of condition 2 (when the clue was after the gap), punctuation signals as well as syntactic markers of discourse were used that

were not utilised in some texts in condition 1 (clue before gap) and as noted earlier (cf. Section 7.1.1) this may have acted as a stimulant for forward clue searching.

For instance:

T2 The National Park Committee reported this year another _____; a man drowned while swimming in the underground waters of a cave in Derbyshire (...).

T6 (...) The soldiers' _____ was inadequate: there was too great a proportion of nylon in their army socks (...).

T20 (...) These events will endanger the prospects for further growth in a _____ which had not been much on the news before - north west India.

T10 (...) This is a form of _____ which they often discuss for it is used to discriminate against them at work.

T18 (...) The _____ sought to imitate life and the first comic actor was recorded in the city in 211 BC.

Thus, the explicitness of certain texts in condition 2 expressed via semi-colon, colon, dash, conjunct for, may have been responsible for the students' high performance on the 'clue after gap' condition and, with the combined effect of the shortness of the texts, may have yielded the present results. In order to obtain a clearer idea about the effect of a proactive clue on inferencing, this experiment would have to be repeated taking account of the size of the texts. It may be possible that a clue not immediately after the gap would be less likely to be spotted by the reader presumably because of memory

constraints. As a matter of fact, subjects experienced difficulty inferencing meaning for the gap in Text 14 for the reason that Romans, the proactive (clue) candidate for the gap, occurred several clauses after the gap. Thus, certain texts discriminated better between conditions than other texts and the results seem to have reflected this tendency.

The third finding relates to the learners' background language. The fact that A subjects obtained scores in condition 2 as high as those of F subjects in the same condition suggests that some possible further research be done on the influence of L1 reading direction on L3 inferencing.⁹ It would be illuminating to see whether the reading direction of the Arabic language did influence inferencing in English or whether the results were purely coincidental other factors relating to text designs (shortness, easiness) having contributed to these results. Crothers et al (1966) had pointed out that familiarity with the material at perceptual level may facilitate reading (comprehension). The authors found that English-speakers learning Russian reacted more slowly to the Cyrillic characters of Russian than to English letters.

This experiment conducted with Arabic dominant students, who used different scripts, suggests that the L1 may have acted as a perceptual facilitator.

3.7.2 Experiment 2: 'Distance of clue'

A null hypothesis was formulated for this experiment as follows:

"There is no effect of the distance of a clue relative to a gap - whether in the immediately-preceding-the-gap arrangement or in the distantly-preceding-the-gap arrangement on the success of learners inferencing a plausible filler for the gap".

3.7.2.1 Comparing mean scores for 'experimental condition' IV

Highest scores overall were obtained in condition 1, that is, when "the clue was immediately preceding the gap" (9.30) as compared with condition 2 (8.180) when "the clue was distantly preceding the gap". This seems to imply that as expected the nearer the clue to the gap, the more likely it is to inference meaning correctly. The difference in scores between the two conditions was important ($D = 1.12$) and is confirmed by the statistical significance of the ANOVA result:

'CONDITIO': $p = .002$ (99.8%)

The further question is: was the high performance in condition 1 due to the contributory effect of the year or of the language variable, or both? (See the interactions in 3.7.2.4, 3.7.2.5 and 3.7.2.6 below.)

3.7.2.2 Comparing scores for the 'yeartime' variable alone

Highest scores overall were obtained by year 2 subjects (9.312) in comparison to year 1 (7.77) and year 3 subjects (9.145). The difference in mean scores between the years is large in two years, (Difference in scores between:

Y1 and Y2 = 1.542

Y1 and Y3 = 1.375

Y2 and Y3 = 0.167)

and is confirmed by the significant ANOVA result:

'YEARTIM': $p = .004$ (99.6%)

The year variable has affected the overall scores in both conditions in spite of the different time allowances. However, year 2 highest performance was somewhat expected in view of the previous results in a separate experiment (see experiment 1). The same pedagogical reasons may be invoked, viz. considerable practice in reading comprehension exercises as evidenced by year 2 syllabus.

Was the significance of these results due to the experimental variable or to the language variable or both? See the interactions.

3.7.2.3 Comparing the scores for the 'language dominance' IV alone

Highest scores were obtained by A group students (that is, 8.791) in comparison to F group students who obtained 8.694. However, the difference in mean scores between these two groups is very small ($D = 0.097$). Both groups had almost equal performance overall. This confirms the highly non-significant ANOVA result obtained for this difference:

'AF': $p = .806$ NS

Unlike the yeartime variable, the language variable had no effect on the performance of the experimental variable overall. This seems to imply that success in inferencing a filler for the gap was not dependent upon the subject's background knowledge but on their number of years of study and on the content of the syllabus currently taught. This result tends to show the positive consequences of intensive training in the 'mechanics' of text analysis, and to play down the effect of the background language.

3.7.2.4 Interaction between 'year' and 'language':

'YEARTIM BY AF'

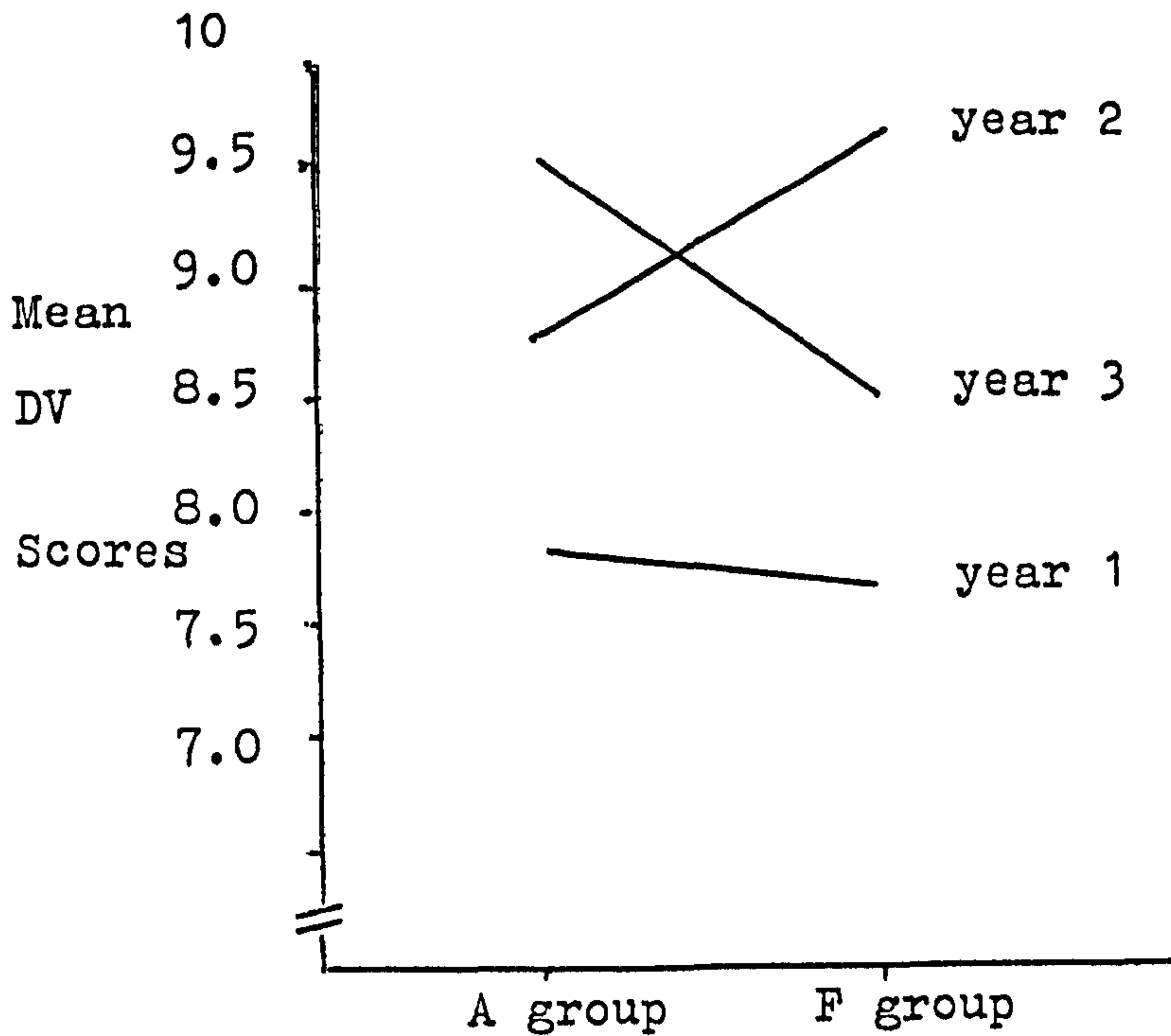
Was the non-significant difference between overall scores of the two language groups much the same in the different years or did the difference of language dominance have a markedly greater effect in some years than others?

These questions may be answered visually by the table and graph below.

Table 05

Mean scores obtained by Y1, Y2, Y3, in A and F					
Y1	A = 7.833	Y2	A = 8.875	Y3	A = 9.666
	F = 7.708		F = 9.749		F = 8.625

Graph 04



Comments

This graph enables a visual appreciation of the difference between mean scores of the two language groups. This difference is small in year 1 but relatively important in year 2 and year 3. Difference in mean scores between A and F:

Year 1 = 0.125

Year 2 = 0.874

Year 3 = 1.041

Note that this pattern is similar to that of Graph 01 above in two recurrent features: year 2 has steeper rising than the others and year 1 is far below year 2 and year 3. Given their previous effect in a separate experiment, it may be predicted that these two years will have the same behaviour in all four experiments. But

unlike Graph 01 pattern (describing the interaction between year and language in connection with the experimental variable 'order of clue') year 1 performance is higher in the A group than in the F group in the present experiment, which may be due to the experimental variable itself. The small interaction specially due to year 2 performance was found to be non-significant by the ANOVA.

YEARTIM BY AF: $p = .149$ NS

which implies that the yeartime variable had no significant joint effect with the language variable on the performance of the DV. But taken separately, as shown earlier, the effect of the year on the DV was significant.

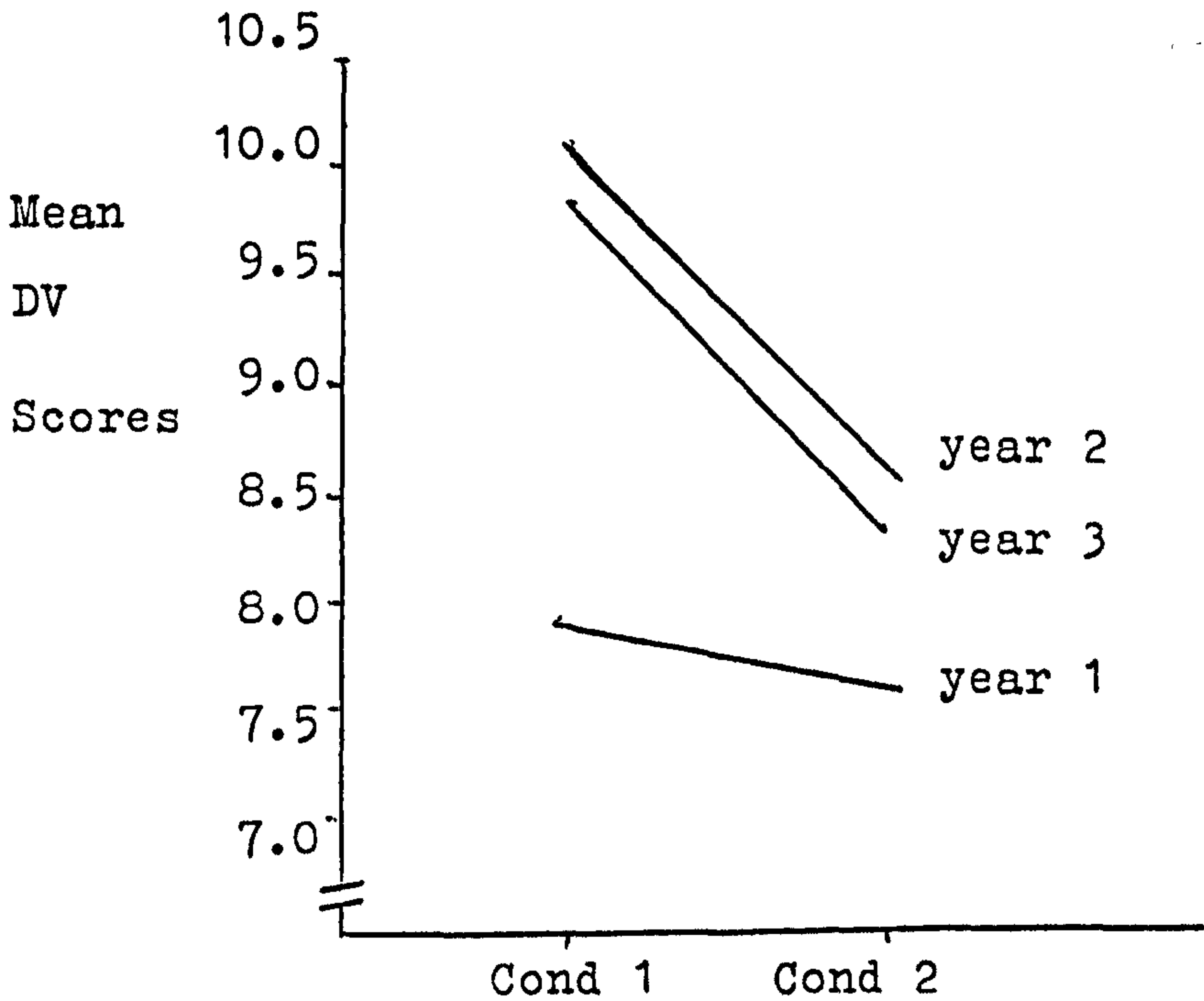
3.7.2.5 Interaction between yeartime and condition
'YEARTIM BY CONDITIO'

Analysing this interaction enables us to answer the question: was the significant difference in performance in the two conditions much the same in all years or not? See the table and graph below.

Table 06

Mean scores obtained by Y1, Y2, Y3, in C1 and C2					
Y1	C1 = 7.875	Y2	C1 = 10.166	Y3	C1 = 9.875
	C2 = 7.666		C2 = 8.666		C2 = 8.416

Graph 05



Comments

The difference in mean scores between the two conditions is considerable in two years. Difference between C1 and C2:

$$Y1 = 0.209$$

$$Y2 = 1.5$$

$$Y3 = 1.459$$

but very small in year 1. Year 2 subjects performed better in both conditions. This is now predictable. There is no interaction between the years' overall performance as the graph shows, and the ANOVA reported a non-significant result for this interaction:

$$\text{YEARTIM BY CONDITIO: } p = .183 \text{ NS}$$

This seems to imply that the year variable had no

significant joint effect with the experimental conditions on the DV. These results are an indication that condition 2 was most difficult to perform whatever the year of the subjects (that is, the further the clue occurred relative to the gap, the more unlikely it was for the subjects to inference acceptable meaning). This, thus, suggests a 'local' inferential strategy (Cohen et al, 1979). 'Subjects' tendency was to search for clues in the vicinity of the gap and overlook those situated further ahead from the gap. It seems that the distance in words (between 5 and 14 words) in the 'neutral' clause was a powerful factor in the clue-searching/inferential process as lexical cohesion seemed less obvious beyond this point. Indeed it appears to have been invisible or non-existent to the testees. It is interesting to note that most lexical semantic ties were substitutional or referential (they included verbatim repetition and co-semy. A few ties only were lexical pragmatic) but it appears that these lexical devices of cohesion could not successfully operate as clues at a distance from the gap, and implies that a strategy to help their recognition should be given particular attention in language instruction.

3.7.2.6 Interaction between language and condition

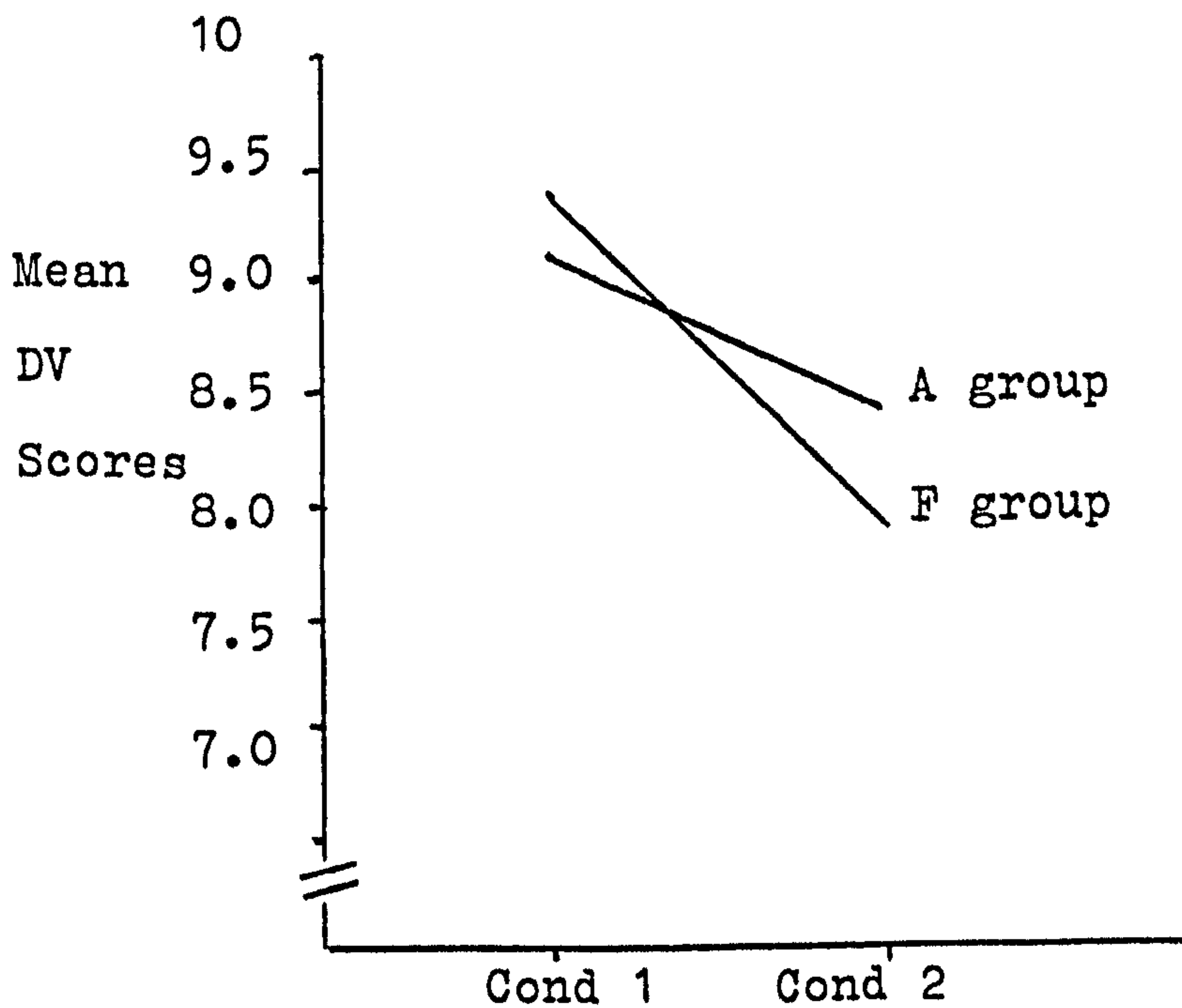
'AF BY CONDITIO'

Was the significant difference in performance in the two conditions different for different language groups or was it much the same? See Table 07 and Graph 06 below.

Table 07

Mean scores obtained by A and F in C1 and C2			
A	C1 = 9.166	F	C1 = 9.444
	C2 = 8.416		C2 = 7.944
	D = 0.750		D = 1.50

Graph 06



Comments and discussion

The difference in mean scores between the two conditions is important in one language group only. Difference between C1 and C2 in A and F:

D = 0.75 in A

D = 1.5 in F

and the interaction between language and condition is small. It was found to be statistically non-significant in the ANOVA:

'AF BY CONDITIO': p = .290 NS

This implies that the language factor had no significant interactive effect on the DV with the experimental conditions overall. Looking at the scores on Graph 06, the A group obtained highest performance in condition 2. So it seems clear that A students were responsible for highest scores in year 2 (see Graph 05 earlier). This result does not go along with the expectation of highest scores by F subjects. Rather it shows once again that the European language factor has not been beneficial to F subjects, performance-wise. Thus, some of the inferential strategies at work in Arabic dominant students have been revealed in this study. A students did not seem to be disturbed by the big distance of the clue relative to the gap and were in general capable of inferencing meaning relatively successfully. This may be related to the rhetorics of the Arabic language which makes greater use of repetition, rewording and restatements as devices to communicate ideas clearly

(see Thompson-Panos and Thomas-Ružić, 1983) and as a rhetorical strategy, a way of persuading (see Koch, 1983b). Also digressions are decoded by the Arabic reader as text-enriching and argumentative devices rather than irrelevant information inserted into the text to distract the writer's purpose and confuse the reader. As a result, A students in this experiment assumed repetition of the same item or restatement of the same concept throughout the text and were therefore capable of recovering the missing item independently of the distance of the clue item in relation to the gap. While the triple interaction between year, language and condition was reported to be highly non-significant in the ANOVA:

'YEARTIM BY AF AND CONDITION': $p = .981$ NS

some significant main effects are worthy of attention and may be of practical importance to language teaching. The lowest scores obtained in condition 2 overall suggest that students may be unaware of the pervasiveness of lexical cohesive ties in text, that is, that members of the same tie ('clues') do not occur at a specific distance from the missing item by necessity. Practice in inferencing meaning can involve tracking down clues across clause and sentence boundaries which often means ignoring nearby albeit cohesive information.

3.7.3 Experiment 3: 'Conjunctive clue'

A null hypothesis was formulated for this experiment as follows:

"There is no effect of the logical connection whether overtly signalled by a conjunct (syntactic linker) or a lexical equivalent to a conjunct, or covert (that is, not signalled) relative to a gap on the success of learners inferencing a plausible filler for the gap".

3.7.3.1 Comparing mean scores for the 'experimental condition' IV

Highest mean scores overall were obtained in condition 3, that is, when the conjunctive clue was an overt lexical equivalent to a conjunct (or paralexical linker) m.s. = 5.570, as compared with the other two conditions, viz. condition 1 which involved an overt syntactic linker (m.s. = 5.361) and condition 2 when the conjunctive clue was covert (m.s. = 4.793). But the difference in mean scores between the three conditions overall is not important, as can be seen below:

Difference between C1 and C2 = 0.568

Difference between C1 and C3 = 0.209

Difference between C2 and C3 = 0.777

This difference (under one mark) between conditions was reported in the ANOVA as non-significant:

CONDITIO: p = .121 NS

This non-significant result is somewhat deceptive as a significant difference even small, was also expected between conditions for this experiment. However, the

difference in mean scores between the three conditions does go along with the expectation of highest scores with the two conditions in which the clue was overt, that is, condition 1 - syntactic clue (m.s. = 5.361) and condition 3 - paralexical clue (m.s. = 5.570). Furthermore, the fact that highest scores were obtained in condition 3 at last does justice to the value of paralexical clues in inferencing conjunctive meaning.

3.7.3.2 Comparing mean scores for the yeartime IV alone

Highest scores overall were obtained by year 3 subjects (5.764) in comparison to year 1 subjects (4.321) and year 2 subjects (5.640). This makes a change from the previous two experiments in which year 2 subjects twice obtained highest results overall. The difference in mean scores is important in two years, as can be seen below:

Difference in mean scores between year 1 and year 2	=	1.319
" " " year 1 and year 3	=	1.443
" " " year 2 and year 3	=	0.124

This order of size was confirmed by the statistically significant ANOVA results:

YEARTIM: $p = .002$ (99.8%)

As expected the year factor has affected the overall results. This is an expectation which unfortunately experiments 1 and 2 did not confirm. The increasing order of the mean scores from year 1 through year 3 implies that some serious difficulties encountered during the first year can be overcome by the time the students reach the final year. And the further question concerning

the condition(s) in which year 1 subjects scored lower can be answered by an examination of the interactions in Section 3.7.3.5 below.

3.7.3.3 Comparing scores for the language dominance variable alone

The F group obtained highest scores overall, 5.436, as opposed to the A group who scored 5.047. It can be anticipated that this group were responsible for the highest scores obtained in condition 3 in year 3. However, the difference in scores between the two language groups overall was small ($D = 0.389$) and as shown in the ANOVA, was statistically non-significant:

'AF': $p = .285$ NS

Both language groups have performed almost equally overall. This is an indication that unlike the year variable, the language variable did not affect the dependent variable results in a significant fashion. Highest scores depended upon the subject's number of years of study rather than on his background language.

