

**Mystifying Discourse:**  
**A Critique of Current Assumptions and an Alternative Framework for**  
**Analysis**

**Kieran O'Halloran**

**Languages in Education**  
**Institute of Education, University of London**

**A thesis submitted for the degree of Doctor of Philosophy (PhD)**

**2000**



## ABSTRACT

The thesis is concerned with texts that mystify events being reported. It begins by focusing on Critical Discourse Analysis (CDA), a currently prominent enterprise, one of whose concerns is with the isolation of text which mystifies the nature of events described. When CDA isolates mystifying text, it is usually with the perspective of a non-analytical reader, either explicitly or implicitly in mind. However, the notion of a non-analytical reader in CDA is undeveloped from a cognitive point of view. The general structure of the thesis is as follows. In the *first* section, I show how CDA's approach to highlighting textual mystification is inadvertently bound up with *symbolic* notions of mental representation in cognitive science. In the *second* section, I outline theories of mental representation in connectionism and cognitive linguistics which problematise the symbolic assumptions of CDA and thus what CDA locates as mystifying text. The thesis develops cumulatively towards an alternative framework for highlighting mystification, in the *third* section, which includes compatible elements from connectionism, cognitive linguistics and recent psycholinguistic research on inference generation. My framework predicts how certain text can lead to mystification for a *non-analytical* reader who has little vested interest in a text and is largely unfamiliar with its subject matter. I show how mystification for this non-analytical reader is connected with inference generation but, in contrast to CDA, I provide a detailed processing profile for such a reader. Attitudes in CDA towards inference generation are often inconsistent and are in conflict with recent psycholinguistic research. My framework, rooted in empirical psycholinguistic study, enables a more plausible, comprehensive and thus consistent perspective on inference generation in reading and how this relates to mystification. Finally, my framework also highlights CDA's 'over-interpretation' in text exegesis done by proxy for non-analytical readers.

## ACKNOWLEDGEMENTS

Firstly, my warmest thanks to Professor Guy Cook who, in supervising this thesis, has provided constant challenge and inspiration. The intellectually enriching experience of supervisions, and the good humour in which they have taken place, have meant they were always something to look forward to, enjoy, and now something I shall miss.

I should like to thank the Institute of Education, University of London for providing a scholarship for the duration of the thesis.

Finally, to Elisabeth Leue - for your support and warmth and for putting up with me during the latter stages of writing-up - from the heart, thank you.

## TABLE OF CONTENTS

Abstract.....	2
Acknowledgements.....	3
Table of Contents .....	4
<b>CHAPTER ONE: INTRODUCTION.....</b>	<b>13</b>
1.1 Description of the Thesis.....	13
1.2 Critical Approaches to Language .....	16
1.2.1 Critical Linguistics.....	16
1.2.2 Critical Discourse Analysis <sup>3</sup> .....	19
Why CDA is ‘Critical’ .....	20
1.3 Criticisms of CDA <sup>4</sup> .....	24
1.4 CDA and the Highlighting of Textual Mystification: Sentential Structure.....	26
1.4.1 Text Inference Generation .....	27
1.4.2 Nominalisations .....	32
1.4.3 How Readers Can Confuse Semantic Transitivity with Syntactic Transitivity .....	36
1.5 CDA Highlighting Textual Mystification: Metaphor <sup>5</sup> .....	38
1.6 Cognitive Theoretical Positions found in the Thesis .....	39
1.6.1 Symbolicism .....	39
1.6.2 Connectionism .....	40
1.6.3 Cognitive Linguistics.....	41
1.7 An Alternative Framework: The Analysis of Mystifying Discourse Produced in the Micro-Context of Interpretation by a Non-Analytical Reader.....	42
1.7.1 Developing the Non-Analytical Reader .....	42
1.7.2 Discourse vs Text .....	43
1.7.3 Inference Generation as a Part of Discourse.....	44
1.7.4 Using IR to Highlight ‘Over-Interpretation’ in CDA on Behalf of a Non-Resistant Reader .....	46
1.7.5 Using Empirical Studies to Support My Framework .....	48