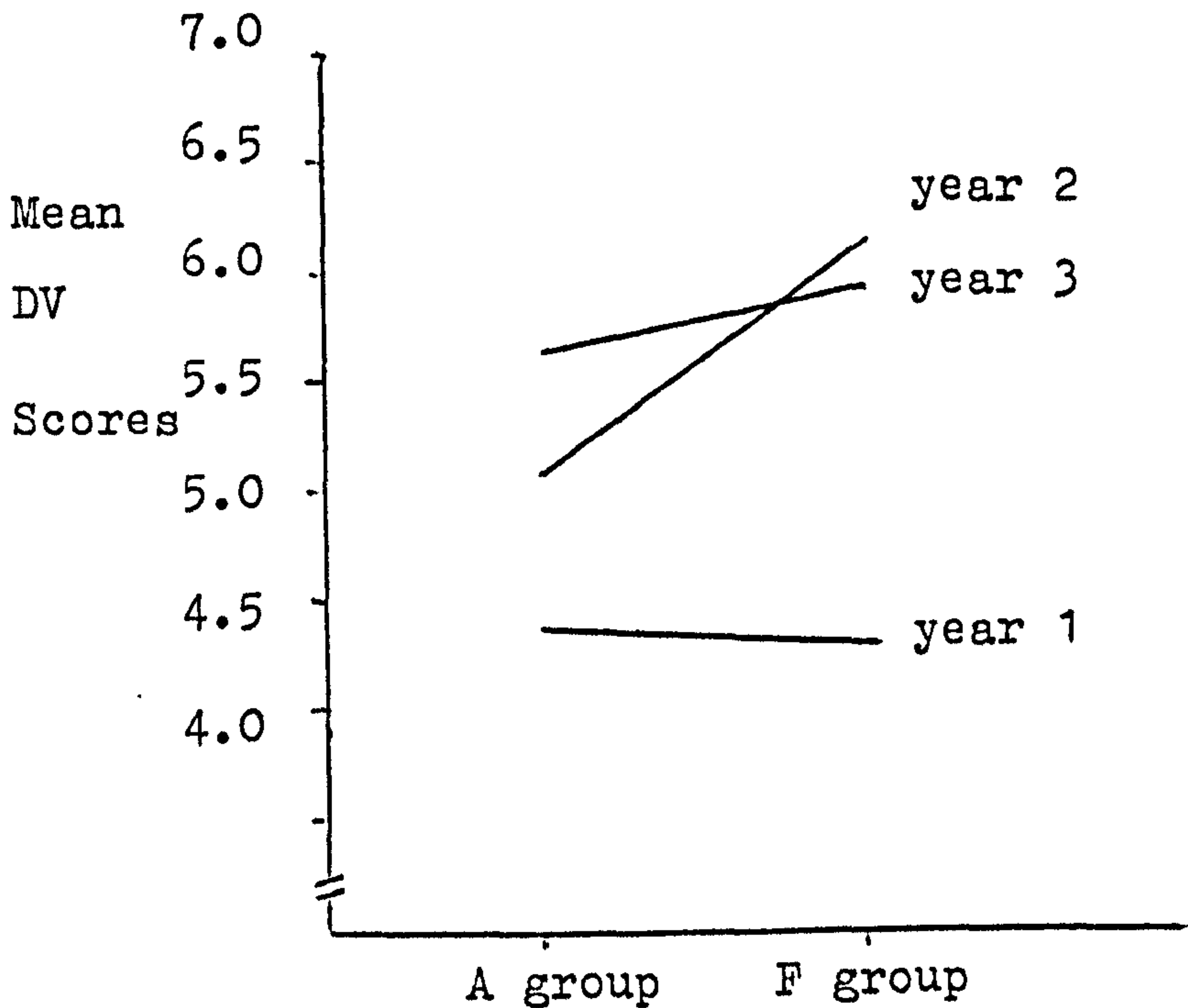
3.7.3.4 Interaction between yeartime and language: 'YEARTIM BY AF'

Was the non-significant difference between scores of the two language dominant groups much the same in the different years or did the non-significant difference of language dominance have a greater effect on some years than others? See Table 08 and Graph 07 below.

Table 08

Mean scores obtained by Y1, Y2, Y3, in A and F					
Y1	A = 4.364	Y2	A = 5.111	Y3	A = 5.667
	F = 4.278		F = 6.129		F = 5.861

Graph 07



Comments

The graph shows that the difference in scores between the two language groups is small in year 1 and year 3 (the lines are parallel) but sizeable in year 2. There is a small interaction between year 2 and year 3 overall scores which confirms the non-significant interaction between year and language reported in the ANOVA:

YEARTIM BY AF: $p = .410$ NS

This implies that the year variable had no significant joint effect with the language variable on the performance of the DV. Note that this graph is equivalent to Graph 01 in experiment 1 and to Graph 04 in experiment 2, and all three graphs showing the non-significant interaction between year and language nevertheless contain the recurrent feature of the steeper rising of year 2 line in direction of the F group. This demonstrates that in year 2 subjects belonging to the F group were better performers than their counterparts in the A group and suggests that the combination of two factors, intensive drilling in 'guided' reading and European language background may have contributed to successful recognition of lexical clues while reading.

3.7.3.5 Interaction between yeartime and condition:

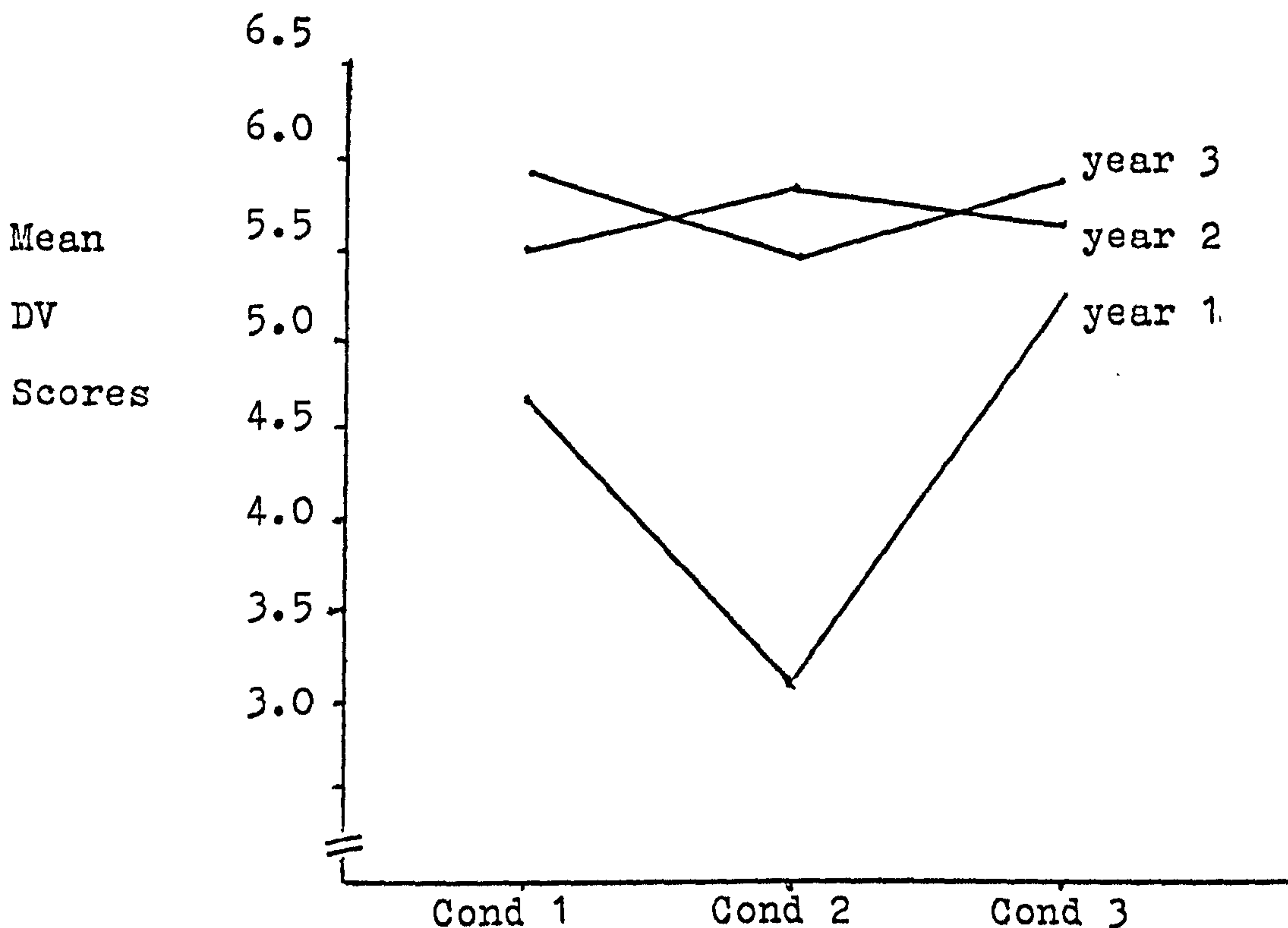
YEARTIM BY CONDITION

Was the non-significant difference in the performance of the three conditions much the same in all years or not? Table 09 and Graph 08 below may give a visual answer to this question.

Table 09

Mean scores obtained in C1, C2, C3, by Y1, Y2, Y3					
Y1	C1 = 4.667	Y2	C1 = 5.50	Y3	C1 = 5.917
	C2 = 3.043		C2 = 5.878		C2 = 5.459
	C3 = 5.253		C3 = 5.542		C3 = 5.917

Graph 08



Comments

The difference in scores between the three conditions is noticeable in year 1 only. It is negligible in year 2 and year 3. There is a small interaction between year 2 and year 3 scores which was reported as near-significant (at $p = .10$ level) in the ANOVA:

YEARTIM BY CONDITIO: $p = .089$ (91.1%)

This implies that the year variable had little effect on the experimental variable and that what effect there was, was mainly due to year 1 performance. However, the yearby condition interaction is important to focus on as it suggests that success in utilising a particular type of clue in inferencing meaning was dependent upon the subjects' overall proficiency which was a factor of the number of years of study and the

syllabus content.

The three years scores in condition 3 reflect familiarity and preference for paralexical clues. It, thus, appears that because the lexical equivalent of the syntactic marker of conjunctive cohesion was basically lexical, with a function similar to a syntactic linker, it was successful as a clue. For instance, inferencing meaning when It may be added was encountered in text was easier than when Furthermore marked the conjunctive meaning of additivity.

The covertness of the clue was problematic for year 1 and year 3 subjects, but not for year 2 subjects. That two year groups out of three found it difficult to inference meaning in the absence of markers of conjunctivity is symptomatic of the fact that the reader's (perhaps unconscious) approach to the text is to look for the explicit signals of logical connection between its different elements, signals which reveal that the writer is being co-operative. Language users like explicitness, a precondition for successful communication, and assume it while reading. Lack of it could mean, even for the native speaker reader, intentional ambiguity on the part of the writer.¹⁰

In this experiment specifically, it may be possible that the reader only expected to use non-signalled lexical clues in lengthier texts than these, and that given the relatively short size of the experimental texts, there was little opportunity for him to utilise this strategy.

Discussion

In view of the results on graph of the yeartime by condition interaction of this experiment and of the interactions of the previous two experiments which demonstrated the falling far below the others of year 1 line, a meaningful pattern appears quite clearly which enables us to conclude that inferencing expertise utilising the lexical semantics and lexical pragmatics of a text is a developmental feature that reaches a plateau in year 2 - 3. This finding has some pedagogical implications with respect to experiment 3 specifically. It seems essential to focus in FL teaching on the potential cohesion of the lexical content of successive sentences and on how to deduce conjunctive meaning from propositional content, whether or not conjunctivity is signalled in text. Researchers on SL and FL learning (Bensoussan and Laufer, 1984) have pointed out that some syntactic markers of cohesion (for example, nevertheless, yet) can be more confusing to non-native speakers than complete absence of signalling. This fact reveals a lack of awareness of semantic and pragmatic lexical clues whose function in the text may illuminate the meaning of certain lexical items.

The following example is an illustration of the way linguistic and pragmatic clues complement each other in the text and can contribute to the interpretation of unknown lexical items (Text 63 of Experiment 3).

The Macedonian capital, Skopje, is fertile, industrialised, with a higher percentage of university students than in almost any other part of the country. Economists regard it as a _____ region. It has unemployment problems.

The cloze item backward is intended to contrast with fertile, industrialised, but its meaning can also be deduced from the lexical content of the proposition expressed by the last sentence via pragmatic knowledge. Indeed, the occurrence of unemployment (problems) has, as its function, to elaborate on the cloze item backward (such elaboration would not exist if the text stopped at region) and the proposition expressed by the last sentence can therefore be viewed as an explanation for preceding discourse.

3.7.3.6 Interaction between language and condition

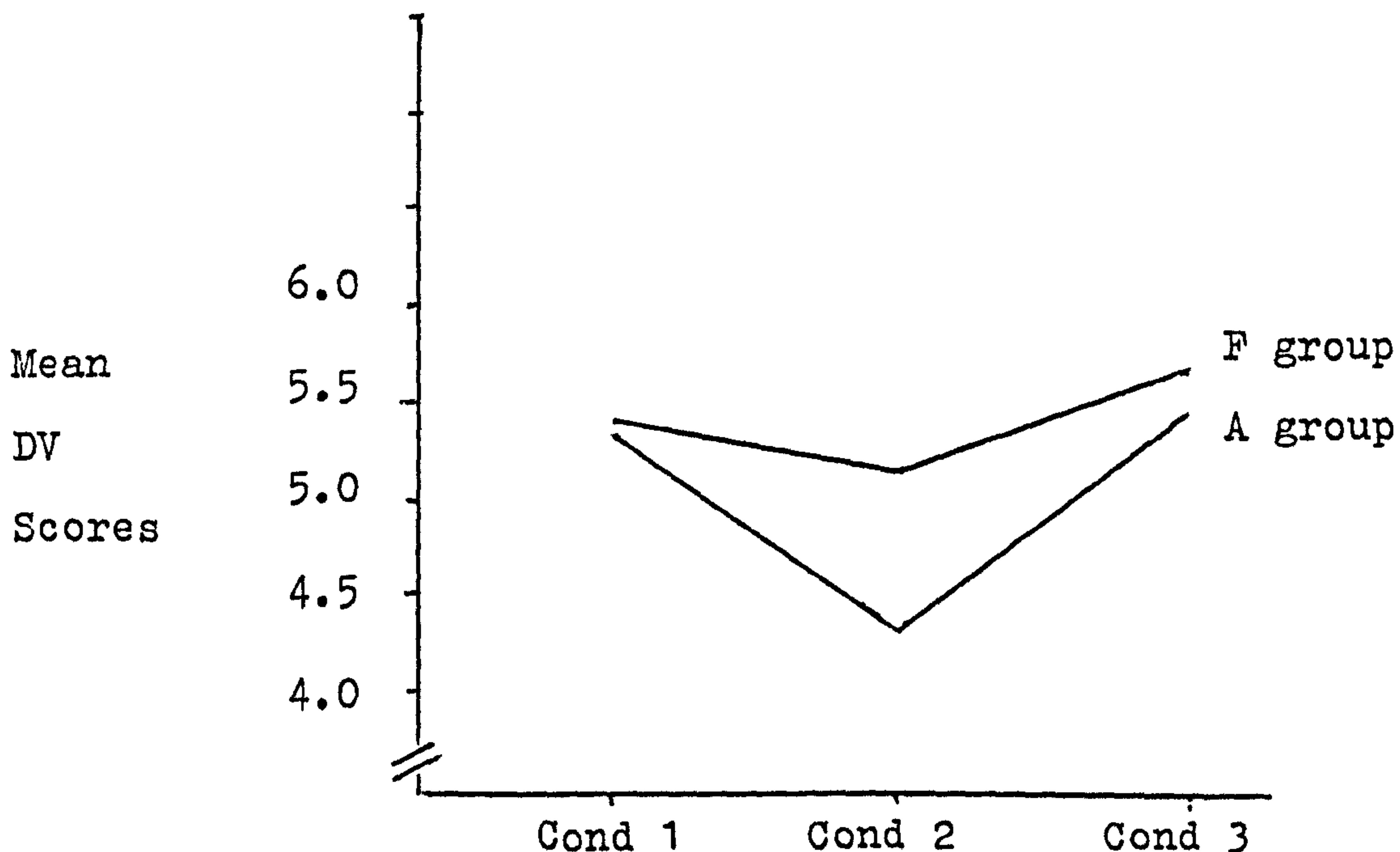
'AF BY CONDITIO'

Was the difference in performance in the three conditions different for different language groups? See Table 010 and Graph 09 below of the mean scores obtained by both language groups in the three conditions.

Table 010

Mean scores obtained by A and F in C1, C2 and C3				
A	C1 = 5.334	F	C1 = 5.389	D = 0.055
	C2 = 4.390		C2 = 5.197	D = 0.807
	C3 = 5.419		C3 = 5.722	D = 0.303

Graph 09



Comments

This graph shows that highest scores were obtained by the F group in the three conditions, but the difference in scores between these two groups in each condition was small, that is, under 1 mark (see Table 010). (Visually the lines on the graph are near together). This difference was found to be statistically non-significant (CONDITION: $p = .121$ NS). Also the interaction between these two independent variables was reported by the ANOVA as non-significant:

AF BY CONDITIO: $p = .618$ NS

This implies that the language variable had no significant combinatory effect with the experimental conditions on the DV scores. However, a closer look at the

'main effects' reveals that low scores were obtained in condition 2 by both language groups and highest scores were obtained by these two groups in condition 3. As noted earlier in Section 3.7.35, the first finding may reflect an absence of awareness of linguistic lexical and pragmatic lexical clues as potential factors of cohesion. On the other hand, the second finding involving preference for lexical signalling of conjunctive cohesion (that is, the presence of paralexical markers in text) over syntactic signalling, reflects an inferencing strategy that is basically lexical and bears implications that the development of this aspect of inferencing should be emphasised in teaching.

3.7.4 Experiment 4: 'linguistic clue'

The following null hypothesis was formulated for this experiment:

"There is no effect of the presence of a linguistic clue or a non-linguistic/pragmatic clue relative to a gap on the success of learners' inferencing a plausible filler for the gap".

3.7.4.1 Comparing scores for the 'experimental condition' IV

Highest scores were obtained in condition 1, that is, when the clue was linguistic: m.s. = 7.374, as against condition 2 which obtained 6.166 when the clue was non-linguistic/pragmatic. The difference in mean scores between the conditions was important:

D = 1.208 and the ANOVA result was found to be statistically significant:

'CONDITIO': p = .023 (97.7%)

These significant results go along with the expectation of highest scores with the linguistic clue condition on account of the explicitness of the clue in this condition. This is an aspect that has often been assumed but no empirical data has been provided in FL acquisition to test this assumption. We have come across the idea of testing the hypothesis that explicit lexical clues are successful predictors of content as a result of a pilot test of reading comprehension in which non-native undergraduates had to supply fillers for gaps in cloze passages. When the text was not sufficiently explicit, subjects had difficulty supplying missing items although much information could be deduced from it.

Was this performance influenced by the year or language independent variables?

3.7.4.2 Comparing scores for the yeartime IV

Highest scores overall were obtained in year 3 (7.270) as against year 1 results (6.062) and year 2 results (6.979). The difference in mean scores overall was important between some years, as the data below shows:

Difference in mean scores between year 1 and year 2 = 0.917
" " " year 1 and year 3 = 1.208
" " " year 2 and year 3 = 0.291

and this is weakly confirmed by the ANOVA near-significant value, at $p = .10$ level of significance:

YEARTIM: $p = .086$ (91.4%)

Thus, the year variable has affected the overall scores in a near-significant fashion. This result was predictable in view of the data obtained from the previous experiments which indicated a significant effect of the year variable on the DV, but the further question is which year has mostly affected the DV? See the interactions.

3.7.4.3 Comparing the scores for the language dominance IV

The F group obtained highest scores on the DV (6.972) in contradistinction to the A group who scored 6.569.

But the difference in mean scores between both language groups is small ($D = 0.403$) and was found to be statistically non-significant in the ANOVA:

'AF': $p = .381$ NS

The two language groups have performed about equally well overall although the F group did sensibly better. Thus, the language variable has not affected the overall scores in a significant fashion and this finding agrees with that of the preceding experiments in which the effect of the language was also found to be non-significant.

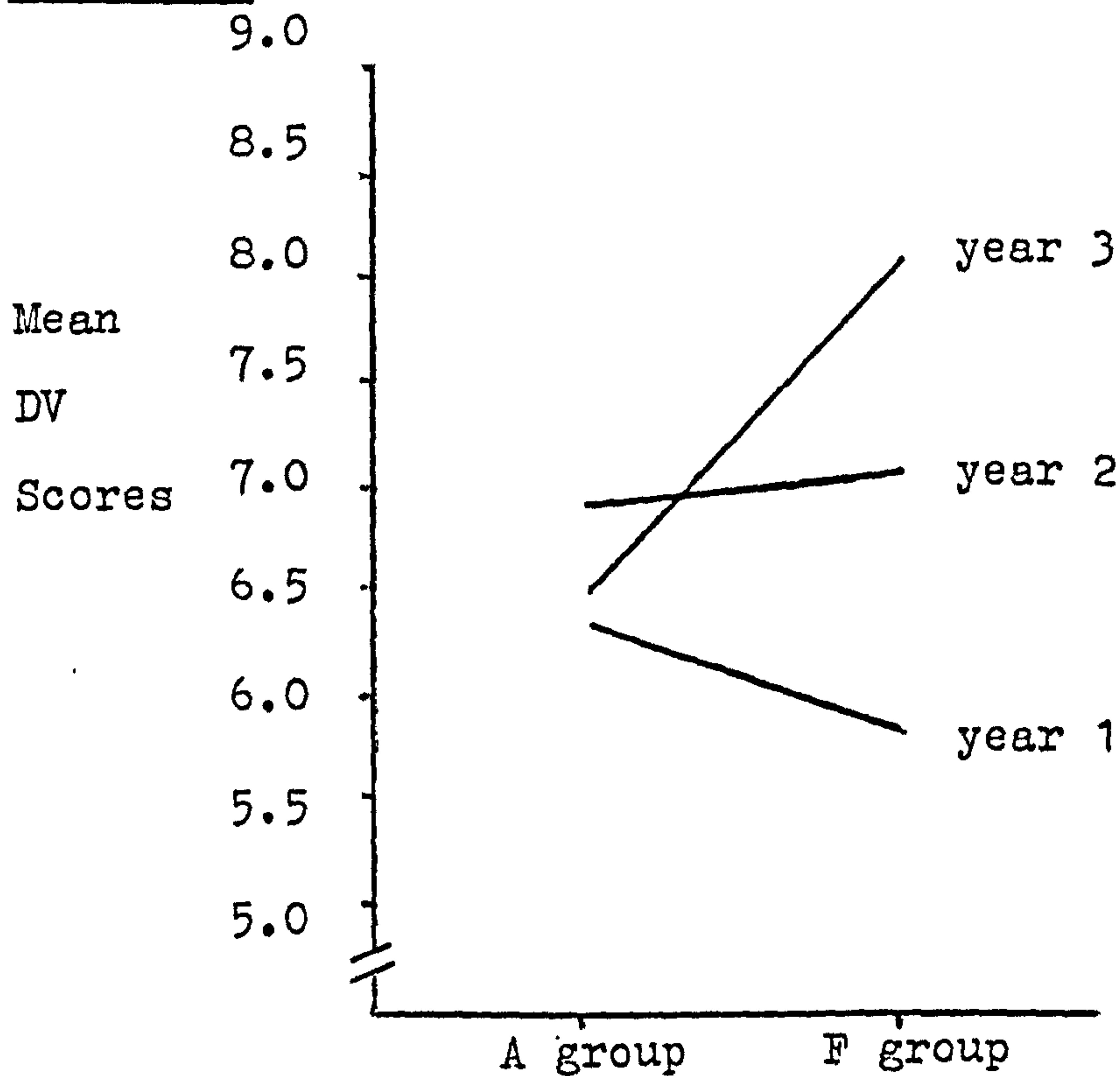
3.7.4.4 Interaction between yeartime and language IV (YEARTIM BY AF)

Was the non-significant difference between scores of the two language groups much the same in the different years or did the difference of language dominance have a greater effect in some years than others? This can be seen on Table O11 and Graph O10 below.

Table 011

Mean scores obtained by Y1, Y2, and Y3 in A and F					
Y1	A = 6.291	Y2	A = 6.916	Y3	A = 6.5
	F = 5.833		F = 7.041		F = 8.041
	D = 0.458		D = 0.125		D = 1.541

Graph 010



Comments

Graph 010 shows that highest scores overall were obtained by the F group in years 2 and 3, but not year 1, but the difference in years' scores between language groups is noticeable in year 3 only, which may result from the experimental variable. We also note that for A group students, the difference in performance between the years is small, in comparison with the difference between years performance by F group students. This seems to imply that instruction had little effect on A's

performance in cloze texts. For the F group, such difference is appreciable. However, the results of the A group do not agree with those obtained in previous experiments which showed a relatively large amount of difference in performance between years in all of them. The cause of this discrepancy in the A group seems to relate to the particular nature of the current experiment which involved recognition of implicit links of 'cohesion' in text and interpretation of implicit links. We also note that in the A group, year 2's performance is highest, a feature not encountered in previous experiments. However, the situation is quite opposed in the F group. Year 3's performance is the highest in this experiment in comparison to the previous experiments where year 2 students have invariably obtained highest scores. This is a feature which may be illuminated by the experimental condition variable itself. However, the yeartime by language interaction reported for this experiment was non-significant in the ANOVA:

YEARTIM BY AF: $p = .192$ NS

which implies that the year variable had no significant joint effect with the language variable on the DV scores. An examination of 'language'-'condition' interaction will show in which experimental condition the performance of the F group was higher.

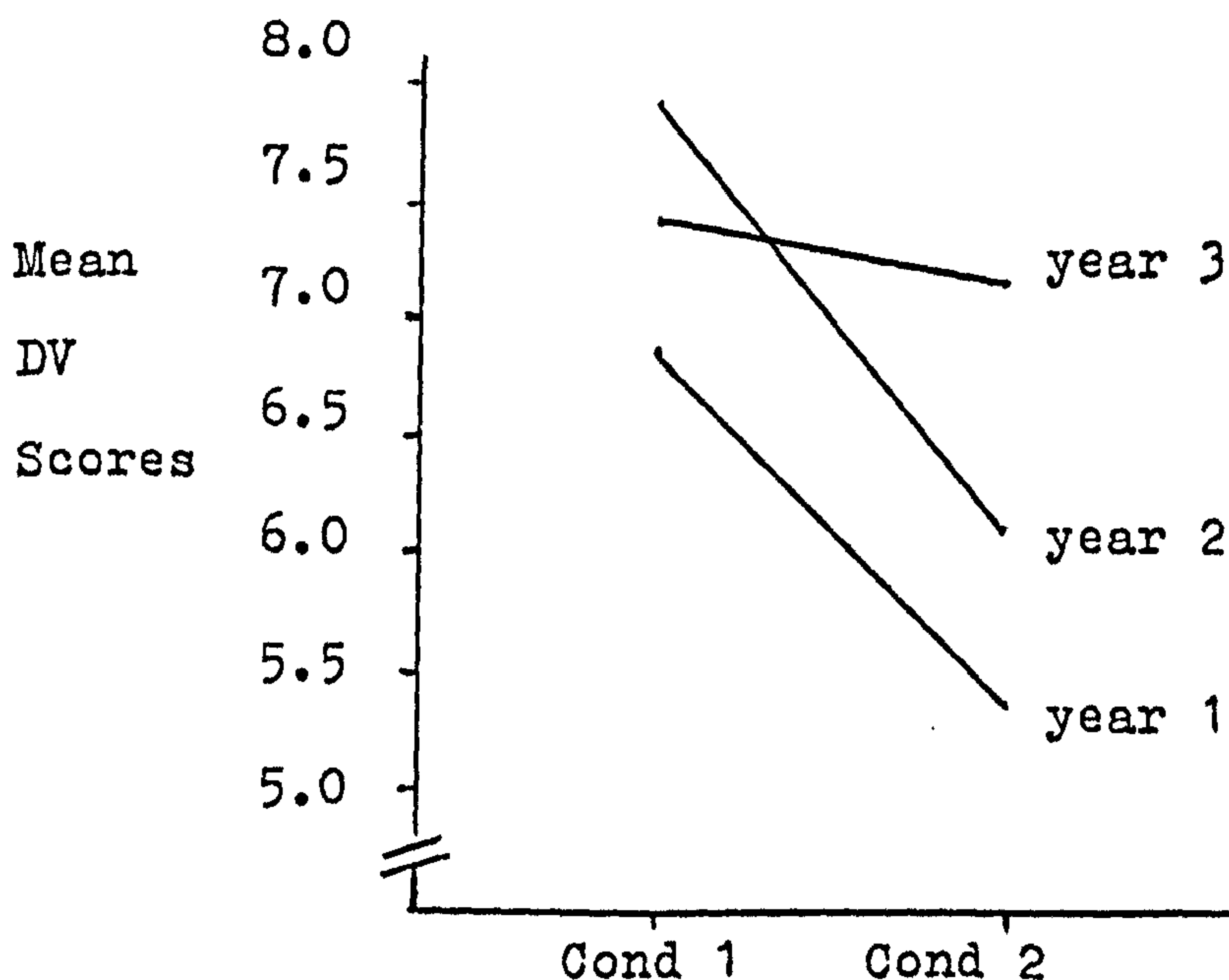
3.7.4.5 Interaction between 'year' and 'condition':
'YEARTIM BY CONDITIO'

Was the difference in performance in the two conditions much the same in all years or not? A study of mean scores numerically and visually will enable us to answer this question.

Table 012

Mean scores obtained by Y1, Y2, Y3, in C1. and C2					
Y1	C1 = 6.833	Y2	C1 = 7.916	Y3	C1 = 7.375
	C2 = 5.291		C2 = 6.041		C2 = 7.166
	D = 1.542		D = 1.875		D = 0.209

Graph 011.



Comments

Graph 011 shows highest performance overall by year 3 subjects in condition 2 only. It shows a large difference between scores in 'experimental conditions'

in year 1 and year 2. The interaction between year and condition was non-significant in the ANOVA:

YEARTIM BY CONDITIO: $p = .390$ NS

which implies that the year variable had no significant joint effect with the experimental variable on the DV.

However, an examination of the 'main effects' enables us to compare these results with those of previous experiments.

Year 2 scores were highest in condition 1. This is now predictable in view of previous results in experiments

1 and 2, although there is still uncertainty

about the background language of these subjects (uncertainty which will be removed by the language-condition

interaction in the next section). Year 2 students enjoy

an intensive course in linguistic "grounding" where

'tools' for text analysis are given particular interest.

But year 3 subjects' performance was higher in condition 2 which involved inferencing pragmatic links of

coherence. It seems possible that because of their

superior experiential maturity in general, they have outperformed year 2 subjects, in this experiment.

Year 1 and year 2 subjects, in comparison, may

not have accumulated background information

that will provide a framework within which to read

(see Carrell, 1982 on this point), but this does not mean

that they are not capable of abstraction. In terms of

reading skills which include inferencing, year 2 (and year 1)

students have not reached year 3 students who can be

described as being at "later developmental stages"

(Flower, 1979).

3.7.4.6 Interaction between 'language' and 'condition'
('AF BY CONDITIO')

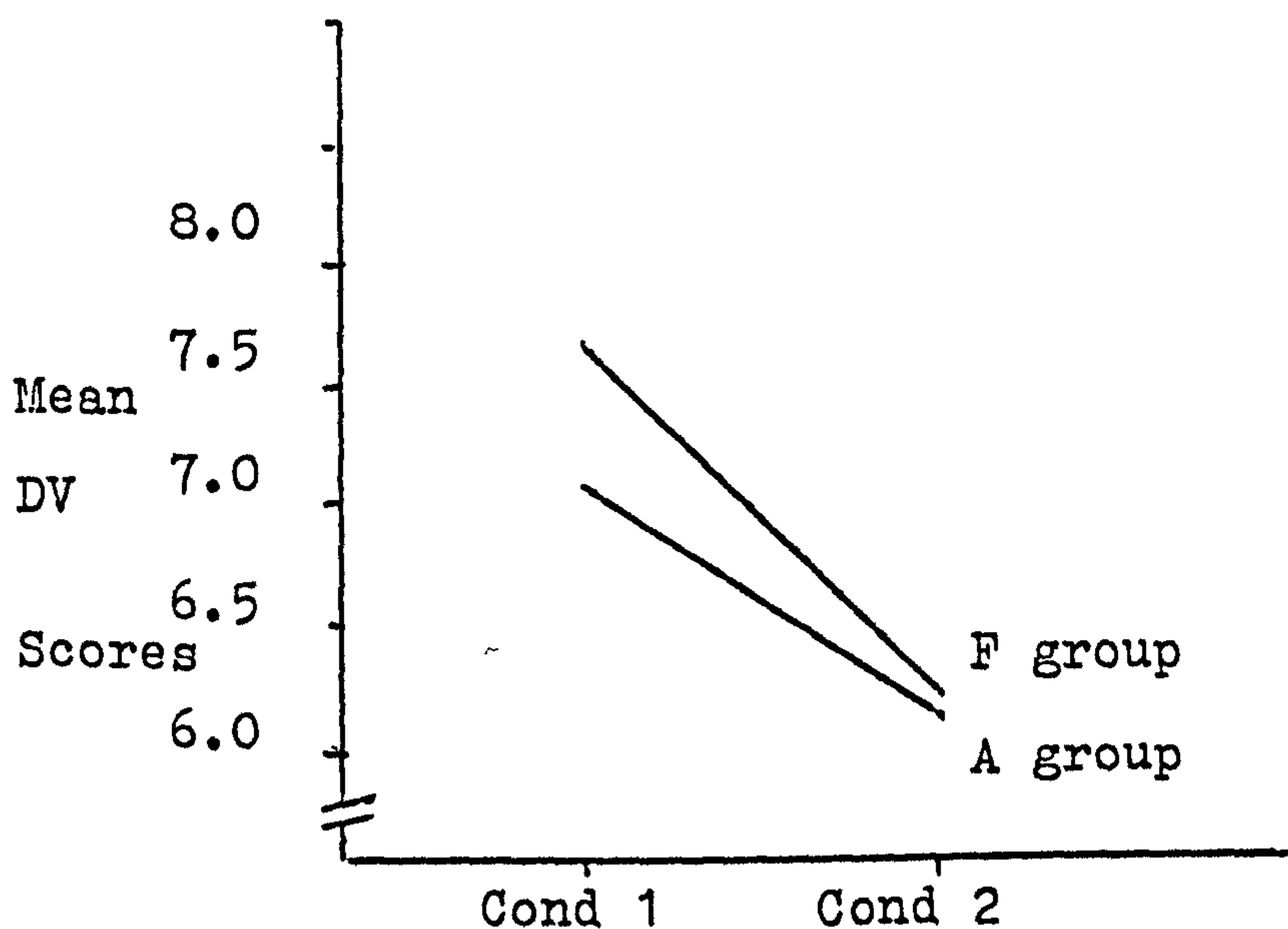
Was the difference in performance in the two conditions different for different language dominant groups or not?

See the table and graph below.

Table 013

Mean scores obtained by A and F in C1 and C2			
A	C1 = 7.027	F	C1 = 7.722
	C2 = 6.111		C2 = 6.222
	D = 0.916		D = 1.5

Graph 012



Comments

This graph shows an appreciable difference between the scores of the F group in the two conditions (D = 1.5) but no interaction between the two IVs 'language' and 'condition'.

The ANOVA reported non-significant statistical interaction:

'AF BY CONDITIO': $p = .57$ NS

which implied that the combinatory effect of the language variable with the experimental variable on the DV was null. Both language groups performed almost equally under both conditions, although F group scores were sensibly higher overall.

General discussion

Despite the non-significance of the results reported by the last two ANOVAs (in this section and in section 3.7.4.5 above), the experimental condition variable has in some ways affected the performance in cloze tests of both language groups in the three years. Condition 2 seemed to have been a complex task to perform for subjects in the three years, whether in the A or the F group, as opposed to condition 1. Subjects were in general capable of identifying linguistic definitional links between lexical elements of text.¹¹ They seemed to have preferred inferencing meaning when the clue was linguistic, which made the propositional content of juxtaposed sentences explicit (for instance, the definitional type of relationship that connects poor to bad shelters). On the other hand, when the degree of cohesion was less explicit, inferencing appeared to be problematic. Widdowson (1978: 26) had remarked that "The difficulty we have in recovering propositional development is a measure of the degree of cohesion exhibited by a particular discourse". Because 'implicit relationships

of lexical cohesion, defined in Chapter 1 as pragmatic relationships of lexical coherence, involve non-definitional types of links between lexical elements of a text (for instance, the link produced by the co-occurrence of poor and no money) they are most difficult to identify generally, perhaps because they require the reader to "co-operate experientially" in the interpretation of the discourse by bringing his own knowledge and beliefs into the text (see Nystrand, 1982 on this point). This is an attitude that some subjects were not willing to have, or that they were simply reluctant to use because of its non-conformity to a test-situation (note the significant difference in scores between experimental conditions). This finding appears to invalidate the argument put forward by Aronowitz (1984) that test-takers are more concerned with the truth-value of their response than with being correct within the linguistic context of the text. Aronowitz (1984) provided some evidence about young native speakers of English answering cloze tests in their native language. They showed a tendency to use a "contextualising strategy" to answer questions and little testwiseness. In a test situation, the "rules of the game" are not to use one's world knowledge to supply answers but to use textual information. But this strategy of test-taking is said to be best understood and utilised by adult test-takers only. Indeed, the present experiment conducted with adult non-native speakers of English shows that subjects did not utilise the "contextualising strategy" when performing in condition 2,

although this would have been mostly relevant. They seem to have tried to show their testwiseness instead. But at some point, the subjects found the material least predictable from a linguistic viewpoint. Reference to pragmatic knowledge was then essential to inferencing and testwiseness was not necessarily the best attitude to adopt. It also seems possible that self-censorship may have been exercised by some of them in their responses. Non-linguistic knowledge relating to one's experiential background and being culture- and language-bound, some subjects were not willing to reveal this background, as it were, and it appears that A students were typical of this attitude generally.

3.7.5 Concluding remarks

The empirical investigation conducted in this study has enabled the drawing of a pattern of inferential abilities involving the identification of lexical resources of cohesion and coherence in text/discourse. This pattern featured some significant values as can be seen on Table 014 below and in the ANOVAs entered in Appendix VI.

Table 014

Type of ANOVA \ Type of experiments	Experiment 1: "Order of clue"	Experiment 2: "Distance of clue"	Experiment 3: "Conjunctive clue"	Experiment 4: "Linguistic clue"
Difference between conditions	S	S	NS	S
Difference between yeartimes	S	S	S	≈ S
Difference between languages	NS	NS	NS	NS
Interaction yeart x language	NS	NS	NS	NS
Interaction yeart x condition	NS	NS	≈ S	NS
Interaction language x condition	S	NS	NS	NS
Interaction yeart x language x condition	NS	NS	NS	NS

At the beginning of this chapter, we raised three related questions (which were further recalled in Section 3.7.):

- 1 How do FL learners utilise lexical resources/links of linguistic cohesion and pragmatic coherence when inferencing unknown meaning while reading?
- 2 How does the use of lexical resources/links of cohesion and coherence vary as a function of FL proficiency?

3 How does the use of lexical resources/links of cohesion and coherence vary as a function of language background?

In order to answer the first question, we have devised a set of four experiments, each of which attempted to enquire into a certain type of inferencing ability. It is interesting that three out of four ANOVA results of the performance on the experimental variables (measured by different inferencing conditions) were significant. This is an indication that some inferencing strategies were more effectively utilised than others and bears implications for language teaching.

The second question derives from the first one and seeks to find differences in performance on the four experimental variables resulting from differences in FL language proficiency (measured by the yeartime variable). ANOVA's results of the difference between years in the four experiments were significant, reflecting a correlation between the cloze scores on the DV and the year factor. This enabled us to conclude that there was an inferencing expertise which reached a plateau in years 2 - 3. The pedagogical implications of such results are diverse and are treated in the next chapter. The third question closely relates to the second research question and also derives from the first one. It seeks to determine the extent to which the background language factor influenced the performance on the DV. But the non-significant ANOVA results of the difference between languages

in the four experiments evidenced absence of correlation between experimental variables and language background. These results suggest that the use of the lexical resources of cohesion (and coherence) vary not as a function of language background but as a function of language proficiency. However, it may be possible that the Placement Test whose aim was to discriminate between language dominance in subjects in effect revealed little information about their actual language dominance. One point is worth mentioning that has recurred in the four experiments is the little attempt on the part of some subjects from both language dominant groups, to use Arabic or French in their responses although they had been allowed and encouraged to do so. Some responses were given in English by A group and F group subjects. Looking at the A group performance in particular, 25 subjects out of 36 who answered in Arabic in the Placement Test, supplied responses in English. It seems possible that for psychological reasons, these students found odd or out of place to use a non-European language for a test written in a European language. This may have acted as an affective "block" in their output. This attitudinal variable was not anticipated and may have affected some results to a certain extent. Furthermore, 6 of the 11 subjects who did use Arabic in their responses, translated then into French, possibly for the "affective" reasons invoked above.¹² Or they may have perceived so strongly certain surface and deep-

structure constraints, which are language-specific, that they did not want to disturb them with the introduction, or rather, the intrusion of a foreign item, in Arabic or French, which itself had its own surface and deep-structure constraints. Ideally, an interview with the testees should have followed the experiments in order to illuminate these diverse and numerous questions with which the experimenter is necessarily confronted.

Certain methodological problems which produced unexpected results may have been avoided. Thus, some texts produced results different from other texts in the same experiment because they discriminated better between conditions than other texts. For instance (see *Misue Anal.*) in T90 the pragmatic clue, annoying little creatures did not seem to operate as a good clue for mosquitoes, the cloze item. In fact, reference to broken romances and marriages (explicit in the text) was picked up by several subjects who supplied women/girls for the gap instead of mosquitoes. Of course, relating the syntagm annoying little creatures to mosquitoes (bees and wasps were treated as near-correct) was a matter of world-knowledge. Those subjects who did not access the pragmatic implicature of this syntagm filled the gap with cats, hooligans, holiday camps, banknotes, tourists, rats, babies, ladies, girls, all of which were treated as incorrect items, although quite likely candidates. On the other hand in T78, the pragmatic clue, no money activated poor in almost all responses, which suggests

that there are "good" pragmatic clues which can guarantee original meaning and there are "bad" pragmatic clues which cannot. But should such responses as women/girls for T90 be treated as typically wrong or would it be more objective to consider them as "simply ethnocentric"?, as Jarrett (1984: 171) suggests: "To label a text incoherent if one does not understand its source and its purpose is simply ethnocentric". Other texts appear to have been "easier" because they contained a considerable amount of topical redundancy. For instance, in Texts 15 and 16, the redundancy of the linguistic and non-linguistic clues (occurring before or after the gap), all referring to the common topic of the war, made it clear to the testees that the cloze item (correct: diseases/epidemics, and near-correct: war/hunger/famine) should be topically related to the war (see App. II, III).

An interesting feature relating to the yeartime independent variable is the consistency of low mean scores by year 3 subjects in the first two experiments as compared to highest mean scores by year 2 subjects in the first two experiments, which reveals certain important aspects of FL learning. It seems that the learner's use of acquired competence reached a dormant state by the end of his 'language' training in year 2 as final year subjects were no longer to perform as successfully as middle year subjects. It may be that year 3 subjects have acquired sufficient "functional" competence or what Corder (1967) called "transitional competence" and hence not exactly stop learning but "go on to learn in

a minor way" (Selinker, 1972: 217).

This empirical study which had the purpose of testing how the lexical resources of a text set up various expectations in the FL reader, has revealed some of the underlying psychological principles involved in reading in a FL and at the same time it has enabled the characterisation of some of the reading problems encountered by FL learners which may be solved through adequate instruction. The next chapter is devoted to this aspect particularly. However, this study will not be complete without an analysis of some experimental items or 'Miscue Analysis' which may cast light on some of the errors made by testees and hence provide data for pedagogic research.

Part Two

3.8. Miscue Analysis

3.8.1 Preliminaries

The concept of Miscue Analysis is said to have been born in the U.S.A. at Wayne State University in 1965. From that period through the present time, a team of researchers then led by K.S. Goodman have been constantly making intensive studies of children's miscues in oral reading in their native language and into the complex mental process that constitutes reading generally. These studies were descriptive and attempted to analyse the observed oral reading responses of readers within a psycholinguistic framework. They utilized the Goodman Taxonomy of Reading Miscues as their instrument of analysis. The present miscue analysis however differs from the above-mentioned studies in two respects: (i) it does not concern young native speakers' performance but adult non-native speakers' performance. (ii) it does not concern oral reading but silent reading. Furthermore this study utilizes the cloze procedure. Nevertheless the basic notions utilized in those studies may account for non-native adult readers miscues in reading English as a FL such as the notion of 'miscue' defined as the deviation between the response of the reader (Observed Response) and the expected response of the text (Expected Response). Our basic assumption may be recalled here: Every response which the reader makes is cued in specific ways -by relating it to the experimental variables involved. Therefore responses will vary qualitatively.

3.8.2 Miscue Research and Analysis

A study of miscues produced by testees in the four experiments has revealed five characteristics that are being examined in the following sections viz:

- 1 Inadequate clue-searching strategy and inadequate response
- 2 Adequate clue-searching strategy but inadequate response
- 3 Reading without reasoning
- 4 Inadequate linguistic knowledge
- 5 Inadequate pragmatic knowledge

3.8.2.1 Inadequate clue-searching strategy and inadequate response

This feature accounts for responses that were counted as incorrect because the strategy selected for clue-searching was inadequate. This inadequacy was measured in relation to the four variables under study, viz "order", "distance", "linguisticness" and "conjunctiveness of clue. For instance, in T4 (Experiment 1 Condition 2), testees who produced the miscue cannot sell instead of disintegrate demonstrated backward clue-searching i.e. sell in pre-gap position was wrongly regarded as a clue item, whilst forward clue-searching was required to recover the item disintegrate. Similarly in T11 (Experiment 1 Condition 1) Cordoba (a miscue) was cued from right-located information relative to the gap, that is Cordoba was mistakenly treated as a clue item. It may also be the case that wrong responses were due to poor lexical (semantic/pragmatic) and syntactic knowledge generally and that for such cases, testees simply relied on chance. However it is interesting to note that the

lowest number of wrong responses due to incorrect clue-searching strategies was recorded in Experiment 2 (in both conditions), probably on account of the length of the texts which seemed to have enabled the use of appropriate clue-searching strategies and hence the production of correct responses.

3.8.2.2 Adequate clue-searching strategy but inadequate response

This feature seems to have been mostly present in Experiments 1 and 2. Because lowest scores were obtained when the clue was in pre-gap position (i.e. condition 1 in Experiment 1), this led us to examine the nature of miscues produced under that condition. Miscue analysis has thus revealed that wrong responses did not relate to the clue-searching strategy itself but to the type of clue item being selected by testees. Thus, responses judged semantically/pragmatically incorrect were "strategically" correct. For instance, in Text 1 (Experiment 1 condition 1), year, swim, cave, report, time were all items occurring before the gap (i.e. the strategy was adequate) but were mistakenly selected as clues by testees and often utilized as cloze items themselves. The ratio of miscues "strategically" correct being relatively high (2:3 on average in Experiments 1 and 2), this implied that failure to recover missing items had to be accounted for not in terms of "strategic failure" but in terms of lexical (semantic and pragmatic) incompetence. Indeed if the exploitation of ^{the} appropriate strategy did not generate adequate responses, it was often because subjects utilized any textual material anterior to the gap generally and thus revealed lack of knowledge of some 'key' items.

(e.g. reported, drowned, while in T1), or of lexicosemantic relationships holding between elements of the text. Relevance of an item as a clue does not seem to have been perceived "meaningfully" in terms of degree of semantic/pragmatic distance or proximity of this item to the gap. Further examples: in T5, incorrect responses such as Falklands, suffering, dangerous reflect adequate clue-searching strategy but lexicosemantic and pragmatic incompetences. Likewise in T21 (Experiment 2, condition 1) British and journalists treated as incorrect responses reveal adequate strategy although semantically and pragmatically incongruent. On the other hand, committees reflects inadequate strategy.