1.8 Chapter Outline .....	49
Notes .....	51
<b>SECTION A: THE RELATIONSHIP BETWEEN SYMBOLICISM, LOGICAL EMPIRICISM, THE CLASSICAL APPROACH TO CATEGORIES AND CDA ..</b>	<b>55</b>
<b>CHAPTER 2: OUTLINE OF SYMBOLICISM, LOGICAL EMPIRICISM AND THE CLASSICAL APPROACH TO CATEGORIES.....</b>	<b>56</b>
2.1 Introduction .....	56
2.2 Symbolic Architecture.....	57
2.2.1 Turing Machines and von Neumann Machines.....	57
2.2.2 Turing Machines, Symbolicism and Philosophy: General .....	58
2.2.3 Assumptions about the Mind-Brain in Symbolicism .....	59
2.2.4 Logical Empiricism <sup>1</sup> .....	61
2.2.5 The Influence of Logical Empiricism upon the Symbolic Paradigm .....	66
2.3 Chomsky and Symbolicism / Logical Empiricism.....	69
2.3.1 Synopsis of Chomsky's Early Positions.....	69
2.3.2 Affinities Between Chomsky's Early Position and Symbolicism / Logical Empiricism. ....	72
2.4 The Classical Theory of Categories.....	75
2.4.1 Aristotle.....	75
2.4.2 Chomsky / Twentieth-Century Linguistics.....	76
2.5 Approaches to Language Processing which Place Emphasis on Syntax.....	78
2.5.1 Chomsky as a Model of Processing - Derivational Theory of Complexity [DTC] .....	78
2.5.2 Canonical Sentence Structure (CSS).....	80
2.5.3 Modularity Hypothesis .....	80
2.5.4 Late Closure and Minimal Attachment Strategies.....	81
2.6 A By-Product of Syntax First Processing: Inference Last .....	84
Notes .....	85
<b>CHAPTER 3: SYMBOLIC, LOGICAL EMPIRICIST, CLASSICAL UNDERPINNINGS OF HOW CDA HIGHLIGHTS MYSTIFYING TEXT.....</b>	<b>87</b>

3.1 Orientation.....	87
3.2 CDA and Chomsky.....	87
3.2.1 Hodge, Kress and Transformations.....	87
3.2.2 Linguese.....	90
3.2.3 Hodge and Kress (1993) and The Derivational Theory of Complexity .....	91
3.2.4 Difficulties with Transactives as a ‘Basic Model’ .....	92
3.2.5 Difficulties with Transactives as the Basis of a Scientific Language.....	94
3.3 Perceptual Heuristics and CDA.....	98
3.4 CDA and Compositionality .....	101
3.4.1 Orientation.....	101
3.4.2 Fairclough (1995a: 110-3).....	101
3.4.3 Fowler (1986: 20).....	103
3.4.4 CDA’s Use of Lakoff and Johnson (1980) .....	104
3.5 CDA and the Classical Theory of Categories.....	105
3.5.1 Nominal Form, Objectifying Effects and Mystification.....	105
3.5.2 Nominals and ‘Scientific Language’.....	107
3.6 CDA, Text and Symbolicism / Logical Empiricism .....	108
3.6.1 Lexical Relations in Text.....	108
3.6.2 Sentential Relations in Text.....	109
3.7 CDA / CL, Inference Generation and Logical Empiricism / Symbolicism.....	113
3.7.1 Orientation.....	113
3.7.2 Returning to Fairclough (1992a); Clark (1992).....	114
3.7.3 Fairclough, Schank and Abelson (1977) and Inference Generation.....	115
Notes .....	118