5.8.2.3 Reading without reasoning

Some miscues have shown that subjects simply filled in empty slots with any item taken at random from any stretch of language preceding or following a gap with no reasoning or real comprehension. It may be possible that clozefied texts do not allow reasoning to the extent needed in (reading) comprehension but rather compel the reader to surface reading. Clozefied texts seem to create artificial obstacles in the reader's mind and as they appear to interfere with his encoding process, this may inhibit his general comprehension and hence produce surface reading. Furthermore, clozes may prevent the reader from extrapolating, from going beyond the text (and using his pragmatic knowledge for instance). But it seems evident that this approach to text/discourse comprehension does not concern the good reader whose ability to make predictions is a deeply rooted reading habit. The

average to low-ability reader who does not possess these habits in the same degree of proficiency as the high-ability reader seems to be more concerned with text-decoding than with encoding its message, regarding text as having a form rather than a function. Decoding (or deciphering) is counterproductive comprehensionwise. The reader's syntactical faculties and general reasoning faculties are inhibited because his mental activity is totally absorbed by analytical, word-by-word-reading. Thus in Experiment 1 the subjects who supplied man, swim, as cloze items for Text 1 demonstrated their inability to build up the synthetic notion of accident (the missing item) from the event described by the lexical items man, swim, drowned, that is to reason that a man who drowns while swimming in the waters of a cave is a case of accident, not a mere description of a man or of a swim. We may note in passing that there is a somewhat incomprehensible attitude among some subjects (low-ability subjects generally) which seems to underlie a myth relating to the cloze exercise. Subjects believed that whatever linguistic form was provided by the author and hence was in the text should be reutilized in the response because it was likely to be correct. Hence some items in pre- and post-gap position were simply used to fill the gap despite their nonsuitability. Whether this phenomenon observed with cloze testing in SL/FL is typical of and restricted to this area or whether it affects other areas of SL/FL learning is a question worth investigating further. Moreover there seems to be a frequency factor ^{attached} to the production of this type of miscues to the extent that an item mentioned twice in the text (whether the same form was being repeated or not) was likely to be selected for the third time as a response. For

instance, the items history and historical having both been utilized by the author in T87 and T88, this may have led the students to the production of history, an incorrect item for the gap.

3.8.2.4 Inadequate linguistic knowledge

(i) Syntactic knowledge

It appears from analysis that some syntactic markers were problematic from the point of view of meaning and function and hence errors occurred from misunderstanding them. For instance, miscues such as violent, mobilised, calm (in T41) seem to have resulted from misunderstanding the syntactic marker however.

Similarly, the marker besides may have caused the production of the miscues less money, growth, contraception, more children, effects, no population.

Referential cuing was also an important factor of performance in inferencing lexical meaning. Subjects who had difficulty finding referents for deictics this, that seem to have encountered problems when 'spotting' clues to the cloze item and understanding the author's message generally.

(ii) Semantic knowledge

Lack of lexical semantic (or word) knowledge often resulted in testees' selecting any item occurring before or after the gap (The requirement of the Experiments was not to leave empty gaps), or utilizing incorrect clue-searching strategies. For example, it seems that responses such as divorce and old age in T53 were produced from misunderstanding of couples, reproductive age and contraceptives.

(iii) Discourse thematicity

This point closely relates to point 5.3.2.3 above insofar as inadequate reading skills(i.e. slow reading akin to deciphering more analytic than synthetic) may result in inability to perceive the thematicity (or topicality) of discourse. Themes and rhemes have specific parsing within and beyond sentences in texts and absence of sensitivity to these elements often resulted in the selection of wrong (thematic and rhematic) clues and hence the production of incorrect cloze items. For instance, subjects supplying cave when accident was expected in T1, or sell for fry in T7, or boys for thoughts in T30 thereby demonstrated their unawareness of the thematic value of drowned (T1), fried (T7) and memories (T30). Recognizing the thematicity of discourse is a reading strategy that should be developed early in the reader if he is to be a successful reader-anticipator .

(iv) Discourse functions

Miscue analysis has revealed issues relating to the recognition of the "discourse" relation underlying relationships between lexical items. Thus some discourse functions such as defining, exemplifying, specifying, classifying, contrasting, emphasizing seem to have been problematic for some readers. For instance, miscues such as year, swim, cave, report for T1 reflect not an inadequate inferencing strategy since all of these miscues occur in pre-gap position but an inability to spot the 'defining' function underlying the clue drowned . Similarly the discourse function 'contrasting' has created problems to some readers

even when it was signalled via discourse markers. For instance, miscues such as together, complementary produced in T50 indicate that the discourse function of adversativity underlying the discourse marker yet failed to be noted and utilized appropriately by the testees.

3.8.2.5 Inadequate pragmatic knowledge

It is interesting to note that few errors were made as a result of linguistic knowledge when linguistic information was available. Thus, in Experiment 4 when linguistic and pragmatic cueing were controlled systematically, most errors occurred on account of pragmatic clues as pragmatic information contained in texts was not always utilized efficiently by testees. Whether their preference for linguistic rather than pragmatic clues was due to psychological or sociological factors as suggested earlier (see 'General Discussion on p. 253) can only be speculative and verbalization may bring insight into these enigmatic questions.

Miscues described as pragmatically inadequate were those which did not take account of the pragmatic clues supplied by the writer and hence present in the text. Rather those miscues reflected subjects' pragmatic knowledge which was often irrelevant to the meaning intended by the writer. For instance, in Experiment 4, miscues such as crops (T98), power cuts (T83), telephone (T98), holidays (T94) reveal that the pragmatic clues supplied by the author viz those with no money (T78), events back after four years (T83), teenagers treat buses and phoneboxes (T88), women not at home all day (T94) were not perceived as pragmatically relevant to the cloze items in question by testees. Such responses are admittedly irrelevant both linguistically/semantically and pragmatically. Similarly in Experiment 2, the miscue quiet (T41) seems to reflect lack of knowledge about violent crowds demonstrating in streets: events and army are pragmatic clues indicating 'violent' rather than 'peaceful' demonstrations. Thus by failing to exploit those pragmatic clues supplied by the author and by utilizing their prior background knowledge, subjects imposed their own perception of the situation/event onto the text. Some cases of slogan- or cliché-based responses are quite interesting to report for they reflect an area of (reading) comprehension strategy little investigated. For instance, in T9, miscues such as communism, class struggle seem to have been "triggered" by the item Russian and have ipso facto become the testees' focus of interest. Likewise, miscues nomads and clean in T11 reflect clichés about the social life and religious practices of Arab Moslems.

Identification to and involvement with the subject-matter are additional features underlying some miscues or pragmatically irrelevant responses. Thus some testees have demonstrated strong feelings for some themes (in a non-technical sense) and have interpreted them in a 'personal' manner. This is an indication that reading is not only interacting with the text but also getting involved with it 'personally'. The theme of teenagers was topical in T88 and the text implied that teenagers' destructive behaviour was bad and had negative effects on society. The subject of teenage was reinterpreted by some testees rather subjectively as a 'social phenomenon' and hence teenage was supplied for the gap instead of vandalism. This type of personal involvement with the text although sometimes inevitable may have a blinding effect on comprehension.

3.9 Concluding remarks

This study has proposed to investigate empirically what aspects of the 'structure' clues the reader actually 'spots', how he uses them in reading (comprehension) and how schematic knowledge intervenes when pragmatic clues underlying lexical coherent links are to be identified in order to interpret the whole text/discourse. By examining what strategies non-native readers employed when they were faced with the experimental material --given that they were reading for meaning and relying upon their own resources in handling the material, and by inventorying and classifying ^{some} miscues, it has been possible to determine some of the possible factors which have

contributed to the production of correct and incorrect responses. Miscue analysis has uncovered some problematic areas of FL reading such as discourse thematicity, inadequate exploitation of textual material and lexical semantic and pragmatic incompetences which are worth investigating further. Thus "positive" and "negative" inferencing/reading strategies have been identified from miscues. Subjects who made errors that were reasonable, logical demonstrated that they were making use of all the available information i.e. they were trying to get meaning from the text in order to recover missing items. This inferencing/reading strategy can be described as "positive". On the other hand, subjects who did not bring common sense and previous language experience to their inferencing were apparently not reading for meaning as little thought was involved in their inferencing/reading process. Such strategy could be referred to as "negative".

It has been necessary sometimes to arbitrarily assign a causal relationship to some of the miscues. Also certain unclear cases have necessarily remained unclear. On the other hand, some cloze items may have been invalid because subjects were able to recognize them without reading the passage. Then verbalization and/or retrospective/introspective interview with testees would have been most useful. These methods of performance analysis will have to be carried out systematically in future investigation for optimum results.

Other factors apparent from miscue analysis and which may have interfered with subjects' performance are text editing and time constraints. These variables, although controlled, have produced unexpected results. Text editing has created a certain imbalance

between the different parts of the textual material, i.e. pre-gap and post-gap material (in Experiments 1 and 2 mainly). This may have produced a kind of visual conditioning in the testees (first and second year students mostly) as miscues implied. Thus if more textual material occurred before the gap, there was a tendency to search for clues in that part of the text, and vice versa. Such shortcomings in text design may have generated had hoc inferencing strategies. Nevertheless this can be avoided in future experiments as longer passages will have to be included in the design and the textual material will have to be distributed equally on either part of the gap (whether or not it contains clues). Overconsciousness of time on the part of some subjects seems to have generated errors that they would probably not make under normal conditions. Thus third year subjects' miscues were mostly pragmatic (i.e. overuse of pragmatic world knowledge) probably on account of the limited amount of time allowed to this category of students who did not exploit the linguistic material sufficiently and therefore relied on non-linguistic information primarily. These are contingencies in empirical research which often do not appear before experimentation. Their effects felt and weighed when results are analyzed should be viewed as guidelines for better design.

Thus the results of the present investigation have inspired miscue research and analysis and the latter have brought insight into the utilization of some specific strategies and may provide a framework for future research in inferencing and in reading in a FL generally.

Notes on Chapter 3

- 1 The Placement Test also indirectly tested general knowledge. The subjects were required to supply a minimum of ten words per topic, but interestingly, the number of responses per topic increased (over ten responses) with the year of course. This may be evidence of increasing general knowledge with years.
- 2 E A Levenston et al (1982). "Discourse Analysis and Reading Comprehension". Paper presented at the International Symposium on LSP, Eindhoven, Holland. Quoted in Deyes (1984: 128).
- 3 Perhaps one of the disadvantages of having similar patterns for all experimental texts is that it may produce a 'psychological set' in the testees, a certain kind of expectation which may interfere with their performance.
- 4 Thanks are due to native speakers at the University of Sheffield and at the Walkley Institute for Continuing Education. Special thanks to Liz Kirby, Larry Furlong, Sally and John Capes, Mrs. Raffo and Mrs. Scott.
- 5 Acknowledgements to Dr. C. Whitaker (University of Bangor) and Dr. M. Djeddi (University of Sheffield) for statistical and computing advice.
- 6 There are two semesters per year in the Algerian university system, and six semesters of the undergraduate course are needed for the obtention of the

"Licence d'Anglais". During the first half of semester one, emphasis is laid on the learning of the 'mechanics' of the English language as a prerequisite for the learning of reading and writing skills. The syllabus then includes the teaching of punctuation, paragraph boundaries, logical structures, that is, the physical properties of an English text. 'Guided' reading which involves answering questions in oral and/or written form on short passages, is attended to only later during the first year, but it is largely developed in the second year from semester three throughout semester four. During that year the syllabus focusses on contextual vocabulary building, denotative and connotative meaning awareness, different types of text recognition (descriptive and argumentative) which contribute to the building up of 'strategies' of text comprehension.

7 We may consider the following time allowances for each year for comparison: First year students receive a total of 192 hours per academic year of tuition in 'reading', that is, 6 hours per week. Second year students enjoy 384 hours per week, that is, 12 hours per week, and final year students also receive 384 hours per week, that is, 12 hours per week.

8 An error analysis would indeed bring light into the most problematic kinds of clues and this would enable the student to receive a more appropriate type of teaching in terms of lexical devices of cohesion.

9 L1 is Arabic, L2 is French and L3 is English for the A group. Arabic refers to the two varieties of Arabic common in diglossic situation, that is, "H" or a "high" superimposed variety which is "highly codified (...) the vehicle of a large and respected body of written literature (...) which is learned largely by formal education and is used for most written and formal spoken purposes but is not used by any sector of the community for ordinary conversation " (Ferguson, 1964: 435) and "L" or a low variety which comprises the (spoken) dialect(s) of the language. Arabic dominant speakers therefore tend to use "L" at home and "H" outside home. French dominant speakers, on the other hand, tend to use "L" and French at home and outside home, that is, French in spoken and written form. A few exceptions, entered in the category of French dominant speakers, had no knowledge of "L" or "H". They included speakers of Berber, a language used by a relatively large speech community, and students who did their (primary and) secondary schooling in France, and therefore had virtually no knowledge of "L" or "H". This group is usually referred to as "étudiants émigrés" in the Algerian educational system. Also some mature students who did their secondary education in Algeria through the medium of French before the "Secondary System Reform" of the 1970s. Thus, the Placement Test yielded results which enabled us to classify the population into two main language groups: Arabic dominant group whose L1 was

considered to be Arabic (H or L) exclusively, and French dominant group whose L1 was considered to be French exclusively.

- 10 However, this does not rule out that some legal texts containing an impressive amount of syntactic marking (yet, nevertheless, notwithstanding, etc.) are as ambiguous to the lay reader. This bears the implication that excessive use of these devices may, in effect, impair text/discourse comprehension.
- 11 In Hasan's (1984) recent study on native speakers' awareness of cohesion (and coherence) in text, she seems to argue that cohesive texts are unquestionably more easily interpretable by native speakers than coherent "texts" (quoted by Urquhart, 1984, in Reading in a Foreign Language, Vol.2/2, pp 295-304, reviewing J. Flood's (ed.), 1984, Understanding Reading Comprehension, International Reading Association, Newark, Delaware).
- 12 Here are some examples of their responses, where one response, either in English, Arabic or French, would have been sufficient.

T15/T16 maladies (French)/ maRad (Arabic: مرض)
T23/T24 vin (French)/ XamaR (Arabic: خمر)
T21/T22 police (French)/ شُرطَة (Arabic: شرطة)

CHAPTER 4

PEDAGOGICAL AND OTHER IMPLICATIONS

4.0 Introduction

The present study was based on the postulation of two distinct though complementary levels of linguistic analysis, the level of 'text' and the level of 'discourse', (which is a practical distinction of the more fundamental separation between 'la langue' and 'la parole'), and has enabled us to differentiate between two distinct though related phenomena, cohesion and coherence. In the light of Widdowson's analysis of cohesion as a feature of "text" and coherence as a feature of "discourse", we have further analysed these two notions as lexicosemantic cohesion and lexicopragmatic coherence, thereby considering the lexical element only in cohesion and coherence, and we have taken a reader's viewpoint regarding his use of vocabulary relations, rather than a writer's viewpoint. Lexical semantic relations holding in text are an inherent part of cohesion. As noted in Chapter 1, "However luxuriant the grammatical cohesion displayed by any piece of discourse, it will not form a text unless this is matched by cohesive patterning of a lexical kind" (Halliday and Hasan, 1976: 292). Furthermore, discourse non-linguistic relations determine coherence to the extent that any piece of text will not form a discourse without lexical pragmatic relations. Thus, lexicosemantic cohesion and lexicopragmatic coherence were viewed as

relational concepts concerned with relationships among lexical items in text and their values in the interpretation of the discourse, that is, how the "literal" meaning of text underlying certain vocabulary items ("co-textual" meaning) is interpreted in discourse ("contextual" meaning). The lexical devices responsible for the production and the interpretation of cohesion and coherence have been studied within a 'functional' or 'communicative' approach, an approach which favours vocabulary as a discourse component because, as demonstrated in Chapter 1, the phenomenon of lexical cohesion could not solely be described within a 'semantic' theory of text analysis à la Halliday and Hasan, and the interpretative aspect of lexical relations of cohesion (and coherence) underlying reading comprehension (in a native language or in a foreign language) could only be accounted for within a general theory of discourse and pragmatics.

In the attempt to analyse lexical cohesion in the light of pragmatics, a study of the concept of "background knowledge" was proposed in Chapter 2 where notions utilised in Cognitive Psychology such as frames, scripts and scenarios (all distinguished under the superordinate notion of 'schemata') were introduced to account for the linguistic and pragmatic levels involved in the interpretation of relations of lexical cohesion and lexical coherence. Thus, in Chapter 2, content and formal schemata were shown to underlie linguistic lexical

and pragmatic lexical competences, and could potentially explain why certain types of lexical linkage were easier to interpret than others. Empirical investigation in the previous chapter brought evidence about the activation of these schemata in inferencing unknown meaning when clues available to the reader were linguistic semantic or non-linguistic pragmatic. Miscue Analysis has also cast light on some of the problematic areas testees have encountered when inferencing / reading in a FL.

In the present chapter, we suggest to examine one type of miscue in particular viz miscues relating to discourse functions. It is hoped that the outcome of this research gives insight into 'communicative' language teaching, vocabulary teaching essentially, and inspires future research on this aspect of discourse comprehension.

Thus, two areas of applicability of this study are considered, viz. FL teaching and applied linguistic analysis. We also plead, in this chapter, for the development of a "performance model of cohesion" to complement competence models of cohesion as expounded by Halliday and Hasan (1976). Performance models are more realistic accounts of language behaviour than competence models as they emphasize language processes rather than language structures. It is hoped that our contribution to the understanding of lexical cohesion will bring insight into the development of such a model.

4.1 Implications for FL teaching: developing lexico-communicative competence

4.1.1 Analysis of items involved in cloze experiments: a problem in recognising discourse functions

Evidence from our investigation has uncovered one of the perceptual strategies¹ at work while reading in a FL, viz. inferencing unknown lexical meaning. Most importantly, an experimental item analysis conducted together with an analysis of subjects' responses on these items has revealed that the heuristic strategy² of lexical clue searching has been interfered with by another component of this strategy, that is, recognition of the "discourse" relation underlying relationships between lexical items. The analysis of items has yielded a correlation between the type of "rhetorical act" in Widdowson's sense, or "communicative function" to use loosely Wilkins' terminology, underlying certain lexical relationships and the type of cloze item to recover. In other words, whether the testees were capable of inferencing unknown meaning seems to have been dependent upon their ability to recognise the 'discourse' function of defining, exemplifying, specifying, classifying, emphasising, generalising (these functions are often utilised in ESP/EST written discourse. See Widdowson, 1978, 1979, on this point) being performed by the propositions containing these items and the data indicates that some "communicative functions" have been more difficult to identify than others. Of course, to prove this point, a

new set of experiments in which different communicative functions were an independent variable, will have to be conducted. This additional factor of difficulty may be explained by the FL learning history of the learners. It seems possible that because these "communicative functions" were not taught explicitly at secondary level, but learned incidentally, it was difficult for the learner to predict their occurrence in the passages. As Widdowson (1979: 16) rightly pointed out, "You do not necessarily teach rhetorical acts when teaching linguistic elements and vocabulary items, as we all know, and what communicative competence the learners do acquire tends to be picked up incidentally". It is suggestive that first year students experienced such difficulties in the most acute way, which seems to agree with Cziko's (1978) claim that a relatively high level of competence in a language is a prerequisite to the ability to use discourse constraints as a source of information in reading.

Since this study is concerned with the use of lexical information to inference unknown meaning and since it seems that such use is dependent on an understanding of specific types of coherence relations in text, the question of whether vocabulary can be taught within a functional/communicative competence framework is ultimately raised. We shall attempt to answer it in Section 4.1.2 below. To illustrate the remark about the recognition of discourse functions underlying certain lexical relationships, we provide the following example

where the 'defining' function caused difficulty to the testees overall. The poor performance of some subjects in both conditions (in experiment 1) seems to have been the effect, not of the order of the clue (which was the IV being tested), but of the function of the clue and this has somewhat depressed the value of the experimental variable that the experiment sought to investigate. Examining some experimental items, it became fairly clear (although some doubt still remains about the lexical competence of certain testees) that the inferencing problem lay in the specific discourse function underlying the proposition(s) expressed by lexical items. Consider the following example (See also 3.8):

Text 1 The NCP reported this year a man who
(Condition 1) drowned while swimming in the under-
 ground waters of a cave in Derbyshire,
 another _____. This seems to happen
 regularly in the spring season.

Text 2 The NCP reported this year another
(Condition 2) _____; a man drowned while swimming
 in the underground waters of a cave in
 Derbyshire. This seems to happen
 regularly in the spring season.

These two texts attracted the following responses (treated as incorrect) for the 'gap': man, year, cave, season. The clue drowned was not easy to spot, whether it occurred retrospectively (as in text 1) or prospectively (as in text 2) to the gap, presumably because the subjects

could not recognise that the whole proposition expressed by the sentence preceding the gap or following it, had as one of its functions, to define the correct cloze item, accident.

4.1.2 Recognising discourse functions and using them in reading and writing: on the role of discourse markers

It is often assumed that the recognition of the function of lexical relations in reading is a function of the number of discourse markers present in the text, and as a result, a text light in these signalling devices (grammatical or paralexical) is likely to give rise to difficulties in FL reading comprehension. In other words, the recognition of the "communicative" function of lexical relationships seems to be subordinated to the presence of discourse markers in text. Their facilitative role in reading comprehension clearly emerges when FL learners are required to reconstruct passages from jumbled up sentences as part of a reading comprehension test. In their absence the learners seem to be bewildered and their only resource seems to be to rely on pragmatic/factual knowledge to organise the experience/reality described by the discourse.

On the other hand, learners often do not rely on these devices to achieve cohesion and coherence in writing and implicitly expect the reader to make appeal to his world-knowledge to interpret their production (the use of

discourse markers in writing is appreciated by language instructors for it gives smoothness to a piece of writing and reduces inferencing on the part of the reader).

The following two examples illustrate these points. Text A is taken from an undergraduate test paper³ in which the student was asked to reconstruct a text about a burglary with murder from sentences that had been mingled.

Text A The policeman asked her how she discovered her aunt's body. Mrs. Smith was upset and sobbed: "my aunt has been killed. She is in her bedroom. The knife is lying on the bed beside her". The policeman asked her who was in the house that night. Mrs. Smith replied she discovered the body when she went into the bedroom to show her aunt her new dress. Mrs. Smith sobbed and said she was the only relative her aunt had. "Nobody was in the house" she said. (K.S. First year)

Clearly this reconstruction⁴ demonstrates that perhaps because of the absence of discourse markers (grammatical and paralexical) the producer of this text opted for a personal narration of the events, that is, relying on her own (schematic) knowledge about murders. The learner was aware that the passage had to be reconstructed along the lines of a conversational exchange between a policeman and a Mrs. Smith but little regard was given to the linguistic lexical devices that held the text together. Admittedly, there is little use of lexical and grammatical devices of cohesion (for example, lexical reiteration of Policeman, presence of converses, asked,

replied, anaphoric pronouns). The reader of text A is somewhat puzzled by this incoherent narration which reflects a process often encountered among FL learners, as Steffensen et al (1984: 60-61) observe: "they distort meaning as they attempt to accommodate even explicitly stated propositions to their own pre-existing knowledge structures".

Another pilot study has evidenced that learners of English do not use discourse markers in their productions and have preference for pragmatic links of coherence over linguistic links of cohesion. A study of 48 student-produced texts⁵ has demonstrated over reliance on implicit links underlying lexical relations of coherence and underuse of explicit links of cohesion, that is, little use of discourse markers of "conjunction" such as so, then, however.

Text B "I can tell you this true story. We have this family of six children. The father does not work. He earns no money. They are poor. They need the help of all the family. One day one of his children is ill. Who will pay the doctor and the medicine? This family is in a dramatic situation. How can she (sic) save money, etc. ..." (B.Y. First year)

It is interesting to note that the producer of text B, a FL learner described as an "underachiever" by the language instructor, was parsimonious with regard to

lexical links of cohesion. (His use of grammatical links of cohesion, however, violated the Maxim of Clarity (Leech, 1983: 66) with the misuse, or rather, the ambiguous use of anaphoric pronouns which resulted in a fair amount of inferencing on the part of the reader). Greater appeal is made to the reader's world-knowledge to build a mental picture of the text and to see what kinds of "communicative" acts are being performed by the propositions underlying the text (for example, causal relation inferencable from does not work - no money, from no money - poor, and poor - (need) help. That FL learners prefer (lexical) coherence links to (lexical) cohesive links in production seems to be the reflection of what happens in reception whereby the search for coherence links is the guiding principle in reading (comprehension), whether in native or foreign language. Ultimately, the use of lexical cohesive devices seems for some producers, a 'stylistic luxury' rather than part and parcel of the skill of producing coherent compositions.

The producer of text A above created coherence links between lexical elements of the text by projecting her own 'logic' onto the discourse, a logic which does not always coincide with that of the author of the original text. However, had the subject been more attentive to lexical signs of cohesion, such as lexical reiteration, links produced via anaphora, etc., this would have enabled better reconstruction of the original text. On the other hand, text B suggests that coherence links are indeed interpretable in so far as reader and writer share

common background knowledge.

These practical illustrations somehow demystify the role of explicit discourse markers as sole exponents of 'communicative' discourse, and pleads in favour of teaching the pragmatic function of lexical relations in discourse (for example, how to recognise conjunctive meaning expressed lexically as in text B) as a complement to the teaching of overt discourse markers of cohesion and coherence.

4.1.3 Some problems on FL vocabulary teaching

4.1.3.1 Vocabulary development and vocabulary recognition:

The results from our empirical investigation on inferencing unknown lexical meaning in reading comprehension and the findings from item analysis have direct consequences on vocabulary teaching and learning. Although it seems quite agreed now among methodologists and language instructors that the area of vocabulary teaching should be in the priorities of FL teaching rather than relegated to secondary position, the question of how to teach vocabulary for "communicative" purposes remains an essential issue in FL instruction. Teaching de-contextualised vocabulary items with the hope that students will retain them and use them for receptive and productive purposes in reading and writing classes does not seem to have solved the problem of lexical "communicative competence", nor has it proved useful in improving lexical performance specially when new/unknown vocabulary items

are encountered in reading. Thus, adequate teaching of FL vocabulary should take the following criteria into account if it is to develop learners' awareness of relations of cohesion and coherence in text/discourse as part of his lexicocommunicative competence:

- 1 Vocabulary development should concern more the learner than the teacher, and because there is no end to learning FL vocabulary and to forgetting it either (see on this point Ott et al, 1973, on one FL vocabulary learning strategy - the mnemonic strategy) more effort should be made on developing learners' strategies to learn and retain more vocabulary and to solve reading problems when unknown items are encountered. Vocabulary development usually occurs at an early stage of FL learning. It refers to the deliberate and systematic expansion by the teacher of the learner's stock of words on semantic lines usually⁶, as for instance, by means of lists of words thematically selected, dictionary exercises and form and class exercises. But these vocabulary activities provide the learner with no knowledge of the way vocabulary items acquire meaning in context and have purpose in discourse, a function as they communicate meaning to the language user, and give little incentive to the learner to develop strategies of meaning recovery when unknown vocabulary items are encountered. This leads us to the second criterion of vocabulary recognition.

2 Vocabulary recognition in actual reading is unlikely to be helped by decontextualised vocabulary teaching as described in 1 above. Indeed, inability to cope with reading materials often lies not in the learners' insufficient knowledge of vocabulary but in their inability to deduce meaning from context, and context, specifically to understand the meaning relationships between vocabulary items in text and their function in the interpretation of the discourse. Thus, in order to ensure that adequate vocabulary development will result in the development of successful 'tools' for reading (comprehension), vocabulary instruction should include not only "usage" instruction (which can be paradigmatic, focussing on how word meaning is a function of other words with which it contrasts in the language system, or syntagmatic, focussing on how word meaning is a function of (syntactic and) lexical relationships among the words in text) but also on "use" instruction, taking account of the "communicative" function of contextualised vocabulary, as a learner may know the meaning of every word in a passage without being able to understand it as discourse.

4.1.4 Towards 'communicative' vocabulary teaching

The form and function of vocabulary have recently received attention by lexicologists and language instructors concerned with the teaching of FL vocabulary "communicatively". Thus, basing his instances of dialogic/conversational discourse, McCarthy (1984)

points to the pragmatic potential of lexical relations and notes they they are worthy of attention for language teaching. What the language teacher needs to communicate to his learners are the communicative effects of lexical relations such as synonymy, antonymy, hyponymy, across discourse boundaries. In conversational discourse, lexical relations realise important functions such as concurrence, divergence, topic-change, transaction-closing. Most of these functions are embodied in the semantics of the text, specifically in lexical cohesion. For instance, the pragmatic function of 'agreement' may be signified in text by the device of lexical repetition and synonymy. Consider the following examples:

1 A = This engine is useless.

B = Yes it is.

2 A = This engine is useless.

B = Yes, useless.

3 A = This engine is useless.

B = Hopeless.

'Agreement' is achieved via grammatical cohesion as in 1. where useless was ellipted, or via lexical cohesion as in 2. with the lexical repetition of useless and 3. with the use of synonym hopeless. The learner needs to know that these three examples are functionally equivalent although formally different.

'Agreement' may also be achieved via antonyms:

4 A = Joe didn't stick to the subject.

B = He wandered off too much.

via more general words (hyperonyms):

5 A = The cat is great company.

B = All pets are.

or via more specific words (hyponyms):

6 A = Pets are great company.

B = Especially cats.

As McCarthy (1984: 19) rightly comments, "at the level of discourse, categories such as synonymy and antonymy cannot simply be dealt with in a monolithic, abstract way as fossilised relationships within the language code, but must be incorporated into a model of continuous re-classification. In practical terms this requires making the learner aware that the speaker/hearer's lexical choices are in constant relation with one another and affect the communicative function of utterances."

It is interesting to note that such relations, subject to continuous re-classification in discourse, produce text-hyponymy and text-synonymy that are not always easily reversed because they become pragmatically 'marked'. This feature does not seem to affect written discourse to the same extent. For example:

7 A = Were you angry?

B = Yes, I was absolutely furious.

8 A = Were you furious?

B = *Yes, I was absolutely angry.

In order to produce 7. rather than 8. the learner must have knowledge of the (lexical) semantics and the (lexical)

pragmatics of the FL, that is, knowledge of the relatedness in semantic meaning of angry and furious (they are text-synonyms when defined on 'broad' semantic lines), knowledge of the semantic coreness of each item (angry is more 'core' than furious), and knowledge of the communicative effect of their coreness in an interchange: certain questions containing non-core items (for example, furious) are 'marked' and subsequently cannot accept responses containing core items, which are 'unmarked' (for example, angry). 'No' would be a more natural response for 8B: "No, I was absolutely angry". Thus, angry and furious are also "value" hyponyms (not "signification" hyponyms, in Widdowson's terminology), which cannot be reversed. (This example demonstrates aspects of lexical cohesion that tend to arise more in conversation than in written text and may be the object of a separate study). A similar example, worthy of attention, concerns written discourse which involves non-reversible lexical items.

9 My sister is married and her husband works for NASA.

10 *My sister's husband works for NASA and she is married.

For reasons of presuppositional meaning, married and husband cannot be reversed in 10. This presuppositional incongruity has a pragmatic origin: husband implicates married, in which case the occurrence of married in 10. is unnecessary because it disturbs the 'logic' or rather, the 'pragmatics' of the utterance.

Thus, knowledge of relations of scale and intensity⁷

of intonation (when the dialogue is in its normal spoken form) and of the potential of items for creating text synonyms is of considerable value to the learner. These are the lexical pragmatics that should be of concern to every language teacher in order to develop in the learner lexical competence and lexicocommunicative competence. Current vocabulary teaching methods are characterised by an atomised approach to vocabulary which results in a decontextualised abstract teaching of sense relations between lexical items and are therefore inadequate for reading comprehension purposes.

Other pragmatic functions can be achieved via lexical relationships, as for instance:

11 A = John looks happy.

B = He's won the pools.

The relationship between happy and won is not semantic but expresses the pragmatic function of "explanation". A, by stating John's happiness, is probably seeking an explanation, and B realises this and supplies it to A. Thus, adequate vocabulary instruction should develop the learner's awareness of how lexical semantics and lexical pragmatics interact in text to make up the discourse, and how they should be exploited for reading comprehension purposes.

Williams (1980) and Bramki and Williams (1984) propose to exploit the "lexical familiarisation devices" supplied by the writer. Such devices

include "exemplification", "explanation", "definition", "stipulation" and "illustration" which are pragmatic functions of (scientific) discourse often explicitly signalled in the text via discourse markers but supported lexically. For instance, "exemplification" can be achieved by contrasting the newly-introduced term with a situation that the author believes the reader already understands.⁸

4.1.5 Developing inferencing strategies as part of "receptive" communication strategies involved in lexicocommunicative competence development

4.1.5.1 On "receptive" communication strategies

Thus, alongside the teaching of vocabulary itself, the teaching of receptive "communication strategies" (Tarone, 1974, 1981) appears essential to build up and shape the learner's receptive comprehension competence. Such strategies may include recognition and exploitation of links between lexical elements of juxtaposed sentences. The learner's attention should be drawn simultaneously to the relatedness of vocabulary items in the text which results from semantic linkage, as noted earlier, and to the relatedness of pragmatic origin which expresses the communicative import of lexical relations in the discourse. Awareness of these two aspects constitutes the language user's linguistic competence and pragmatic or "communicative competence" (in the sense defined by Allen and Widdowson (1975), Criper and Widdowson (1975) and Widdowson (1978), that is, the ability to interpret discourse, to realise what

"communicative acts" are being performed by propositions containing linguistic elements), competences not necessarily co-existent and interactive in the FL learner. "Communicative competence" is often acquired at a relatively later stage of FL learning. As noted earlier, students' "threshold level" reached in secondary schools is usually entirely based on knowledge of the language system but not on how this system is effectively put to use in communicative discourse, their knowledge of lexis and syntax is often the result of meaningless rote learning. If the introduction of 'new' teaching methods based on the notional-functional approach seems to have achieved the development of syntactico-communicative competence in the learner, their handling of vocabulary teaching is still unsatisfactory to achieve lexico-communicative competence in the learner. Some "communicative" course books which have flourished over the past decade are still structural/slot filling in their methods and often do not offer a rich variety of lexical content in relation to the communicative needs of the learner. However, this last point will not be pursued further.

Thus, because "communicative competence" is acquired relatively late, this may be one of the factors which can explain why it is most problematic to the FL learner, although his lexical competence may be well developed. For instance, sensitivity to formal and functional links of lexical cohesion and coherence in a given passage is unconscious in the native-speaker reader. The FL learner, on the other

hand, may master the former but not the latter. Consider again the following example (Carrell, 1982: 484):

12 The picnic was ruined. No one remembered to bring a corkscrew.

The native speaker's linguistic (semantic competence) will enable him to recognise this piece of language as a text made up of two sentences. His "non-linguistic" communicative competence will enable him to see that both utterances refer to the conditions attendant on the act of explaining. The first utterance states a fact. It contains the lexical item ruined which may be viewed as a "predictive verb" (in Winter's (1977, 1982) sense) as it anticipates a clause which will provide an explanation to the facts stated in the first utterance. The second utterance hence gives the reason for the picnic's ruin. Thus, the link between these two utterances lies in their relation to the communicative act of explaining (act signalled or 'predicted' by ruined) and therefore corkscrew and picnic are perceived as causally connected. Supposing that the meaning of corkscrew is unknown to the reader, he will be able to approximate its meaning only if he understands the communicative function (that is, explanation) underlying the two utterances, and also by reference to his world-knowledge.

Granted that the cause of the FL learner's lack of sensitivity to discourse constraints has to be sought at secondary school level where the development of "communicative competence" is often inhibited, "remedial" teaching ought to be done at tertiary level in order to

allow linguistic and communicative/pragmatic competence to develop in harmony and to be used for receptive purposes. Thus, in order to ensure that "Greater concern should be given to seeing that what is learned has communicative value and that what has communicative value is learned" (Wilkins, 1979: 92) and also that what is learned can be used receptively, exercises aimed at improving receptive competence (which involves both semantic and pragmatic competence) must be included in a reading syllabus.

It would have been useful to have required from the testees responses in English only so as to be able to conduct an error analysis of their responses and thereby access their "interlanguages". However, an analysis of testees' products would have interfered with the study of the process and would have distracted the main purpose of the investigation. Although responses supplied in the target language (or in the native language when the subject could not find the right word in his working and long term memory) may be viewed as an index of the learner's production strategic competence, we were trying to avoid testing production as well as comprehension. Nevertheless, the experiments have all involved testing learners' receptive/comprehension strategic competences, whether the responses were in NL or in FL. Future research may involve the testing of native speakers' receptive strategic competences in identical experimental conditions and the comparison of their performances with non-native speakers.

4.1.5.2 Improving inferencing strategies for reading comprehension: some practical exercises

The acquisition of inferencing strategies as part of the development of receptive strategic competence appears to be essential at any level of reading activity, scanning (or the quick search for a particular piece of information in a given text), skimming (or the perusal of a text, picking out elements of information here and there to obtain an overall picture) and in-depth reading (which involves a closer examination of a text to obtain even more information and to build a more complete picture of the text). It is interesting to note that the data obtained in our investigation suggests tendencies of learners to process specific inferencing strategies at different developmental stages of their proficiency in the target language (which express different stages of their "interlanguages"). Thus, they appear to have relied more on lexical cohesion embodied in text-defined or text-oriented meaning at an earlier stage of development of FL proficiency (the first two years) and less on lexical relations of coherence which are embodied in discourse-defined or discourse-oriented meaning. At a later stage they seem to have relied on both types of relations. Hence, final year students' performance was higher overall probably because of their ability to work out coherence links implicit in the text (and as, successfully, links signalling cohesion). Thus, sensitivity to the fact that lexical meaning occurs in "linguistic clusters" as well as "discourse clusters"

(Anthony, 1975: 24) seems to be a factor of "communicative competence" development. The implications for the language instructor are obvious: exercises to improve the reader's lexicocommunicative competence, and "communicative competence" generally should be introduced into the reading syllabus at early stages of foreign language learning. Also adequate FL vocabulary teaching should include inferencing strategies (or clue-searching strategies) as a way of eliminating wild guesses when unknown items are encountered. On this point, it is worth noting that the concern of structural methods of teaching EFL focussed on WHAT to teach (phonological grammatical and lexical elements), the notional-functional method dealt with this WHAT in a more satisfactory fashion, but neither have attempted to answer the HOW to use the pedagogical material, namely, the vocabulary it contains, in order to ensure a better start to the learner or to provide him with a "jumping board" for his achievement of "communicative competence". Rather than attempting to deal with the text globally in reading comprehension, the language instructor should induce the learner to utilise its lexical content in a systematic way, viz. to exploit its lexical devices of connection because, as Galisson suggests (1983: 3), the text can be in the service of the words in the same way as the words are in the service of the text. ("En période d'apprentissage au moins, je suggère que le texte soit mis au service des mots, comme les mots sont mis au service du texte").

The following are types of exercises to develop

learners' lexical (linguistic) competence and lexico-communicative competence at higher intermediate to advanced level. If the number of unknown items encountered is high (for instance, four to five within a span of three sentences), the learner will need to improve his lexical competence/knowledge. "Lexical grids" (Harvey, 1983) or 'schemata building exercises' related to building mental "frames" or "schemata" (see Chapter 2) can be devised to organise and categorise vocabulary into related areas. As a vocabulary building exercise, it can be an effective way of approaching topics in reading comprehension as it develops awareness of semantic and pragmatic meaning. For instance:

	<u>Example</u>	<u>Person (N)</u>	<u>Place (N)</u>	<u>Action (V)</u>
Transport	bus	driver	street	drive
Animal	sheep	shepherd	farm	watch
Sport	tennis	umpire	tennis court	score
Job	typing	typist	office	type
	teaching	teacher	school	teach
	cooking	cook	restaurant kitchen	cook

The following type of exercise (adapted from McCarthy, 1984) will develop the learner's "micro" and "macro-abilities" (Baltra, 1983: 27) by learning to recognise lexical relationships of cohesion and coherence. The subject may be given a set of random sentences including one or two distractors, that he will arrange following lexical/grammatical cohesion and coherence as organisational criteria for interpretable text/discourse.

- 1 The kitchen is being redecorated.
- 2 Cauliflowers are very nice.
- 3 Pets are great company.
- 4 The bathroom is almost finished.
- 5 They are cheaper this week.
- 6 The tiles are very plush.
- 7 We want it finished for the barbecue party.
- 8 Specially cats.
- 9 Large mirrors always give elegance to bathrooms.
- 10 Mine were Siamese.
- 11 Have you tried them in croquettes?
- 12 The poor little things died of food poisoning.
- 13 They are mixed together.
- 14 Some people like them raw.
- 15 Maybe they'll come back.
- 16 This will be impossible: the plumber has gone on holiday.
- 17 Siamese cats are frail creatures.
- 18 Or in a curry?
- 19 John bought the kitchen units last Christmas.
- 20 Vegetables are good value this summer.

Thus, the purpose of this exercise is to make the learner aware of the "company words keep", and of the cohesive power of such relations as synonymy, antonymy and hyponymy across sentence boundaries, even in the absence of more obvious discourse markers. For example,

1 - 19 - 7 - 16 - 4 - 6 - 9, displays a chain of lexical elements which relate one to the other through collocation generally:

Kitchen - bathroom (hyponymy), bathroom - tiles -
mirrors (partonymy):

"The kitchen is being redecorated. John bought the kitchen units last Christmas. We want it finished for the barbecue party. This will be impossible: the plumber has gone on holiday.

The bathroom is almost finished. The tiles are very plush. Large mirrors always give elegance to bathrooms".

Thus, instance 2 is unlikely to follow instance 1 because kitchen units is a more appropriate co-textual and contextual element than cauliflowers, this basing one's judgement of appropriacy on linguistic as well as factual knowledge.

This third type of exercise may be devised to help the learner compare paraphrases "out of context" and "in context", that is, how paradigmatic paraphrase equivalents may be used "in context" as syntagmatic paraphrases, thereby focussing on semantic-definitional and pragmatic-non-definitional meaning. For instance:

(Paradigmatic) paraphrase:

You are tipsy

= you are slightly drunk

(= synonyms)

John is thick

= John is not clever

(= antonyms)

(Syntagmatic) paraphrase:

You are tipsy. Have you

been drinking at Harry's?

John is thick. He never

understands a joke and

cannot play any game.

(Paradigmatic) paraphrase:

I went to the haberdasher's
= I went to the shop that
sells thread, ribbons and
buttons

(Syntagmatic) paraphrase:

I went to the haberdasher's.
I needed some blue thread
to sew the pocket of my
trousers.

This last type of exercise may be devised also to improve the learner's inferencing ability when encountering unknown vocabulary items. It involves practising the use of a dictionary: learn how to read and test dictionary definitions (for example, to sneeze = to make explosive sound in sudden involuntary expiration to expel anything that irritates interior of nostrils - O.E.D., 1976), and learn to familiarise with non-standard definitions (Scholfield, 1979) (to sneeze: you do this when you smell a strong spice, as for example, black pepper). Exercises based on the principle of definition-finding, reflect "l'approche sémasiologique" (Galisson, 1983: 15) from the words to the notions they depict, and may improve the learner's linguistic lexical competence. Exercises involving word-finding (as for instance, cross-words) reflect "l'approche onomasiologique" (ibid, p15), where the process is from the notions or ideas to the words or linguistic forms. The latter type may improve the learner's lexico-communicative competence in so far as it goes beyond the linguistic to integrate what is extra or non-linguistic, that is to say, the psychological, sociological and cultural aspects of communication so vital in FL learning.

4.2 Implications for applied linguistics research

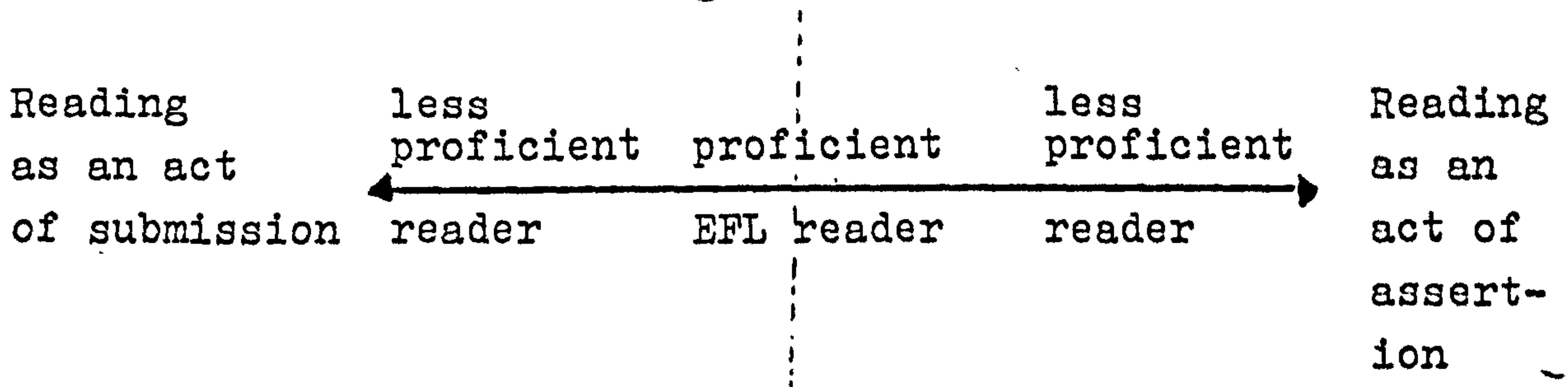
In 1965 (Chapter 1), Chomsky stressed that the study of language is one of increasing our understanding of cognitive processes in general because language is just one aspect of human cognition and cannot be properly understood apart from it. But the analyst to date seems to be deprived of integrated theories which relate linguistic structure and function to cognitive processes in general. In the expectation of an integrated theory which would relate linguistic (lexicogrammatical) structure and function to cognitive processes in general, one can attempt, through empirical investigation, to bring insight into processes involved in specific areas of language comprehension. The present study was aimed at shedding light on one aspect of the on-going mental processes involved in inferencing unknown meaning in reading in a FL, that aspect being the awareness of lexical relations of cohesion and coherence in reading comprehension.

In the light of Widdowson's (1984) account of the reading process which views the textual object as "schematically organised and so represents a structural order which the reader has to reconcile with his own" (p225), and which describes the reading process as an "act of assertion" and/or an "act of submission", it seems possible to describe reading in a foreign language along these lines and view it as an "act of assertion" and/or an "act of submission" depending on how the reader chooses to consider it.

With reference to the present study, the interpretation of lexical relations of cohesion and coherence may be defined in terms of these two readers' attitudes, viz. assertive and submissive. A priori it would seem that there are tendencies towards encouraging 'submission' or 'dependence' when linguistic knowledge is involved because lexical relations will key in closely with the reader's prediction⁹ and towards encouraging 'assertion' or 'dominance' when non-linguistic/pragmatic knowledge is involved in the interpretation of these lexical relations, the latter being relatively unpredictable. Thus, the interpretation of lexical relations of cohesion which involves some specific aspect of schematic knowledge, viz. linguistic knowledge seems to require the reader to recognise the writer's "territorial claim" to use Widdowson's image, and will seek to recover the underlying discourse from the textual clues provided by him (the writer). Therefore the reader will be dependent and will adjust to the writer's scheme in a submissive fashion when dealing with devices of cohesion. Recognising the writer's textual clues implies that the reader can accommodate the writer's conceptual scheme into the pattern of his own world and means that the reader will follow the text like a script.

On the other hand, the interpretation of lexical relations of coherence which typically appeals to non-linguistic/pragmatic knowledge seems to be an "act of assertion" which manifests itself in the reader's projection of his own scheme on what he reads and will

change the direction of accommodation so that the text, but not his pattern of things, is adjusted to fit the patterns of his own significance. This assertive attitude towards reading is often the result of the absence of textual clues. Then the reader will tend to be dominant and to assert the primacy of his own conceptual pattern. These two attitudes are presumably interactive in reading (comprehension) and their separation does not imply a sharp distinction between them. They may be best viewed on a continuum, and the successful/proficient FL reader would be defined as someone who keeps a balance between them, and the less able/less proficient FL reader would be situated in the two poles of the continuum as the diagram below shows:



4.3 Concluding remarks

We have been concerned in this chapter with the outcome of the empirical investigation undertaken with non-native readers of English. Although the 3-way interactions between the variables (experimental conditions, yeartime and language dominance) were non-significant, some of the 2-way interactions were significant. Thus, the difference in performance between the "yeartimes" was significant in all four experiments, and the difference between "conditions" was significant in three experiments out of

four. This led us to investigate further the reasons for such significance by conducting an item analysis. It appears that the results were often affected by learners' communicative rather than linguistic competence, and this seems to have had an effect across years.¹⁰ Assuming that the subjects have enjoyed "contextual" language teaching (vocabulary and syntax) at secondary level, and knowing that they still benefit from it at tertiary level (contextual language teaching is one of the principles of the 'modern' approach to EFL teaching), it does not seem to be the case that this type of teaching has served "as a valuable basis for the later development of communicative competence" (Criper and Widdowson, 1975: 210). Evidence suggests that some advanced learners are not yet aware of the complex relationships between form and function in general and in relation to vocabulary in a FL in particular.