## **SECTION B: CONNECTIONISM AND COGNITIVE LINGUISTICS -**

### **PROBLEMATISING SYMBOLICISM ETC AND IMPLICATIONS FOR CDA .. 120**

#### **CHAPTER 4: CONNECTIONISM - PROBLEMATISING HOW CDA HIGHLIGHTS**

##### **MYSTIFYING TEXT..... 121**

4.1 Introduction .....	121
4.2. Connectionism: The General Picture .....	122
4.2.1 Orientation: The Interaction of Syntactic and Semantic Information .....	122
4.2.2 Taraban and McClelland (1988).....	123
4.2.3 Arguments for Neural Modelling.....	124
4.2.4 Modelling Neural Activity.....	127
4.2.5 Sub-symbols / Microfeatures .....	129
4.3 An Outline of McClelland and Kawamoto’s (1986) Connectionist Model of Sentence Processing and its Implications for How CDA Highlights Mystifying Text.....	131
4.3.1 Outline of McClelland and Kawamoto (1986).....	131
4.3.2 How the Model Shades Meaning.....	134
4.3.3 Implications for CDA of McClelland and Kawamoto (1986).....	135
4.4 Connectionism and Inference Generation.....	141
4.4.1 Orientation.....	141
4.4.2 Inference Generation in Sentential Processing .....	142
4.4.3 Strength of Interpretation in Short Text Processing.....	144
4.4.4 Summing-Up.....	146
4.5 Connectionism and Mental Representation and Implications for CDA.....	147
4.5.1 Gross Descriptivism vs Gross Internalism.....	147
4.5.2 Representation vs Enactment .....	149
4.5.3 Connectionism and the ‘Representational’ Language of Science .....	151
4.6 Connectionist Responses to Fodor and Pylyshyn’s Arguments against the Connectionist Approach .....	153
4.6.1 Orientation.....	153
4.6.2 Compositionality.....	153
4.6.3 Systematicity.....	154
4.6.4 Productivity .....	157
4.7 Endpoints.....	158
Notes .....	159

## CHAPTER 5: COGNITIVE LINGUISTICS - PROBLEMATISING HOW CDA HIGHLIGHTS

<b>MYSTIFYING TEXT .....</b>	<b>160</b>
5.1 Introduction .....	160
5.2 Categories and Prototypicality .....	161
5.2.1 Prototype Effects .....	161
5.2.2 Prototypicality and Syntactic Categories: Implications for CDA .....	165
5.3 Basic-Level Categories .....	168
5.3.1 Orientation .....	168
5.3.2 Direct Understanding of the Basic Level .....	169
5.3.3 Cognitive Economy of the Basic Level .....	171
5.3.4 Challenging Objectivism .....	172
5.3.5 The Role of Context and the Basic Level .....	175
5.4 The Symbiosis of the Basic level and Prototypes .....	175
5.4.1 Orientation .....	175
5.4.2 Causation as Interaction .....	177
5.4.3 Implications for CDA .....	178
5.5 Cognitive Interdependence of Basic-Level Nouns and Basic-Level Action Categories .....	180
5.5.1 Basic-Level Action Categories .....	180
5.5.2 Basic-Level Event Categories .....	181
5.5.3 Implications for CDA .....	182
5.6 Superordinates .....	184
5.6.1 Superordinate Noun Categories .....	184
5.6.2 Superordinate Verb Categories .....	185
5.6.3 Translating Superordinates into the Basic Level .....	186
5.6.4 Implications for CDA .....	187
5.7 Cognitive Linguistics and Compounds: Implications for CDA .....	188
5.8 Endpoints .....	190
Notes .....	192
<b>SECTION C: CREATING AN ALTERNATIVE FRAMEWORK FOR THE ANALYSIS OF MYSTIFYING DISCOURSE .....</b>	<b>195</b>