Thus, the conclusions that emerge from the present study are both pedagogical and methodological as they may inspire future research in the field of FL reading comprehension and on cohesion, addressing in particular the linguistic features of written text. We have attempted to see HOW learners deal with unknown meaning in reading comprehension by examining the type of lexical context and context that surrounds unknown items and that was measured in terms of quality (the clue was immediate or distant; syntactic, paralexical or lexical; linguistic or non-linguistic) rather than quantity. To our knowledge no study has attempted to examine systematically

and qualitatively what constitutes a "clue" when inferencing unknown meaning. Yet linguists and non-linguists alike are unanimous in saying that the presence of "context" is relevant for the interpretation of utterances. Recent research in English as a FL (Khalil, 1985) has shown how the interpretation by a native speaker reader of an utterance produced by a non-native reader may depend on the linguistic and pragmatic "clues" surrounding utterances. The present study, the first to deal with inferencing the meaning of a cloze item as an experimental technique and in which inter-sentential connection of a lexical kind is treated as an independent variable, provides preliminary evidence that the quality of the lexical co-text and context influences FL learners' ability to recover the meaning of unknown vocabulary items. More research needs to be done on the quality rather than the quantity of what is globally referred to as "context", on the impact of other types of clues on reading comprehension and on reading generally. Finally, the use of the cloze technique for future experiments seems to be mostly appropriate for evaluating inferencing abilities as an index of reading comprehension in EFL.

Notes on Chapter 4

- 1 Thus, extraction of the basic semantic elements in order to permit a hypothesis as to meaning to be generated by the hearer/reader is only one perceptual strategy known to be universal and not language specific (see Bever, 1970, and Tarone, 1974, on perceptual strategies).
- 2 Strategy and procedure are used interchangeably in this study although Faerch and Kasper (1983, 1984b), for example, basing their usage of the terms on cognitive psychology, have used "procedure" to denote what one does in order to achieve a goal (in the most general terms) and "strategy" to refer to a particular subset of procedures as those employed for problem-solving.
- 3 This is one sample from the data obtained on a test of reading comprehension conducted with 48 undergraduate EFL students at the University of Algiers in December 1982. The testees were asked to arrange jumbled-up sentences 'into a coherent whole'. 'Coherent' was defined as 'making sense for them as well as the reader'.
- 4 Compare with original text: Mrs. Smith was upset and sobbed: "My aunt has been killed. She is in her bedroom. The knife is lying on the bed beside her". The policeman asked her how she discovered her aunt's body. Mrs. Smith replied she discovered the body when she went into the bedroom to show her aunt her new dress.

The policeman asked her who was in the house that night./
"Nobody was in the house" she said./Mrs. Smith sobbed
and said she was the only relative her aunt had./

- 5 The pilot study was conducted with 48 undergraduate EFL students at the University of Algiers in December 1982. The subjects were asked to write a 100 word "free composition" on one of the two topics: "Should women work?" and "Why do you save/not save money?".
- 6 Meara (1978) remarks that in the learner, the semantic arrangement of FL words seems to be less well-established than in the native speaker. What is implied is that whatever the number of vocabulary items a learner may possess, he may not be aware of their relatedness in the semantic system. In reading, sense relations are not automatically evoked in the learner's mind as they are in the native speaker .
- 7 See on this point Cruse (1975, 1977) who attempts to analyse lexical generality and specificity within a Gricean framework. Cruse (1977) points out that lexical generality and specificity underlying the use of hyponyms cannot always be explained by the Gricean principle of co-operation. For example (Cruse, 1977: 153):

(Said by someone A who is the owner of only one domestic animal - an Alsatian. Hearer B knows this)

- a. I think I shall take the dog for a walk.
- b. I think I shall take the animal for a walk.
- c. I think I shall take the Alsatian for a walk.

a. is a further specification of animal as a dog (because B, the hearer, knows that A, the speaker, owns a dog) and is therefore pragmatically redundant. So a. does not comply with Grice's Maxim of Quantity. But for Grice a. would comply with the Maxim of Quantity, but b. would violate it, because if B knows that A has a dog/an animal/an Alsatian, then A should not use the more general term 'animal' to refer to his dog. In doing so, b. is violating Grice's Quantity Maxim, that is, Be as informative as possible.

8 For example (Bramki and Williams, 1984: 176): "A large percentage of the human race still lives in very small self-sufficient peasant communities. These people experience great poverty, but they can provide, on an individual basis, for their own survival. They have a degree of economic independence. If we now turn to the inhabitants of New York, London, or any other metropolitan area, we must observe the opposite situation - a high standard of living together with an extreme economic dependence. The inhabitants of cities are totally incapable of providing for themselves, directly, the means of their survival".

9 The notion of "prediction" is related to Sanford and Garrod's (1981: 127) notion of scenarios: if a text

conforms to the reader's predictions, because he can recognise a scenario behind it, it is readily interpreted. If it does not, it will be more difficult to understand.

- 10 Cooper's (1984) research also suggests that as one goes up the levels of grammar and discourse the comprehension gap between learners of two different types of target language proficiency widens.

APPENDIX I

Transcription and pronunciation table

Arabic Alphabet	Transcription	Nearest (English) Equivalent	Point and Manner of Articulation
ب	b	<u>back</u>	bilabial stop
ت	t	French <u>thé</u>	dental stop
ث	θ	<u>thing</u>	dental fricative
ج	dz	<u>jump</u>	post-alveolar fricative
ح	ħ	/	pharyngeal fricative
خ	x	Scottish English <u>loch</u>	uvular fricative
د	d	French <u>dame</u>	dental stop
ذ	ð	<u>then</u>	dental fricative
ر	R	Spanish <u>Rio</u>	uvular rolled
ز	z	<u>zero</u>	alveolar fricative
س	s	<u>soon</u>	alveolar fricative
ش	ʃ	<u>fishing</u>	postalveolar fricative
ص	ʂ	/	(emphatic) alveolar fricative
ض	ʒ	/	(emphatic) dental fricative
ط	t̤	/	(emphatic) dental stop
ظ	t̤̥	/	more emphatic than ʒ dental fricative
ع	ɣ	/	velar fricative
غ	ʁ	French <u>robe</u>	uvular fricative

Arabic Alphabet	Transcription	Nearest (English) Equivalent	Point and Manner of Articulation
<p>ف ق ك ل م ن ه و ي</p>	<p>f q K l m n S w J</p>	<p><u>f</u>ew — <u>K</u>ey peo<u>l</u>e su<u>m</u> bu<u>n</u> hou<u>s</u>e <u>w</u>et <u>y</u>et</p>	<p>labiodental fricative uvular stop velar stop alveolar lateral bilateral nasal alveolar nasal pharyngeal fricative bilabial semi-vocalic palatal semi-vocalic</p>
<p>ا ا ا</p>	<p>a ə u i :</p>	<p><u>b</u>ad gi<u>r</u>l pu<u>t</u> pre<u>t</u>ty vowel length</p>	

**PAGE
MISSING
IN
ORIGINAL**

- 6 Some soldiers in the Falklands were reported to have suffered permanent foot damage. The soldiers' _____ clothing was inadequate: there was too great a proportion of nylon in their army socks. Doctors think that certain types of clothing can cause considerable harm.
- 7 Next time you have a few chicken livers to use up, remember that you can make delicious evening snacks with fried livers. Try to _____ them and put them on a piece of toasted brown bread with garlic.
- 8 Next time you have a few chicken livers to use up, try to _____ them. Remember that you can make delicious evening snacks with fried livers. Then put them on a piece of toasted brown bread with garlic.
- 9 Russian women have virtually all the responsibility for the children and the home, and this is used to discriminate against them at work. It is a form of _____ which they often discuss.
- 10 Russian women have virtually all the responsibility for the children and the home. This is a form of _____ which they often discuss for it is used to discriminate against them at work.
- 11 The Umayyad Caliphate in Spain was the greatest period of Al-Andalus. It was a civilisation based principally on the towns. The Muslims were primarily _____. The high achievement of Al-Andalus can be seen in Cordoba and Granada.

- 12 The Umayyad Caliphate in Spain was the greatest period of Al-Andalus. The Muslims were primarily _____. Their civilisation was based principally on the towns. The high achievement of Al-Andalus can be seen in Cordoba and Granada.
- 13 Many peoples of that period came into contact with the Greeks. The beginnings of Roman literature was worthy of admiration. However, only the _____ had the maturity and the imagination to assimilate and carry on the culture of their neighbours.
- 14 Many peoples of that period came into contact with the Greeks. However, only the _____ had the maturity and the imagination to assimilate and carry on the culture of their neighbours. The beginnings of Roman literature was worthy of admiration.
- 15 The lack of organisation in the Crimean war was appalling. Cholera and dysentery were widespread. Many soldiers died of _____. Mary was disturbed by the dreadful stories about the war which came back to her.
- 16 The lack of organisation in the Crimean war was appalling. Many soldiers died of _____. Cholera and dysentery were widespread. Mary was disturbed by the dreadful stories about the war which came back to her.

- 17 The Italian popular theatre is said to have arisen in Rome in the 2nd century BC. The first comic actor was recorded in the city in 211 BC. The _____ sought to imitate life and this theatrical genre may have had a certain social value.
- 18 The Italian popular theatre is said to have arisen in Rome in the 2nd century BC. The _____ sought to imitate life and the first comic actor was recorded in the city in 211 BC. This theatrical genre may have had a certain social value.
- 19 The greatest tragedy of the events at the Golden Temple will be felt by the people. These events will endanger the prospects for further growth in north west India. This is a _____ which had not been much on the news before.
- 20 The greatest tragedy of the events at the Golden Temple will be felt by the people. These events will endanger the prospects for further growth in a _____ which had not been much on the news before - north west India.

(Exp. 2)

21. A debate has raged about the control of the British Police. To the _____, such control should be assured by the superior members of their staff. To the critics, there should be locally elected committees responsible to the local community.
22. A debate has raged about the control of the British Police. To the critics, there should be locally elected committees responsible to the local community. To the _____, such control should be assured by the superior members of their staff.
23. We have organised some splendid cookery demonstrations. Several _____ you've watched being made will be on sale this afternoon and you will hear talks about wines from one of our Masters of Wine.
24. We have organised some splendid cookery demonstrations and you will hear talks about wines from one of our Masters of Wine. Several _____ you've watched being made will be on sale this afternoon.
25. For most of those on board, crossing the border represented an expedition into unknown territory. The _____ amused themselves by reading the difficult names. The train staff and the Austrian locomotive engineers were also excited by the border crossing.
26. For most of those on board, crossing the border represented an expedition into unknown territory. The train staff and the Austrian locomotive engineers were also excited by the border crossing. The _____ amused themselves by reading the difficult names.

- 27 There is nothing like a crisp and tasty salad with a sauce. So we have prepared the perfect _____ to meet your needs. Fresh homegrown vegetables are an important ingredient, too.
- 28 There is nothing like a crisp and tasty salad with a sauce, and fresh homegrown vegetables are an important ingredient too. So we have prepared the perfect _____ to meet your needs.
- 29 Thanks to his childhood memories of watching those primitive trains steaming off to distant places, he had some thoughts about making a fortune with railways. They were grandiose _____. He wanted to organise trains that would cross a whole continent.
- 30 Thanks to his childhood memories of watching those primitive trains steaming off to distant places, he had some thoughts about making a fortune with railways. He wanted to organise trains that would cross a whole continent. They were grandiose _____.
- 31 People were keeping an ear open for further noises such as sirens. Sometimes when the sound of a _____ began to die away, the heavy guns started to roar. Planes zigzagged over the city.
- 32 People were keeping an ear open for further noises such as sirens. Planes zigzagged over the city. Sometimes when the sound of a _____ began to die away, the heavy guns started to roar.

- 33 The structure of society in 19th c. Germany was changing rapidly. There were many social areas where change was quite _____ like material conditions, social relations and ideologies. A few areas were slower to change.
- 34 The structure of society in 19th c. Germany was changing rapidly. A few areas were slower to change. There were many social areas where change was quite _____, like material conditions, social relations and ideologies.
- 35 It is hard to believe that more babies are born with impaired or no hearing and that many _____ suffer some hearing loss as they get older. Young people and adults can become deaf through illness or injury.
- 36 It is hard to believe that more babies are born with impaired or no hearing. Young people and adults can become deaf through illness or injury, and many _____ suffer some hearing loss as they get older.
- 37 In the wet processing of coffee, the fresh fruit is pulped by a crushing machine. However, some _____ remains and this residue is removed by fermentation and washing in large containers. The coffee seed is usually dried by exposure to the sun.
- 38 In the wet processing of coffee, the fresh fruit is pulped by a crushing machine. The coffee seed is usually dried by exposure to the sun. However, some _____ remains and this residue is removed by fermentation and washing in large containers.

39 Little red flames were seen early in the morning. Then, after a quick rush of noise, _____ started to crack from the entire building. The windows of the kitchens were covered with thick smoke.

40 Little red flames were seen early in the morning. The windows of the kitchens were covered with thick smoke. Then, after a quick rush of noise, _____ started to crack from the entire building.

(Exp. 3)

- 41 The security forces who operated after the violence in Cairo and Alexandria had no instructions to protect the administration buildings and kill the civilians. However, the population was _____. After the events, the army was sent on to the streets.
- 42 The security forces who operated after the violence in Cairo and Alexandria had no instructions to protect the administration buildings and kill the civilians. The population was _____. After the events, the army was sent on to the streets.
- 43 The security forces who operated after the violence in Cairo and Alexandria had no instructions to protect the administration buildings and kill the civilians. Despite this fact, the population was _____. After the events, the army was sent on to the streets.
- 44 The visit of the Duke of Edinburgh to Armagh in Northern Ireland yesterday was _____. Therefore the Irish government in Dublin sent a formal protest about the incident to the British ambassador in Dublin. The Duke was to visit Dublin next month.
- 45 The visit of the Duke of Edinburgh to Armagh in Northern Ireland yesterday was _____. The Irish government in Dublin sent a formal protest about the incident to the British ambassador in Dublin. The Duke was to visit Dublin next month.

- 46 The visit of the Duke of Edinburgh to Armagh in Northern Ireland yesterday was _____. This resulted in the Irish government in Dublin sending a formal protest about the incident to the British ambassador in Dublin. The Duke was to visit Dublin next month.
- 47 In an interview with 'New Society', Anne Sorby reports on a common attitude among parents today: they do not accept that their children's lifestyle can change. Really, the parents have _____. Few only would recognise this fact.
- 48 In an interview with 'New Society', Anne Sorby reports on a common attitude among parents today: they do not accept that their children's lifestyle can change. The parents have _____. Few only would recognise this fact.
- 49 In an interview with 'New Society', Anne Sorby reports on a common attitude among parents today: they do not accept that their children's lifestyle can change. The truth is that the parents have _____. Few only would recognise this fact.
- 50 The Church in Europe and Islam in Arab countries play fundamental roles in their societies. 'Life below' and 'life beyond' are connected in western Europe. Yet they are _____ in many countries of the Muslim world.

- 51 The Church in Europe and Islam in Arab countries play fundamental roles in their societies. 'Life below' and 'life beyond' are connected in western Europe. They are _____ in many countries of the Muslim world.
- 52 The Church in Europe and Islam in Arab countries play fundamental roles in their societies. 'Life below' and 'life beyond' are connected in western Europe. This is not true of many countries of the Muslim world where they are _____.
- 53 The use of contraceptives by couples of reproductive age in developed and less developed areas can slow population growth. Besides it may result in _____. So it is urgent that governments should give more money for research on the effects of the pill.
- 54 The use of contraceptives by couples of reproductive age in developed and less developed areas can slow population growth. It may result in _____. So it is urgent that governments should give more money for research on the effects of the pill.
- 55 The use of contraceptives by couples of reproductive age in developed and less developed areas can slow population growth. It may also be added that it may result in _____. So it is urgent that governments should give more money for research on the effects of the pill.

- 56 Jose de Molina says that Argentine writers should adhere to the tradition of Spanish literature. But I say that Argentine literature can be defined as a desire to become _____ Spain. The search for European themes is a well-known phenomenon in 20th c. literature.
- 57 Jose de Molina says that Argentine writers should adhere to the tradition of Spanish literature. I say that Argentine literature can be defined as a desire to become _____ Spain. The search for European themes is a well-known phenomenon in 20th c. literature.
- 58 Jose de Molina says that Argentine writers should adhere to the tradition of Spanish literature. This disagrees with my definition of Argentine literature as a desire to become _____ Spain. The search for European themes is a well-known phenomenon in 20th c. literature.
- 59 Two sociologists at the University of Illinois have argued that an open-plan office with lots of people in it can create a friendly environment. Furthermore it can _____ frustration. But many office workers who were asked disagree, according to a recent study.
- 60 Two sociologists at the University of Illinois have argued that an open-plan office with lots of people in it can create a friendly environment. It can _____ frustration. But many office workers who were asked disagree, according to a recent study.

- 61 Two sociologists at the University of Illinois have argued that an open-plan office with lots of people in it can create a friendly environment. It may also be added that it can _____ frustration. But many office workers who were asked disagree, according to a recent study.
- 62 The Macedonian capital, Skopjé, is fertile, industrialised, with a higher percentage of university students than in almost any other part of the country. However, economists regard it as a _____ region in the Balkans. It has unemployment problems.
- 63 The Macedonian capital, Skopjé, is fertile, industrialised, with a higher percentage of university students than in almost any other part of the country. Economists regard it as a _____ region in the Balkans. It has unemployment problems.
- 64 The Macedonian capital, Skopjé, is fertile, industrialised, with a higher percentage of university students than in almost any other part of the country. By contrast, economists regard it as a _____ region in the Balkans. It has unemployment problems.
- 65 In many rural areas of the 13th century, you can find the town hall in the market place. However, this type of location was usually reserved for _____. People had to guard against the enemy.
- 66 In many rural areas of the 13th century, you can find the town hall in the market place. This type of location was usually reserved for _____. People had to guard against the enemy.

- 67 In many rural areas of the 13th century, you can find the town hall in the market place. By contrast, this type of location was usually reserved for _____. People had to guard against the enemy.
- 68 Some writers have always criticised John Steinbeck's novels because they _____ his style. John Steinbeck's writings were neither like Kafka's nor like Becket's, and his novels were easier to read.
- 69 Some writers have always criticised John Steinbeck's novels. They _____ his style. John Steinbeck's writings were neither like Kafka's nor like Becket's, and his novels were easier to read.
- 70 Some writers have always criticised John Steinbeck's novels. The reason is that they _____ his style. John Steinbeck's writings were neither like Kafka's nor like Becket's, and his novels were easier to read.
- 71 The presence of a chronically ill child can have a profound effect on all members of the family. It is widely assumed that the other children in the family often develop _____ because much of the parents' attention is directed towards the ill child.
- 72 The presence of a chronically ill child can have a profound effect on all members of the family. It is widely assumed that the other children in the family often develop _____. Much of the parents' attention is directed towards the ill child.

- 73 The presence of a chronically ill child can have a profound effect on all members of the family. It is widely assumed that the other children in the family often develop _____. The reason is that much of the parents' attention is directed towards the ill child.
- 74 Food and drink can be consumed at shop prices in this restaurant, or you can take them home. Mr. Torrino's new invention is _____ office workers because it is open throughout the lunch period.
- 75 Food and drink can be consumed at shop prices in this restaurant, or you can take them home. Mr. Torrino's invention is _____ office workers. It is open throughout the lunch period.
- 76 Food and drink can be consumed at shop prices in this restaurant, or you can take them home. Mr. Torrino's new invention is _____ office workers. This can be explained by the fact that it is open throughout the lunch period.

(Exp. 4)

- 77 The World Health Organisation reports that most major disasters happen in the tropics to those who live in bad shelters on dangerous ground. Disasters are getting bigger and more frequent and the _____ inevitably die when earthquake and cyclone strike.
- 78 The World Health Organisation reports that most major disasters happen in the tropics to those with no money. Disasters are getting bigger and more frequent and the _____ inevitably die when earthquake and cyclone strike.
- 79 Legends and myths are part of the culture and heritage of the home, and are in general invented by the elder relatives in the family. They mould a child's life and teach him the principles of life. Legends are mostly created by the _____.
- 80 Legends and myths are part of the culture and heritage of the home and are in general invented by those who have special relationships with children. Legends are mostly created by the _____.
- 81 Egypt now imports half of its food. The density of its population in towns is amongst the highest in the world, and will probably double by the next twenty years. Centuries ago, it was a veritable _____.
- 82 Egypt has to face many problems. The density of its population in towns is amongst the highest in the world, and will probably double by the next twenty years. Centuries ago, it was a veritable _____.

83 The international games are back again, but they should be shown at convenient times. I also think that the BBC should do more to explain all those confusing _____ on our TV screens.

84 Our favourite events are back again after four years, but they should be shown at convenient times. I also think that the BBC should do more to explain all those confusing _____ on our TV screens.

85 The Roman Livius Andronicus translated the Odyssey for use in schools because he wanted better educational methods. For the next hundred years that followed, epic and drama remained the main concern of the _____ poets such as him.

86 The poet Livius Andronicus translated the Odyssey for use in schools because he wanted better educational methods. For the next hundred years that followed, epic and drama remained the main concern of the _____ poets such as him.

87 Some people are capable of vandalizing their country, transforming it into a place without history or beauty. They live in the immediate present and are unaware of historical continuity and without culture. _____ today is a widespread social phenomenon.

88 Some people are capable of treating their country the way some teenagers today treat buses and phoneboxes. They live in the immediate present and are unaware of historical continuity and without culture. _____ today is a widespread social phenomenon.

- 89 A British journal reports that 62% of the people who return from their holiday abroad complain about something. One of their common complaints is about _____. These stinging insects seem to be always out to get them and are often the cause of broken romances and marriages.
- 90 A British journal reports that 62% of the people who return from their holiday abroad complain about something. One of their common complaints is about _____. Those annoying little creatures seem to be always out to get them, and are often the cause of broken romances and marriages.
- 91 Sheep first appeared in the Ice Ages. But they can adapt to climates ranging from hot deserts to the regions of the Arctic. Originally, in those _____ places, they were monstrous creatures, as large as oxen.
- 92 Sheep needed their thick coats when they first appeared. But they can adapt to climates ranging from hot deserts to the regions of the Arctic. Originally, in those _____ places, they were monstrous creatures, as large as oxen.
- 93 A change in hours may mean that the time has come for working women to acquire different habits. It is believed that a new pattern of hours will suit them as well as their _____.

94 A change in hours may mean that the time has come for women who are not at home all day to acquire different habits. It is believed that a new pattern of hours will suit them as well as their _____.

95 In a South African camp. Leah came out of a hut she shared with her widowed mother and a sister.

'Does your mother get a grant or a pension?'

'No'

'What do you do for food?'

'We _____ food'

'Have you ever returned any of the food people lend you?'

'No'.

96 In a South African camp. Leah came out of a hut she shared with her widowed mother and a sister.

'Does your mother get a grant or a pension?'

'No'

'What do you do for food?'

'We _____ food'

'Have you ever thought they may need the food themselves?'

'No'.