<b>CHAPTER 6: PSYCHOLINGUISTIC EVIDENCE FOR SHALLOW PROCESSING WITH PARTICULAR REGARD TO INFERENCE GENERATION .....</b>	<b>197</b>
6.1 Introduction .....	197
6.2 Inferences in Text Comprehension and Likelihood of Their Generation .....	198
6.2.1 Coherence vs Elaborative .....	198
6.2.2 Psycholinguistic Evidence Supporting the Likelihood of Generation of Inferences .....	199
6.2.3 Antecedent (Backward) Causal Inferences and the Likelihood of Their Generation.....	206
6.2.4 Causal Consequent Inferences and the Likelihood of Their Generation.....	208
6.2.5 Instantiation and Likelihood of Generation .....	212
6.2.6 Instrument Inferences and the Likelihood of Their Generation .....	215
6.2.7 Summing-Up of the Psycholinguistic Research on Inference Generation.....	217
6.3 Shallow Processing of Text through Top-Down Processing .....	219
6.3.1 Sanford and Garrod's (1981 / 1994) Primary Processing Principle.....	219
6.3.2 'Moses Illusion' .....	221
6.3.3 Flores d'Arcais .....	223
6.3.4 Summing-Up.....	224
6.4 Classical Processing and Shallow Processing.....	225
6.4.1 Semantic Value.....	225
6.4.2 Formal Deductive Inferences .....	226
6.5 Selectional Processing of Narratives .....	229
6.5.1 Main vs Secondary Characters.....	229
6.5.2 Proper Names vs Role-Descriptions .....	230
6.6 Implications of This Chapter for CDA .....	233
6.6.1 Orientation.....	233
6.6.2 How the Idealised Reader Problematises what CDA Regards as Mystifying Text .....	234
6.6.3 Conflict between the Discourse of the Idealised Reader and CDA Explanation (Interpretation-2) for a Non-Resistant Reader .....	243
6.7 Summary.....	248
Notes: .....	250

**CHAPTER 7: HIGHLIGHTING COMPATIBLE ELEMENTS IN CONNECTIONISM,**

## **COGNITIVE LINGUISTICS, AND PSYCHOLINGUISTIC EVIDENCE FOR SHALLOW**

<b>PROCESSING.....</b>	<b>253</b>
7.1 Introduction .....	253
7.2 Compatibility of Connectionism with Cognitive Linguistics .....	253
7.2.1 General Compatibilities .....	253
7.2.2 Surface Structure (Linguistic Input) Not Equivalent To Process Output.....	255
7.2.3 The Functionalism of Cognitive Linguistics and Connectionism .....	257
7.2.4 Prototypes .....	259
7.3 Cognitive Linguistics and Shallow Processing .....	261
7.3.1 Cognitive Linguistics, Instantiation and Shallow Processing.....	261
7.3.2 Cognitive Linguistics, Instrument Inferences and Shallow Processing.....	263
7.3.3 Cognitive Linguistics, Causal Consequents and Shallow Processing .....	265
7.4 Connectionism and Shallow Processing .....	266
7.4.1 Connectionism, Moses Illusion and Shallow Processing.....	266
7.4.2 Connectionism, Sanford and Garrod's Primary Processing Principle and Shallow Processing.....	267
7.4.3 Connectionism, The Minimalist Hypothesis and Shallow Generation of Forward Elaborative Inferences .....	268
7.4.4 General Corollary .....	274
7.5 Tensions between Connectionism and Cognitive Linguistics: Implications for CDA .....	275
7.5.1 Tensions between Connectionism and Cognitive Linguistics in Treatment of Concepts.....	275
7.5.2 Explanatory vs Constitutive Metaphors .....	281
7.5.3 Implications for How CDA Treat Metaphor.....	283
7.6 Endpoints.....	290
Notes .....	291

## **CHAPTER 8: AN ALTERNATIVE FRAMEWORK FOR THE ANALYSIS OF MYSTIFYING**

<b>DISCOURSE PRODUCED BY A NON-ANALYTICAL READER .....</b>	<b>295</b>
8.1 Introduction .....	295
8.2 Why an Alternative Framework for the Analysis of Mystifying Discourse is Necessary .....	295
8.2.1 An Alternative Non-Analytical Reader.....	295
8.2.2 An Alternative Vocabulary for Inference Generation.....	296