APPENDIX III

Sample of miscue inventory: Table of Expected Responses (C) and Observed Responses (NC and W)

Texts	C responses	NC responses	W responses
1	accident	victim-death-case-of drowning-and incident- bad news-tragedy	year-swim-cave- season-report-time- committee-was
2	accident	bad news-incident- victim-tragedy	man-time-game-season- just after
3	disintegrate	die-disappear-rot- become useless	dry-get dry-get oily- are oily-have oil- become dusty-catch fire-burn-cannot sell- are replaced-are left- eventually
4	disintegrate	(ibidem)	sell-naturally-burn
5	synthetic	nylon-military-foot- type of	Falklands-socks- dangerous-suffering- damage-new-harmful-
6	synthetic	(ibidem)	socks-reported-worn out
7	fry	cook-grill-recycle- toast	degustate-sell-taste- remember-make-crush- use-cut-keep-throw- prepare
8	fry	(ibidem)	keep-remember-take care of-make-snoek- mix-cut
9	inequality	discrimination- injustice-racist- oppression-domination- slavery-sexism-social division-apartheid	responsibility-work- Russian system-virtue- communism-class struggle-handicap- life-argument-
10	inequality	(ibidem)	tradition-discipline- discussion-emancipa- tion
11	townspeople	rulers-powerful- rulers-not nomads- architects	civilized-Spanish- merchants-at its basis- great-Andalusians-in the South-in Cordoba- from Cordoba-in Andalusia-from Granada- in Granada-visitors- nomads-countrymen

Texts	C responses	NC responses	W responses
12	townspeople	merchants-architects- great achievers- powerful	from Cordoba-from Granada-great-caliphs- civilized-Andalusians- believers-warriors
13	Romans	literate-bright-sages- enlightened-minority- intellectuals-Italians	Greeks-people-Turks- beginners-Arabs-older- Europeans-rich-Moslem- Umayyad
14	Romans	Italians	intelligentsia- educated-rich- Egyptians-cultured- admirers-Moslems
15	diseases	illnesses-cholera- hunger-epidemics-lack of hygiene-dysentry- starvation-the plague- malaria-feed poisoning- infections	crime-war
16	diseases	(ibidem)	war-dreadful stories
17	comedy	comic theatre-popular theatre-plays-actors- playwrights-Romans- Italians-dramatists	Comedia dell Arte- followers-second- tragedy-social value- other-aim
18	comedy	dramatic theatre-actors- plays-Romans-Italians- dramatists-playwrights-	second-musicians- tragedy-genre- history
19	region	place-country-tragedy- story-event-news item	prospect-danger- temple-war
20	region	place-country	tragedy-century- thing-disaster
21	police	authorities-government- ministers-police officers-state	British-committees- journalists- solution
22	police	(ibidem)	elected-journalists- local community- British
23	recipes	cakes-dishes-products- samples	things-wines
24	recipes	(ibidem)	wines-books-plants objects

Texts	C responses	NC responses	W responses
25	passengers	people-children-staff- engineers-crew- explorers-tourists	customs officers
26	passengers	people-train staff- engineers-children- crew-explorers- tourists	foreigners-unknown- customs officers- Austrians
27	sauce	mixture-thing-recipe- food-one	salad
28	sauce	thing-food-recipe- mixture-salad	ingredient-vegetable- taste
29	thoughts	ideas-plans-dreams- trains-railways- inventions-initiatives- indeed-but failed_and beautiful-	memories-fortunes
30	thoughts	(ibidem)	toys-boys-memories- places
31	siren	rifle-fire machine- bomb-bazuka	plane-train-firebell
32	siren	rifle-fire machine- bomb-plane-bazuka	train-ambulance-man
33	rapid	fast-perceptible- noticeable-obvious- remarkable-clean- apparent-visible	difficult-material- ideological-slow
34	rapid	(ibidem)	slow-few-rich
35	babies	others-young people- children	people
36	babies	people	old people
37	pulp	fruit-fresh fruit- still	residue-coffee-seed- crumbs
38	pulp	fruit-fresh fruit	coffee-seed-residue
39	flames	fire-the roof-wood- timber-walls-glass debris	windows-smoke
40	flames	fire-the roof-walls- windows-timber-glass debris-gas bottles- gas pipes	smoke

Texts	C responses	NC responses	W responses
41	massacred	murdered-killed-exterminated	violent-mobilized-calm-quiet-arrested-shocked-out of control-resisting-demonstrating-triumphant
42	massacred	(ibidem)	terrified-resisting-arrested-not
43	massacred	(ibidem)	not affected
44	unexpected	unplanned-not scheduled-a catastrophe-a disaster-unwelcomed	cancelled-postponed-short missed-formal-announced
45	unexpected	(ibidem)	cancelled-short-late-incident-al-agreeable
46	unexpected	(ibidem)	short-formal
47	changed	prejudices-different mentalities-conservative ideas	denied it- problems-no authority-rejected it-objected-suffered-not recognized this
48	changed	prejudices- no authority-retrogressed-conservative ideas-not recognized this-objected to this	nostalgia-stubborn minds-accepted
49	changed	(ibidem)	new ideas-accepted
50	separated	independent-divided-unconnected	not recognized-together-problematic-social problems-complementary
51	separated	(ibidem)	different- not obeyed-not recognized-fundamental
52	separated	(ibidem)	together-practiced
53	cancer diseases	illnesses-dangerous illnesses-sterility-infertility-poor health	less money-effects-more children-depopulation-old age-divorce-growth-contraception-no population
54	cancer diseases	(ibidem)	less money-less people-underdevelopment-no result-more research
55	cancer diseases	(ibidem)	more children-more research-twins-small babies

Texts	C responses	NC responses	W responses
56	separated from	not part of-not like-different from	with-like-in-famous in- literature of-against- another
57	separated from	(ibidem)	one of-with-another- common in-really- wellknown in-European in
58	separated from	(ibidem)	really-researched in- like-wellknown in
59	eliminate	prevent-stop	provoke-give-create-be
60	eliminate	(ibidem)	develop-cause-give- study
61	eliminate	(ibidem)	agree with-add-bring- create
62	backward	underdeveloped- poor-jobless- difficult	university-small-holy- developed
63	backward	(ibidem)	rich-industrial-small- productive-strategic- Saharian
64	backward	(ibidem)	expanding-agrarian
65	watchtowers	the army-guards- soldiers-defence- protection	enemies-the church-the shops-them-foreigners- the rich-writers
66	watchtowers	(ibidem)	the poor- the enemy- churches-wars-them
67	watchtowers	(ibidem)	nonrural areas-everyone- merchants
68	hated	disliked-disagreed with-did not understand	copied-read
69	hated	(ibidem)	admired-used-defended- developed-ignored- wrote in-understood- did not know
70	hated	(ibidem)	read-did not know
71	problems	illnesses-complexes- jealousy	effects-this idea- healthily-attention
72	problems	(ibidem)	quarrels-well-better
73	problems	(ibidem)	rapidly

Texts	C responses	NC responses	W responses
74	successful with	especially for-welcomed by-good for	not for-a shop for-run by
75	successful with	(ibidem)	shut for-not for- next to-a shop for-poen to
76	successful with	(ibidem)	fact for-expensive for
77	poor	miserable-people-resourceclass-underprivileged	animals-women
78	poor	(ibidem)	crops-trees
79	grandmothers	grandparents-elders-family-ancestors-parents	teachers-ignorants-priests-friends-imagination-
80	grandmothers	(ibidem)	other children-uncles-society-culture-teachers-historians-time-past
81	granary	power-paradise-heaven-Goldmine-land of opportunities-Eden-wonderland	country-desert-problem
82	granary	(ibidem)	kingdom of Faroes-disaster-overpopulated place-colony-tourist place-museum
83	sports	games-sporting events-results-programmes-matches-competitions	power cuts-times-waves-
84	sports	(ibidem)	films-people-times-thoughts-discussions-political debates-noises
85	Roman	famous-great-wellknown-dramatic	new-lonely
86	Roman	(ibidem)	modern-educated-ancient-translator

texts	C responses	W responses	W responses
87	vandalism	destruction- carelessness	illiteracy-society
88	vandalism	(ibidem)	the phone-modern life- continuity-history-so- teenagers-the present
89	mosquitoes	insects	broken marriages-food- foreigners-hotels
90	mosquitoes	insects	women-dogs-cats-money- conditions-diseases-dirt
91	freezing	Arctic-icy-cold- remote-frightful	strange
92	freezing	(ibidem)	desertic-hot-new-different- two-historical
93	employer	husbands-families- children-jobs	customs-beliefs-choices
94	employer	(ibidem)	habits-holidays-home-life- likings-clothes
95	borrow	are lent-receive- are given-ask for	do nothing for-eat-buy- have no-lack-don't need- prepare-save up

APPENDIX VI

RESULTS OF STATISTICAL ANALYSIS

Analysis of variance

Experiment 1

Tests of significance for C1 using sequential sums of squares

Source of variation	Sum of squares	DF	Mean Square	F	Sig. of F
Within cells	623.75000	66	9.45076		
Constant	8070.02778	1	8070.02778	853.90274	0.0
Yeartim	83.84722	2	41.92361	4.43601	.016
AF	20.25000	1	20.25000	2.14269	.148
Yeartim by AF	30.12500	2	15.06250	1.59379	.211

Tests of significance for C2 using sequential sums of squares

Source of variation	Sum of squares	DF	Mean Square	F	Sig. of F
Within cells	298.25000	66	4.51894		
Conditio	36.00000	1	36.00000	7.96647	.006
Yeartim and Conditio	5.79167	2	2.89583	.64082	.530
AF and Conditio	21.77778	1	21.77778	4.81922	.032
Yeartim by AF and Conditio	2.18056	2	1.09028	.24127	.786

Experiment 2

Tests of significance for C1 using sequential sums of squares

Source of variation	Sum of squares	DF	Mean Square	F	Sig. of F
Within cells	370.87500	66	5.61932		
Constant	11007.50694	1	11007.50694	1958.86878	0.0
Yeartim	68.72222	2	34.36111	6.11482	.004
AF	.34028	1	.34028	.06055	.806
Yeartim by AF	22.05556	2	11.02778	1.96248	.149

Tests of significance for C2 using sequential sums of squares

Source of variation	Sum of squares	DF	Mean Square	F	Sig. of F
Within cells	293.208333	66	4.44255		
Conditio	45.56250	1	45.56250	10.25593	.002
Yeartim and Conditio	15.50000	2	7.75000	1.74449	.183
AF and Conditio	5.06250	1	5.06250	1.13955	.290
Yeartim by AF and Conditio	.16667	2	.08333	.01876	.981

Experiment 3

Tests of significance for C1 using sequential sums of squares

Source of variation	Sum of squares	DF	Mean Square	F	Sig. of F
Within cells	463.05556	66	7.01599		
Constant	5932.51852	1	5932.51852	845.57073	0.0
Yeartim	92.23148	2	46.11574	6.57295	.002
AF	8.16667	1	8.16667	1.16401	.285
Yeartim by AF	12.69444	2	6.34722	.90468	.410

Tests of significance for within cells using sequential sums of squares

Source of variation	Wilks Lambda	Lambda	Sig. of F	Averaged F	Sig. of F
	Approx	Mult F			
Conditio	.95043	1.69490	.192	2.14786	.121
Yeartim and Conditio	.90925	1.58326	.183	2.06607	.089
AF and Conditio	.98770	.40458	.669	.48308	.618
Yeartim by AF and Conditio	.98215	.29406	.881	.30160	.876

Experiment 4

Tests of significance for C1 using sequential sums of squares

Source of variation	Sum of squares	DF	Mean Square	F	Sig. of F
Within cells	494.54167	66	7.49306		
Constant	6601.56250	1	6601.56250	881.02410	0.0
Yeartim	38.16667	2	19.08333	2.54680	.086
AF	5.84028	1	5.84028	.77943	.381
Yeartim by AF	25.38889	2	12.69444	1.69416	.192

Tests of significance for C2 using sequential sums of squares

Source of variation	Sum of squares	DF	Mean Square	F	Sig. of F
Within cells	644.70833	66	9.76831		
Conditio	52.56250	1	52.56250	5.38092	.023
Yeartim and Conditio	18.66667	2	9.33333	.95547	.390
AF and Conditio	3.06250	1	3.06250	.31351	.577
Yeartim by AF and Conditio	18.50000	2	9.25000	.94694	.393

APPENDIX VII

Sample of students' verbal protocols on Pilot Test S = Subjects

- S7 "Feathers can ... time, duration ... good quality ... some firms offer cushions ... good quality foam with a ten year guarantee will last longer ... good quality ... yes, it's in the text ... value for money ... duration ... can last longer, set A ... yes set A, feathers can last longer" (T1).
- S14 "... it's about alcoholic drinks ... it is wine ... because wine has been repeated three times - so it's wine ... people drink wine to accompany the food ... their dinner. Set A" (T3).
- S2 "Yes, it is demonstrations ... about how to make wine ... so it's set B ... set of instructions, how to make wine ... you must watch the instructions ... and ... then, you can try it, the wine" (T3).
- S10 "It puts an end to bending and ... place ... position ... centre ... localise ... yes you bend because ... heavy dishes from a low level ... and you ... she puts them in the centre of the oven ... inside the oven ... that is ... localise ... find room for them ... the dishes are heavy ... she's on her knees ... perhaps ... and puts the dishes ... right in the centre ... maybe the dishes are on the floor ... she has to bend ... to low level ... life is easy yes ... it is set C ... set C" (T2).

S12 "... makes life easier as it puts an end to bending and ... preparing food ... preparing heavy dishes from a low level ... I choose set B ... because she's bending and preparing heavy dishes ... on the floor ... low level on the floor, set B ... an oven ... at a convenient level ... makes life easier than ... preparing dishes on the floor ... ease of operation" (T3).

S7 "... Claire had hoped the ... conversation ... they sat down ... Claire felt unsatisfied ... she's disturbed by the subject ... it's an annoying subject ... about marriage ... she doesn't feel at ease to talk about marriage ... husband ... so set A is good ... women are not at ease with this subject ... they will not pursue the subject ... set A" (T8).

S1 "... at the bright bars of her empty ... mistress of the room ... of her empty, empty furniture ... with concern ... no - of her empty ... we put room because ... we have room here and room twice in the text ... it's set B, the word is room" (T6).

S8 "... Tahitians ... don't ignore their traditions ... set A ... they mix the old life with modern life ... the past ... set B ... adoption, acceptance ... no, not suitable ... accept is wrong because it's opposite meaning ... in the sentence we have but, but will absorb ... it's set A ... refuse, reject" (T9).

S14 "... conversation would continue ... a little longer ... continue ... which conversation ... I'm not sure of set A, not marriage, not really, no ... hadn't pursued the subject ... They've changed the subject ... talked about something else. I'd say same conversation, I'm not sure, same is not in the sets ... I don't know, but it is same, same conversation" (T8).

APPENDIX VIII

PILOT TEST

Instructions to students

You have been given 10 passages. Each passage contains ONE missing word (or blank) and is followed by 3 sets A, B, and C, with some of the features or characteristics of the missing word. What you have to do is:

1 Look at the features of set A, set B and set C in order to reconstruct the meaning of the missing word.

2 Then decide whether set A, set B or set C fits in the blank in the best way.

3 As you work, TELL how you are getting your answer. Just think out loud as you solve the problem. You may think in English, Arabic or French.

Here is an example to help you:

House plants that have enjoyed outside sunshine and fresh air during _____ should be brought indoors. As geraniums become more and more expensive to buy each year, it is even more vital to keep them safely through the winter.

- A
 - period of year
 - Good weather
 - vacation
 - no work
 - laziness
 - seaside
- B
 - season
 - heat
 - sunshine
 - ideal for plants
 - seaside
 - laziness
- C
 - house
 - bad condition
 - repairs
 - painting
 - decorating
 - workers

ANSWER: the set of features that fit the blank is: set B or set A (and the word which could be recovered is: summer or the hot season)

TEXT No.1

Some firms offer cushions with a choice of natural or synthetic fillings. Good quality foam with a ten-year guarantee will last longer than a cheap feather filling. Feathers can _____, the oil dries out, leaving them dry and brittle and eventually they turn to dust.

- A
 - time
 - duration
 - good quality
 - value for money
- B
 - separation into parts
 - breaking, crumbling
 - damage
 - poor quality
- C
 - wings
 - feathers
 - air, wind
 - lightness

TEXT No.2

Separating hob and oven gives you several advantages over a free-standing cooker. An oven at a convenient level makes life easier as it puts an end to bending and _____ heavy dishes from a low level. Here we focus on built-in ovens and test their cooking performance and ease of operation.

- A
 - raising from one level
 - higher level
 - use of hands
 - bending
 - heaviness
- B
 - preparing food
 - eating
 - use of heat
 - oven, fire
 - dishes
- C
 - place, position
 - limited space
 - centre
 - localise
 - find room for something

TEXT No.3

We have organised some splendid cookery demonstrations by the Dutch Dairy Bureau followed by talks about wines from one of the PD's Masters of Wine. Then, putting theory into practice, you'll have a chance of trying at least two of the _____ you've watched being made, and sipping with expert guidance and discussion four beautiful wines to accompany your three-course lunch or dinner.

- A
 - alcoholic drink
 - made from grapes
 - glasses
 - dinner, lunch
- B
 - set of instructions
 - cooking
 - food
 - dishes
 - lunch, dinner, parties
- C
 - food
 - meal
 - cooking
 - daytime
 - guests

TEXT No.4

Perhaps it is the innate good sense and good humour of the Tahitians which have insulated them from some of the more damaging pressures of the twentieth-century lifestyle. They do not readily _____ practices and methods which have served them well for generations, but will absorb the new with the old and make both work in harmony.

- A
 - rejection
 - no acceptance
 - no agreement
 - refusal
- B
 - adoption
 - taking in
 - acceptance
 - agreement
- C
 - choice, preference
 - suitability
 - special treatment

TEXT No.5

In her first glow of pride, Diane looked at her watch. Hardly fifteen minutes had passed since she had come into the house. How had she done it? She held the top end of the basket up and pressed the _____ of her right hand against its floor. It gave out firm-sounding squeaks. It had sprung in it.

- A
 - B
 - C
- object
 - made of metal
 - sewing
 - thread
 - tailor
 - sharp
 - part of body
 - in hand
 - inner surface
 - between finger and wrist
 - flat
 - object
 - made of metal
 - up and down
 - movement
 - elasticity

TEXT No.6

Adele, quite surprised, endeavoured to show herself mistress of the room by her civilities, and looked at the bright bars of her empty _____ with concern. 'Thank you, I am quite warm, very warm', said Mrs M. 'Allow me to stay here a little while, and do have the goodness to hear me my third act. I have brought my book'.

- A
 - B
 - C
- object
 - made of iron or brass
 - in fire place
 - bars
 - shiny
 - coal or wood
 - warmth
 - place
 - door
 - windows
 - furniture
 - carpets
 - fireplace
 - warmth
 - object
 - made of wood
 - four legs
 - flat top
 - write on
 - read
 - book
 - sit

TEXT No.7

You can imagine what a lovely, surprise gift that a _____ would make. Especially for a relative who lives a long way away, or who has emigrated or whom you simply don't see very often.

- A
 - B
 - C
- travelling
 - reservation, buying
 - air
 - railway
 - theatre
 - television
 - programmes, films
 - recording
 - magnetic tape
 - rebroadcasting
 - object
 - metal
 - vehicle
 - travelling,
 - riding

TEXT No.8

Then they talked about something else. The cost of heating, if they used their cottages at all in the coming winter. Claire had hoped the _____ conversation would continue a little longer. Around rounds of drinks, while she put their cold supper on the table. They sat down. Claire felt unsatisfied. Or was it a sense of disturbance, just because they hadn't pursued the subject?

- A
 - B
 - C
- marriage
 - husband
 - wife
 - ceremony
 - entertainment, party
 - container, basket
 - platted wigs
 - weaving
 - human beings
 - young
 - family
 - father
 - mother

TEXT No.9

No one knows as much about pasta as the _____. There is evidence that it was a revered dish as early as 5000 BC and its popularity is certainly not decreasing. Not unaturally therefore, when we decided to start selling our own range of pasta we knew exactly where to go.

- A
 - B
 - C
- human beings
 - cooking
 - good food
 - human beings
 - Algeria
 - cooking
 - good food
 - pasta (macaroni, spaghetti)
 - human beings
 - Italy
 - cooking
 - good food
 - pasta (macaroni, spaghetti)

TEXT No.10

People want television programmes that help them with difficulties in their lives, or help them grow in the enjoyment of their lives. They want programmes that'll show them how to cope with the _____, or get more out of their gardens. In other words, programmes must leave them with something and not simply wash over them.

- A
 - B
 - C
- human being
 - collecting
 - money, taxes
 - state
 - difficulties
 - worries
 - puzzles
 - problematic
 - solutions
 - entertainment
 - television
 - radio
 - arrangement
 - programming

Appendix IX: Sample of A and F students' responses on Placement Test.

	I	II	III	IV	V	VI	VII	
	<p>١- كوكبة ٢- حمارين ٣- تشبينة ٤- الرنين</p>	<p>١- الرنين ٢- المشهور ٣- الصقار ٤- الشترن</p>	<p>١- المريض ٢- السرورية ٣- الرطب ٤- الشائعة</p>	<p>١- الطبخ ٢- الأواني ٣- الإنجيل</p>	<p>١- الحركة ٢- اللقود ٣- الكسوف</p>	<p>١- الزيت ٢- الموترات</p>	<p>١- الرجال ٢- التلفزة ٣- الألات</p>	<p>١- السموم ٢- الزنابق ٣- الرضعة ٤- الرض ٥- الطلبة ٦- النقاد ٣- الصحافيات</p>
	<p>East Hill : but , bus-ya, you , arbil, god , hails , even , Camp - franc , results .</p>	<p>hospital : infirmière , médecin , médicament , soignée , ambulance , démographie , statistique</p>	<p>Arnie : plat , bouteille , bouteille , nourriture , plat , table . Ushim : cuisine , pain , pain-bon , viande , choucroute , viande , Ame .</p>	<p>politique : intervention , extérieur , sénateur , député , président , maire ,</p>	<p>Visite T.V : Télévision , écran , pièces , électronique , Télé cathodique ,</p>	<p>Armed : acier , veste , paine , argent , muelle , communication université : étudiants , étudiants , université , étudiants , étudiants salle , Ample ,</p>	<p>Livings : film , écran , scénario , dialogue , images , montage production .</p>	<p>Voysage : bagages , été , avion , train , voiture , table .</p>

BIBLIOGRAPHY

- ALDERSON, J.C. (1979). The cloze procedure and proficiency in English as a foreign language. Tesol, Vol.13, No.2, pp219-227.
- ALDERSON, J.C. and URQUHART, A.H. (eds) (1984). Reading in a Foreign Language. Longman, London.
- ALLEN, J.P.B. and WIDDOWSON, H.G. (1975). Grammar and Language Teaching. In Edinburgh Course in Applied Linguistics, Vol.2, pp45-97; ed. by JPB Allen and SP Corder, Oxford University Press, London.
- ANDERSON, R.C., REYNOLDS, R.E., SCHALLERT, D.L. and GOETZ, E.T. (1977). Frameworks for comprehending discourse. American Educational Research Journal, Vol.14, No.4, pp367-381.
- ANTHONY, E.M. (1975). Lexicon and Vocabulary: some theoretical and pedagogical aspects of word meaning. Regional English Language Centre Journal, Vol.6, No.1, pp21-29.
- ARONOWITZ, R. (1984). Reading Tests as Texts. In Coherence in spoken and written discourse, ed. by D. Tannen. Advances in Discourse Processes Series, Vol.XII, pp245-264.
- BACHMAN, L.F. (1982). The trait structure of cloze test scores. Tesol, Vol.16, No.1, pp61-70.
- BALTRA, A. (1983). Learning how to cope with reading English for academic purposes in 26 hours. Reading in a Foreign Language, Vol.1, No.1, pp20-34.

- BARTLETT, F.C. (1932). Remembering: A study in Experimental and Social Psychology. CUP, Cambridge, England.
- DE BEAUGRANDE, R. (1980). Text, Discourse and Process: Toward a multidisciplinary science of texts. Longman, London.
- DE BEAUGRANDE, R. and DRESSLER, W. (1981). Introduction to Text Linguistics. Longman, London.
- BEARDSLEY, G. (1982). Context cues in early reading. Journal of Research in Reading, Vol.5, No.2, pp101-112.
- BELLERT, I. (1970). On a condition of the coherence of texts. Semiotica, Vol.2, pp335-363.
- BENSOUSSAN, M. and LAUFER, B. (1984). Lexical guessing in context in EFL reading comprehension. Journal of Research in Reading, Vol.7, No.1, pp15-32.
- BEVER, T.G. (1970). The cognitive basis for linguistic structures. In Cognition and the development of language, ed. by J. Hayes. John Wiley and Sons, New York, pp279-369.
- BRAMKI, D. and WILLIAMS, R.C. (1984). Lexical familiarisation in economics text and its pedagogic implications in reading comprehension. Reading in a Foreign Language, Vol.2, No.1, pp169-181.
- BRANSFORD, J.D. and FRANKS, J. (1971). The abstraction of linguistic ideas. Cognitive Psychology, Vol.2, pp331-350.
- BRANSFORD, J.D., BARCLAY, J.R. and FRANKS, J. (1972). Sentence memory: A constructive versus interpretative approach. Cognitive Psychology, Vol.3, pp193-209.
- BROWN, G. and YULE, G. (1983). Discourse Analysis. Cambridge University Press, Cambridge, England.