8.2.3 The Incompatibility of Symbolic Postulates on Which CDA is Based with Psycholinguistic Evidence of Chapter 6 .....	297
8.3 An Alternative Framework for the Analysis of Mystifying Discourse.....	299
8.3.1 Orientation.....	299
8.3.2 The Processing Profile for the Idealised Reader: The Framework Principles.....	301
8.3.3 Reinforcing the Distinction between Discourse Analysis and Text Linguistics.....	307
8.4 ‘Protest Mob Storm Tube HQ’ News Text: Evening Standard 7 August 1996 .....	309
8.4.1 How the Text Leads to Mystification for IR as to the Nature of Causal Antecedence and Causal Consequence.....	309
Using I.R as a Yardstick for the Success of a Text.....	319
8.4.2 Contrast with Other ‘Reclaim the Streets’ Texts.....	321
8.4.3 Mystification through Automatic Selective Process around Main Characters .....	325
8.5 News in Brief Text - ‘Sex Assault Sailors Jailed’: The Mirror October 8 1997.....	327
8.5.1 Instantiation of the Text Body by the Headline .....	327
8.5.2 Contrast with Other News Texts.....	332
8.6. ‘Interpreter ‘Was Prepared to Prostitute Herself’ News Text: Evening Standard 21 November 1997 .....	334
8.7 ‘The Thin End of the Veg’ Text: Today October 20 1995 .....	340
8.8 Fairclough’s (1989) ‘Quarry Load-Shedding Problem’ News Text .....	343
8.8.1 Highlighting how the Text Can Lead to a Different Mystification in Reading Compared to that Found by Fairclough.....	343
8.8.2 The Interpersonal Function of the Headline .....	347
8.9 Providing a More Comprehensive Explanation of Trew (1979: 100-1) / Montgomery (1995).....	349
8.10 Endpoints.....	351
Notes .....	352
<b>CHAPTER 9: CONCLUSION.....</b>	<b>356</b>
9.1 Introduction.....	356
9.2 Shallow Reading vs Consumption Reading.....	356
9.2.1 The Consumption Metaphor as Constitutive Metaphor .....	356
9.2.2 The Consumption Metaphor as Misleading Discourse .....	357

9.2.3 The Consumption Metaphor as Legitimation of CDA .....	358
9.2.4 Saying One Thing: Doing Another .....	358
9.2.5 Consumption Constitutive Metaphor vs the Shallow Explanatory Metaphor .....	359
9.3 Specific Demonstrations of the Thesis .....	360
9.3.1 The Problematic Processing Assumptions in CDA.....	360
9.3.2 Situating CDA Processing Assumptions in an Historical Context.....	364
9.3.3 The Partiality of Interpretation Principle .....	364
9.3.4 Exploring the Micro-context of Interpretation to Show CDA ‘Over-Interpretation’ .....	365
9.3.5 Where CDA is an Inappropriate Mode of Discourse Analysis .....	365
9.3.6 Highlighting Recent Work on Inference Generation in Psycholinguistics.....	366
9.3.7 An Alternative Framework for the Highlighting of Mystifying Discourse .....	366
9.3.8 The Distinction between Text Linguistics and Discourse Analysis .....	368
9.4 Endpoint .....	369
Notes .....	369
<b>BIBLIOGRAPHY .....</b>	<b>370</b>

## CHAPTER ONE: INTRODUCTION

### 1.1 Description of the Thesis

Critical Linguistics (CL) and Critical Discourse Analysis (CDA) [see for example: Fowler, Hodge, Kress and Trew (1979), Kress and Hodge (1979), Fairclough (1992a), Hodge and Kress (1993), Caldas-Coulthard and Coulthard (1996), Chouliaraki and Fairclough (1999)] are enterprises broadly concerned with highlighting the traces of cultural and ideological meaning in spoken and written texts. CDA is now the most common contraction of the two, CL usually seen as a precursor of CDA. Both enterprises have become established text analytical strategies, extending their scope to educational practice, i.e., Kress (1989a), Fairclough (1992b), Wallace (1992) etc. One concern of both CL and CDA is the highlighting of how certain syntactic and semantic choices can *mystify* the nature of the events being described in a text. Indeed, CL and CDA have in recent years been the most dominant strategies in applied linguistics for the exposure of textual mystification. The theoretical base of such highlighting lies in Fowler, Hodge, Kress and Trew (1979) and Kress and Hodge (1979) whose ideas continue to inform detections of textual mystification in current CDA. Another facet of CL and CDA has been the incorporation of the approach to metaphor in Lakoff and Johnson (1980), two of the pioneers of *cognitive linguistics* (see 1.6.3). Common in CL and CDA is the notion that textual metaphor helps to construct a particular view of a situation which can mystify the actuality of events.

There has been very little inspection of the processing assumptions within CL and CDA, which underwrite how these enterprises highlight mystifying text. Some of these processing

assumptions derive from psycholinguistic work on language processing from the 1960s, but reassessment of such assumptions has not taken place. Indeed, there has been little drawing upon of more *recent* psycholinguistic evidence of text comprehension to substantiate claims as to how certain text can be mystifying for the reader. Moreover, there is very much an absence of a global perspective on text processing in CDA, with different CDA authors often in conflict with one another with regard to how text is processed, particularly with regard to the issue of inference generation; see 1.4. This lack of inspection of processing assumptions in CDA and lack of citation of work in psycholinguistics has been the case, no doubt, because of the weight CDA attaches to *social theory*, particularly the work of Foucault, e.g. Foucault (1972). With regard to text which can lead to mystification in reading, my thesis counters both the paucity of appreciation of cognitive matters and inconsistencies in CDA. Despite inconsistencies on the issue of text processing, what most CDA authors *do* have in common, however, is a cognitive / philosophical position on mental representation which underlies their perspectives on text processing. This particular notion of mental representation is known as *symbolicism* [see 1.6.1 for details]. However, this position in CDA is very much an *implicit, unrecognised* one. In *section A* of the thesis, I show how symbolic mental representation not only underlies CDA but also influences what CDA regards as being mystifying text. So the initial focus of the thesis might then be construed loosely as a piece of ‘CDA of CDA’, seeking to ‘denaturalise’ its assumptions of mental representation.

The second large focus of the thesis, which constitutes *section B*, outlines theories of mental representation in two relatively recent, prominent and broadly-speaking complementary enterprises - *connectionism* and *cognitive linguistics* (see 1.6.2 and 1.6.3 for more details). These enterprises present a direct challenge to notions of mental representation in

symbolicism. In cognitive psychology, symbolism and connectionism are, broadly speaking, the dominant approaches to cognitive modelling. Section B of the thesis *problematizes* the taken-for-granted symbolic postulates of mental representation in CDA with enterprises, which taken together, can be regarded as the natural antithesis of symbolism. In problematising the symbolic assumptions of CDA, I thus problematise how CDA locates mystifying text.

In section B, in effect, I produce a timely situating of CDA within a wider appreciation of the issue of mental representation. A brief history will make this clearer. While the socio-theoretical base of CDA, broadly speaking, is different to that of CL (see 1.2.1 and 1.2.2), notions of mental representation are carried through from CL into CDA of the eighties and nineties. Many of the bearings for the detection of mystifying text in CDA in the nineties derive from the CL work of Fowler et al. (1979) and Kress and Hodge (1979). For example, Fairclough (1995a) contains a chapter on representation that makes explicit citation and use of Fowler et al. (1979) and Kress and Hodge (1979). Hodge and Kress (1993), the second edition of Kress and Hodge (1979), actually does not revise the earlier work, including instead a new final chapter which is more CDA oriented and accordingly a revised bibliography (see: Hodge and Kress, 1993: xii). Now, connectionism only became a serious challenger to symbolism in the late eighties, continuing to flourish in the nineties. Because of this, looking at CDA in term of ‘neglect’ of the issue of mental representation would be an unfair *ex post facto*. The symbolic assumptions operative in the above authors would have seemed obvious in the absence of another cognitive perspective. It is timely then to consider issues of mental representation in CDA from the perspective of connectionism, as well as cognitive linguistics, given their contemporary prominence and the challenge they pose to symbolicists.

After problematising, in *section B*, the symbolic assumptions of CDA and thus how CDA locates mystifying *text*, in *section C* I undertake the creation of an alternative framework for the highlighting and analysis of mystifying *discourse*. I shall discuss the contents of *section C* in 1.7 and there comment on the text / discourse distinction. Let me now, though, in 1.2 flesh out the theoretical background and tactics of CL and CDA in general and then, in 1.3 and 1.4, their perspectives on textual mystification.

## **1.2 Critical Approaches to Language**

### *1.2.1 Critical Linguistics*

Critical linguistics (CL) is the term used to describe the application of a particular set of linguistic procedures to texts with a view to uncovering concealed cultural and ideological meanings, (see for example: Fowler et al. 1979; Kress and Hodge 1979; Fowler 1991.)