- CARRELL, P.L. (1982). Cohesion is not coherence. Tesol, Vol.16, No.4, pp479-488.
- CARRELL, P.L. (1983). Some issues in studying the role of schemata or background knowledge in SL comprehension. Reading in a Foreign Language, Vol.1, No.2, pp81-92.
- CARRELL, P.L. (1984). Inferencing in ESL: Presuppositions and implications of factive and implicative predicates. Language Learning, Vol.34, No.1, pp1-21.
- CARTON, A.S. (1971). Inferencing: a process in using and learning language. In Psycholinguistics and Second Language Learning, ed. by P. Pimsleur and T. Quinn. CUP, Cambridge, England, pp45-58.
- CHAFE, W.L. (1972). Discourse structure and human knowledge. In Language Comprehension and the acquisition of knowledge, ed. by J.B. Carroll and R.O. Freedle. Wiley, Washington, pp41-70.
- CHAPMAN, L.J. and STOKES, A. (1980). Developmental trends in the perception of textual cohesion. In Processing of Visible Language, ed. by P.A. Kolers, M.E. Wrolstad and H. Bouma, Vol.2. Plenum, New York, pp
- CHAROLLES, M. (1978). Introduction aux problèmes de la cohérence des textes. Approche théorique et étude des pratiques pédagogiques. Langue Française, No.38, pp7-41.
- CHOMSKY, N. (1965). Aspects of the Theory of Syntax. MIT Press, Cambridge, Massachusetts.

- CLARK, H.H. (1977). Inferences in comprehension. In Basic Processes in Reading: Perception and Comprehension, ed. by D. Laberge and S.J. Samuels. Lawrence Erlbaum Associates, Hillsdale, New Jersey, pp243-264.
- CLARK, H.H. and GERRIG, R.J. (1983). Understanding old words with new meanings. Journal of Verbal Learning and Verbal Behaviour, Vol.22, pp591-608.
- CLARK, H.H. and HAVILAND, S.E. (1977). Comprehension and the Given - New Contract. In Discourse Production and Comprehension, ed. by R.O. Freedle. Ablex Publishing Corporation, Norwood, New Jersey, Vol.1, pp1-40.
- CLARKE, M.A. (1979). Reading in Spanish and English: Evidence from Adult ESL students. Language Learning, Vol.29, No.1, pp121-150.
- CLARKE, M.A. and BURDELL, L. (1977). Shades of meaning. In On Tesol '77: Teaching and Learning - Trends in Research and Practice, ed. by H.D. Brown, C.A. Yorico and R.H. Crymes. pp131-143.
- CLARKE, M.A. and SILBERSTEIN, S. (1977). Towards a realisation of Psycholinguistic Principles in the ESL reading class. Language Learning, Vol.27, No.1, pp135-154.
- COHEN, A., GLASMAN, H., ROSENBAUM-COHEN, P.R., FERRARA, J. and FINE, J. (1979). Reading English for Specialised Purposes: Discourse analysis and the use of student informants. Tesol, Vol.13, No.4, pp551-564.

- COOPER, M. (1984). Linguistic competence of practised and unpractised non-native readers of English. In Reading in a Foreign Language, ed. by J.C. Alderson and A.H. Urquhart. Longman, London and New York, pp122-135.
- CORDER, S.P. (1967). The significance of Learner's errors. International Review of Applied Linguistics, Vol.5, pp161-170.
- COULTHARD, M. (1977). An Introduction to Discourse Analysis. Longman, London.
- COUVES, G. (1978). Inferencing: a survey of one aspect of vocabulary acquisition in reading English as a second or foreign language. M.A. dissertation, University of Bangor, North Wales.
- COWAN, J.R. (1976). Reading, Perceptual Strategies and Contrastive Analysis. Language Learning, Vol.26, pp95-109.
- CRIPER, C. and WIDDOWSON, H.G. (1975). Sociolinguistics and Language Teaching. In Edinburgh Course in Applied Linguistics, ed. by J.P.B. Allen and S.P. Corder. Oxford University Press, London, Vol.2, pp155-210.
- CROTHERS, E., SUPPES, P. and WEIR, R. (1964). Latency phenomena in prolonged learning of visual representations of Russian sounds. International Review of Applied Linguistics, Vol.2, pp205-217.

- CRUSE, D.A. (1975). Hyponymy and lexical hierarchies. Archivum linguisticum, Vol.6 (New Series), pp26-31.
- CRUSE, D.A. (1977). The pragmatics of lexical specificity. Journal of Linguistics, Vol.13, pp153-164.
- CZIKO, G.A. (1978). Differences in First and Second Language Reading: The use of syntactic semantic and discourse constraints. Canadian Modern Language Review, Vol.34, pp473-489.
- D'ANGLEJAN, A. and TUCKER, G.R. (1975). The acquisition of complex English structures by adult learners. Language Learning, Vol.25, No.2, pp281-293.
- DAVIS, F.B. (1972). Psychometric research on comprehension in reading. Reading Research Quarterly, Vol.7, pp628-678.
- DANEŠ, F. (1974). Papers on Functional Sentence Perspective. Mouton, The Hague.
- DEYES, A.F. (1984). Towards an authentic 'discourse cloze'. Applied Linguistics, Vol.5, No.2, pp128-137.
- FAERCH, C. and KASPER, G. (eds.) (1983). Plans and strategies in interlanguage communication. In Strategies in Interlanguage Communication, Longman, London and New York. (Applied Linguistics and Language Studies Series - general editor - C.N. Candlin)
- FAERCH, C. and KASPER, G. (1984a). Two ways of defining communication strategies. Language Learning, Vol.34, No.1, pp45-63.

- FAERCH, C. and KASPER, G. (1984b). Pragmatic knowledge: rules and procedures. Applied Linguistics, Vol.5, No.3, pp214-225.
- FERGUSON, C.A. (1964). Diglossia. Word, Vol.15, pp325-340.
- FIRBAS, J. (1964). On defining the theme in functional sentence analysis. Travaux Linguistiques de Prague, Vol.1: L'Ecole de Prague d'aujourd'hui, Academia. Editions de l'Academie Tchechoslovaque des Sciences, Prague, et Librairie C. Klincksieck, Paris (1964 and 1966).
- FILLMORE, C.J. (1968). A case for case. In Universals in Linguistic Theory, ed. by E. Bach and R. Harms. Holt, Rinehart and Winston Inc., New York, Chicago, San Francisco, pp1-90.
- FILLMORE, C.J. (1977). Topics in lexical semantics. In Current Issues in Linguistic Theory, ed. by R.W. Cole. Indiana University Press, Bloomington, Indiana, pp76-138.
- FLOWER, L. (1979). Writer-based prose: A cognitive basis for problems in writing. College English, Vol.41, pp19-37.
- FREDERIKSEN, C. (1975a). Acquisition of semantic information from discourse: effects of repeated exposures. Journal of Verbal Learning and Verbal Behaviour, Vol.14, pp158-169.
- FREDERIKSEN, C. (1975b). Representing logical and semantic structure of knowledge required from discourse. Cognitive Psychology, Vol.7, pp371-458.

- FREEBODY, P. and ANDERSON, R.C. (1981). Effects of vocabulary difficulty, text cohesion and schema availability on reading comprehension. Center for the Study of Reading, University of Illinois, Urbana, Champaign. Technical Report 225.
- FRIES, C.C. (1957). The structure of English: An introduction to the construction of English sentences. Longman, Green and Company, London.
- GALISSON, R. (1983). Des mots pour communiquer: éléments de lexicométhodologie. CLE International (Didactique des Langues Etrangères, Paris).
- GARROD, S. and SANFORD, A. (1977). Interpreting Anaphoric Relations: The integration of semantic information while reading. Journal of Verbal Learning and Verbal Behaviour, Vol.16, No.1, pp77-90.
- GAZDAR, G. (1979). Pragmatics: Implicature, presupposition and logical form. Academic Press, New York, San Francisco, London.
- GOODMAN, K.S. (1967). Reading: a psycholinguistic guessing game. Journal of the Reading Specialist, Vol.4, pp126-135.
- GOODMAN, K.S. (1973). Psycholinguistic universals in the reading process. In Psycholinguistics and Reading, ed. by F. Smith. Holt, Rinehart and Winston Inc., pp21-27.
- GREENE, F.P. (1965). Modification of the cloze procedure and changes in reading test performance. Journal of Educational Measurement, Vol.2, No.2, pp213-217.

- GRICE, H.P. (1975). Logic and Conversation. In Syntax and Semantics, Vol.3: Speech Acts, edited by P. Cole and J.J. Morgan. Academic Press, pp41-58.
- GUTWINSKI, W. (1976). Cohesion in Literary Texts. Mouton: The Hague, Paris.
- HAGERUP-NEILSEN, A.R. (1977). The role of macrostructures and linguistic connectives in comprehending familiar and unfamiliar discourse. Unpublished doctoral thesis, University of Minnesota, USA.
- HALLIDAY, M.A.K. (1967a). Notes on transitivity and theme. Part 1. Journal of Linguistics, Vol.3, pp37-81.
- HALLIDAY, M.A.K. (1967b). Notes on transitivity and theme. Part 2. Journal of Linguistics, Vol.3, pp199-243.
- HALLIDAY, M.A.K. (1968). Notes on transitivity and theme. Part 3. Journal of Linguistics, Vol.4, pp153-308.
- HALLIDAY, M.A.K. (1970). Language structure and language function. In New Horizons in Linguistics, ed. by J. Lyons. Penguin Books Ltd., Harmondsworth, Middlesex, England, pp140-165.
- HALLIDAY, M.A.K. (1973). Explorations in the functions of language. Edward Arnold, London.
- HALLIDAY, M.A.K. and HASAN, R. (1976). Cohesion in English. Longman, London.
- HASAN, R. (1968). Grammatical cohesion in spoken and written English. Part One. Programme in Linguistics and Language Teaching, Paper 7. Communication Research Centre, Longmans, Green and Company, London and Harlow.

- HAVILAND, S.E. and CLARK, H.H. (1974). What's New: Acquiring new information as a process in comprehension. Journal of Verbal Learning and Verbal Behaviour, Vol.13, pp512-521.
- HARRIS, Z.S. (1970). Papers in structural and transformational linguistics. (Vol.1). D. Reidel Publishing Company, Dordrecht, Holland.
- HARVEY, P.D. (1983). Vocabulary learning: the use of grids. ELT Journal, Vol.27, No.3, pp243-246.
- HATCH, E. and FARHADY, H. (1982). Research Design and Statistics for Applied Linguistics. Newbury House Publishers Inc., Rowley, Massachusetts.
- HOEY, M. (1983). On the surface of discourse. George Allen and Unwin, London.
- HOTTEL-BURKHART, N. (1981). Cohesion in the English essays of native speakers of Canadian French. Ph.D. thesis, University of Texas.
- HYMES, D. (1972). On communicative competence. In Sociolinguistics, ed. by J.B. Pride and J. Holmes. Penguin, Harmondsworth, pp269-285.
- JAMES, C. (1980). Contrastive Analysis. Longman, London.

- JARRETT, D. (1984). Pragmatic coherence in an oral formulaic tradition. Coherence in spoken and written discourse, Vol.XII in the series Advances in Discourse Processes, ed. by D. Tannen. Ablex Publishing Corporation, Norwood, Jersey, pp155-171.
- JOHNSON, P. (1982). Effects on reading comprehension of building background knowledge. Tesol, Vol.16, No.4, pp503-516.
- JUST, M.A. and CLARK, H.H. (1973). Drawing inferences from the presuppositions and implications of affirmative and negative sentences. Journal of Verbal Learning and Verbal Behaviour, Vol.12, pp21-31.
- KAPLAN, R.B. (1966). Cultural thought patterns in intercultural education. Language Learning, Vol.16, No.1-20.
- KAPLAN, R.B. (1972). The anatomy of rhetoric: prolegomena to a functional theory of rhetoric. The Centre for Curriculum Development, Inc. Philadelphia, Pennsylvania.
- KARTTUNEN, L. (1971). Implicative verbs. Language, Vol.47, pp340-358.
- KEMPSON, R.M. (1977). Semantic Theory. CUP, Cambridge, England.
- KEMPSON, R.M. (1975). Presupposition and the delimitation of semantics. CUP, Cambridge, England.
- KINTSCH, W. (1974). The representation of meaning in memory. Lawrence Erlbaum, Hillsdale, New Jersey.
- KOCH, B.J. (1983). Presentation as a proof: the language of Arabic rhetoric. Anthropological Linguistics, Vol.25, pp47-60.

- LEECH, G.N. (1980). Explorations in Semantics and Pragmatics. John Benjamin's BV, Amsterdam.
- LEECH, G.N. (1981). Semantics: The Study of Meaning. Penguin Books (2nd ed.). Harmondsworth, Middlesex.
- LEECH, G.N. (1983). Principles of Pragmatics. Longman, London and New York.
- LEVENSTON, E.A. et al (1982). Discourse analysis and reading comprehension. Paper presented at the International Symposium on LSP. Eindhoven, Holland, quoted in A.F. Deyes (1984).
- LEVINSON, S.C. (1983). Pragmatics. CUP, London.
- LYONS, J. (1977). Semantics. Vols.1 and 2, CUP, Cambridge, England.
- MARCUS, S. (1980). Textual cohesion and textual coherence. Revue Roumaine de Linguistique Appliquee, Vol.25, No.2, pp101-112.
- MCCARTHY, M.J. (1984). A new look at vocabulary in EFL. Applied Linguistics, Vol.5, No.1, pp12-22.
- MEARA, P. (1978). Learners' Word Associations in French. Interlanguage Studies Bulletin, Utrecht, Vol.3, No.2, pp198-211.
- MENOSKY, D.M. (1976). Modes and Materials. In Findings of Research in Miscue Analysis: Classroom Implications, ed. by P.D. Allen and D.J. Watson. ERIC, NCTE and National Institute of Education, Urbana, Illinois, pp102-107.
- MOE, A.J. (1979). Cohesion, coherence and the comprehension of text. Journal of Reading, Vol.23, No.1, pp16-20.

- MORGAN, J.L. and SELLNER, M.B. (1980). Discourse and Linguistic Theory. In Theoretical Issues in Reading Comprehension, ed. by R.J. Spiro, B.C. Bertram and W.F. Brewer. Lawrence Erlbaum Associates, Hillsdale, New Jersey, pp165-200.
- MINSKY, M. (1975). A framework for representing knowledge. In The Psychology of Computer Vision, ed. by P.H. Winston. pp211-277.
- NEISSER, U. (1976). Cognition and Reality: Principles and Implications of Cognitive Psychology. W.H. Freeman, San Francisco.
- NYSTRAND, M. (ed) (1982). What writers know: the language process and structure of written discourse. Academic Press, New York, London.
- OLLER, J.W. Jr. (1972). Scoring methods and difficulty levels for cloze tests of proficiency in E as a SL. The Modern Language Journal, Vol.56, No.3, pp151-158.
- OLLER, J.W. Jr., BOWEN, J.D., DIEN, T.T. and MASON, V.W. (1972). Cloze tests in English, Thai and Vietnamese: native and non-native performance. Language Learning, Vol.22, No.1, pp1-16.
- OCHS, E. (1979). Planned and Unplanned Discourse. In Syntax and Semantics, Vol.12: Discourse and Syntax, ed. by T. Givón. Academic Press, New York, San Francisco, London, pp51-80.
- O'NEILL, R. (1978). Kernel One. Longman, London.

- OTT, C.E., BUTLER, D.C., BLAKE, R.S. and BALL, J.P. (1973).
The effect of interactive image elaboration on the
acquisition of FL vocabulary. Language Learning,
Vol.23, No.2, pp197-206.
- PALMER, F.R. (1976). Semantics - A New Outline. CUP,
Cambridge, England.
- PETÖFI, J.S. (1978). A formal semiotic text theory as an
integrated theory. In Current Trends in Text Linguistics,
ed. by W. Dressler. Walter de Gruyter, Berlin,
New York, pp35-46.
- PETÖFI, J.S. (ed.) (1979). Text versus sentence: Basic
Questions of Text Linguistics. Helmut Buske Verlag,
Hamburg, Vols. 1 and 2.
- PIAGET, J. (1955). The child's construction of reality.
(translated from French by M. Cook). Routledge and
Kegan Paul Ltd., London.
- QUIRK, R. and GREENBAUM, S. (1973). A University Grammar of
English. Longman, London.
- RICHARDS, J.C. (1974). Error analysis and second language
strategies. In New Frontiers in Second Language
Learning, ed. by J.H. Schumann and N. Stenson.
Newbury House Publishers Inc., Rowley, Massachusetts,
pp32-53.
- RIEGER, C.J. III (1975). Conceptual memory and inference.
In Conceptual Information Processing, ed. by R.C. Schank
North Holland, Amsterdam, pp157-268.

- RIESBECK, C.K. and SCHANK, R.C. (1978). Comprehension by computer: expectation-based analysis of sentences in context. In Studies in the Perception of Language, ed. by W.J.M. Levelt and G.B. Flores d'Arcais. Wiley, New York.
- RIVERS, W.M. (1968). Teaching Foreign Language Skills. The University of Chicago Press, Chicago and London.
- ROBSON, C. (1983). Experiment, design and statistics in psychology. Penguin Books, Harmondsworth, Middlesex, England (2nd edition).
- RUMELHART, D.E. (1975). Notes on a schema for stories. In Representation and Understanding: Studies in Cognitive Science, ed. by D.G. Bobrow and A.M. Collins. Academic Press, New York, pp 211-250
- RUMELHART, D.E. (1977). Understanding and summarising brief stories. In Basic Processes in Reading: Perception and Comprehension, ed. by D. Laberge and S.J. Samuels. Lawrence Erlbaum Associates, Hillsdale, New Jersey, pp265-304.
- RUMELHART, D.E. (1980). Schemata: the building blocks of cognition. In Theoretical Issues in Reading Comprehension, ed. by R.J. Spiro. B.C. Bruce and W.E. Brewer. Lawrence Erlbaum Associates, Hillsdale, New Jersey, pp33-58.
- RUMELHART, D.E. and ORTONY, A. (1977). The representation of knowledge in memory. In Schooling and the Acquisition of Knowledge, ed. by R.C. Anderson, R.J. Spiro and W.E. Montague. Lawrence Erlbaum Associates, Hillsdale, New Jersey, pp99-135.

- SANFORD, A.J. and GARROD, S.C. (1981). Understanding Written Language. Wiley, Chichester, England.
- SCARCELLA, R.C. (1984). Cohesion in the writing development of native and non-native English speakers. Ph.D. thesis, University of Southern California.
- SCHANK, R.C. (1975). Conceptual Information Processing. North Holland Publishing Company, Amsterdam.
- SCHANK, R.C. (1972). Conceptual dependency: A theory of Natural Language Understanding. Cognitive Psychology, Vol.3, pp552-631.
- SCHANK, R.C. and ABELSON, R.P. (1977). Scripts, Plans, Goals and Understanding: An enquiry into human knowledge structures. Lawrence Erlbaum Associates, Publishers, Hillsdale, New Jersey.
- SCHOLFIELD, P.J. (1979). On a non-standard dictionary definition scheme. In Exeter Linguistic Studies, Vol.4, ed. by R. Hartmann. Exeter University.
- SCHOLFIELD, P. (1982). Learning word meaning through explanation within English. Interlanguage Studies Bulletin Utrecht, Vol.6, No.2, pp34-63.
- SELINKER, L. (1972). Interlanguage. International Review of Applied Linguistics, Vol.10, No.2, pp209-231.
- SHAUGHNESSY, M.P. (1977). Errors and expectations. OUP, New York.
- SLAKTA, D. (1975). L'ordre du Texte. Etudes de Linguistique Appliquée, No.19, Didier, Paris.

- SMITH, F. (1971). Understanding reading: a psycholinguistic analysis of reading and learning to read. Holt, Rinehart and Winston, New York.
- SMITH, F. (1973). Psycholinguistics and Reading. Holt, Rinehart and Winston Inc., New York.
- SMITH, F. (1978). Reading. CUP, Cambridge, England.
- SMITH, N. and WILSON, D. (1979). Modern Linguistics: The result of Chomsky's Revolution. Penguin Books.
- STEFFENSEN, M.S. (1981). Register, Cohesion and Cross-cultural Reading Comprehension. Technical Report 220, Center for the Study of Reading, University of Illinois, Champaign, Illinois.
- STEFFENSEN, M.S., JOAG-DEV, C. and ANDERSON, R.C. (1979). Across cultural perspective on reading comprehension. Reading Research Quarterly, Vol.15, No.1, pp10-29.
- STEFFENSEN, M.S. and JOAG-DEV, C. (1984). Cultural Knowledge and Reading. In Reading in a Foreign Language, ed. by J.C. Alderson and A.H. Urquhart. Longman, London, pp48-64.
- STOODT, B.D. (1972). The relationship between understanding grammatical conjunctions and reading comprehension. Elementary English, Vol.49, pp502-504.
- STUBBS, J.B. and TUCKER, G.R. (1974). The cloze test as a measure of English proficiency. Modern Language Journal, Vol.58, No.5/6, pp239-241.
- SZWEDEK, A. (1980). Lexical cohesion in text analysis. Papers and Studies in Contrastive Linguistics, Vol.11, pp95-100.

- TANNEN, D. (1978). The effect of expectations on conversation. Discourse Processes, Vol.1, pp203-209.
- TANNEN, D. (1979). What's in a frame? Surface evidence for underlying expectations. In New Directions in Discourse Processing, ed. by R.O. Freedle. Ablex Publishing Corporation, Norwood, New Jersey, pp137-181.
- TANNEN, D. (1984). The pragmatics of cross-cultural communication. Applied Linguistics, Vol.5, No.3, pp189-195.
- TAYLOR, W.L. (1953). "Cloze Procedure": A new tool for measuring readability. Journalism Quarterly, Vol.30, pp415-433.
- TARONE, E. (1974). Speech Perception in SL Acquisition: A suggested model. Language Learning, Vol.24, No.2, pp223-233.
- TARONE, E. (1981). Some thoughts on the notion of 'communication strategy'. Tesol, Vol.15, No.3, pp285-295.
- THOMPSON-PANOS, K. and THOMAS-RUZIC, M. (1983). The least you should know about Arabic: Implications for the ESL writing instructor. Tesol, Vol.17, No.4, pp609-623.
- THORNDYKE, P.W. (1976). The role of inferences in discourse comprehension. Journal of Verbal Learning and Verbal Behaviour, Vol.15, pp437-446.
- THORNDYKE, P.W. (1977). Cognitive structures in comprehension and memory of narrative discourse. Cognitive Psychology, Vol.9, pp77-110.

- THORNDIKE, R.L. (1974). Reading as reasoning. Reading Research Quarterly, Vol.9, No.2, pp135-147.
- URQUHART, A.H. (1977). The effect of rhetorical organisation on the readability of study texts. Ph.D. thesis, University of Edinburgh.
- VACHEK, J. (ed.) (1964). A Prague School Reader in Linguistics. Compiled by J. Vachek, Studies in the history and theory of linguistics, Indiana University, Bloomington.
- VAN DIJK, T.A. (1973). Text Grammar and Text Logic. In Studies in Text Grammar, ed. by J.S. Petöfi and H. Rieser. Reidel, Dordrecht, pp17-78.
- VAN DIJK, T.A. (1977). Text and Context - Explorations in the Semantics and Pragmatics of discourse. Longman, London.
- WARREN, W.H., NICHOLAS, D.W. and TRABASSO, T. (1979). Event chains and inferences in understanding narratives. In Directions in Discourse Processing, ed. by R.O. Freedle. Ablex Publishing Corporation, Norwood, New Jersey, pp23-52.
- WIDDOWSON, H.G. (1978). Teaching language as communication. OUP, Oxford.
- WIDDOWSON, H.G. (1979). Directions in the teaching of discourse. In The Communicative Approach to Language Teaching, ed. by C. Brumfit and K. Johnson, OUP, London, pp49-60.

- WIDDOWSON, H.G. (1984). Reading and communication. In Reading in a Foreign Language, ed. by J.C. Alderson and A.H. Urquhart. Longman, London, pp212-230.
- WILKINS, D.A. (1979). Grammatical, situational and notional syllabuses. In The Communicative Approach to Language Teaching, ed. by C. Brumfit and K. Johnson. OUP, Oxford, pp91-98.
- WILLIAMS, R.C. (1980). Lexical Familiarisation. In The Reader and the Text, ed. by L.J. Chapman. Heinemann Educational Books, London, pp49-59.
- WILLIAMS, R.C. (1983). Teaching the recognition of cohesive ties in reading a foreign language. Reading in a Foreign Language, Vol.1, No.1, pp35-52.
- WINTER, E.O. (1977). A clause - relational approach to English texts - study of some predictive lexical items in written discourse. Instructional Science, Vol.6, Special Issue, pp1-92.
- WINTER, E.O. (1982). Towards a contextual grammar of English: the clause and its place in the definition of sentence. George Allen and Unwin, London.