#### *Whorf*

Crucial to the theoretical grounding of CL has been what has become popularly known as the ‘Whorfian hypothesis’ (or ‘Sapir-Whorf hypothesis’): ‘...differences of linguistic structure cause the speakers of different languages in some sense to ‘see the world’ in different ways’ (Fowler, 1991: 30). Critical linguists cite Whorf as theoretical validation (e.g. Fowler et al., 1979) although with some modification. ‘Language’ is replaced with ‘linguistic varieties’ with an emphasis on text varieties. On this derivation, in CL, texts are regarded as enshrining ideology which in turn is manipulative of the reader’s thought. However, there are a number of problems with citing Whorf as theoretical validation. There



is actually no mention of a hypothesis in Whorf's writings (Whorf, 1956). On this point see Ellis (1993). When we read Whorf we find his writings indicate a more subtle appreciation of the relationship between language, thought and culture than that gleaned from the phrasing of the hypothesis. Another difference lies in the fact that language and thought are blended for Whorf rather than separated as the 'hypothesis' suggests. Indeed, as argued elsewhere, (O'Halloran, 1997), the 'hypothesis' is most probably an ex post facto distillation of his work, and as such a distortion.<sup>1</sup> While we might be suspicious that the 'hypothesis' would have had Whorf's approval, more charitably, one might regard the 'Whorfian hypothesis' as a peg on which to hang the common intuition that use of language can naturalise particular perspectives, a convenient citation point for those who feel this intuition strongly. Whorf is largely omitted now from the theoretical base of CDA (see 1.2.2). Nevertheless, in chapter 3, we shall see Whorfian vestiges in CDA's attitude to language and cognition. Another citation point for mystification in language is the work of Orwell, particularly his essays on the English language. His novel 1984, is also alluded to, especially the factitious dialect 'Newspeak', and the intended 'doublethink' that ensues through its use (Fowler et al., 1979).

### *Halliday*

Various facets of Hallidayan functional linguistics are also drawn upon in CL. Hallidayan linguistics is *functional* since the working premise in Halliday's work is that 'language is as it is because of its function in the social structure' (1973: 65). Underlying this strong functionalism are the postulates that linguistic forms realise particular functions and that speaker selections are systematic and principle-governed. For Halliday, language serves three functions: i) *ideational* - to represent people, objects, events, and states of affairs in the world; ii) *interpersonal* - to express the speaker's attitude to these representations; iii)

*textual* - to array i) and ii) in a cohesive and appropriate manner. Salient parts of the grammar are configured into systems to enable the actualisation of these functions. Speakers make choices from these systems such that the selection of a particular form is against a background of other potential selections. Linguistic selections are significant in that they are woven into the functional demands of the context of the situation. Taking these points as lead, Fowler et al. (1979: 187) 'follow Halliday in requiring that social meanings and their textual realisations be included within the scope of grammatical description'. Halliday's ideational function is crucial to CL since CL maintains that particular grammatical configurations in texts can be ideologically salient. For instance, it has been argued in CL that the occurrence of passives in a text may have ideological significance since passives allow agency deletion (Trew, 1979).<sup>2</sup> The same applies for what is known as *nominalisation* - where a noun form is viewed as being derived by transformation from a verb with the accompanying deletion of argument(s) (see 1.4.2).

### *Why CL is 'Critical'*

Having profiled the machinery of CL, what then is its purpose? Since the underlying premise is that encoded ideology in the text can manipulate or mystify thought, the 'critical' reader is one vigilant to the prospect of reader construction, and seeks to expose:

the ideological level of meaning in texts that are manipulative of their readers and / or mystifying of their subject matter. To read innocently, non-analytically, is to be manipulated and mystified...

Richardson (1987: 146-7)

So, *non-analytical reading* facilitates textual manipulation and mystification. In CL, the general supposition, although sometimes implicit, is that disclosure of ideology in a text is consonant with some sort of exegetic privilege. That is, a 'critical reading' has exegetic

privilege if it leads to the ‘social emancipation’ of the reader, preventing textual manipulation and mystification.

### *1.2.2 Critical Discourse Analysis<sup>3</sup>*

#### *Explanation (‘Interpretation-2’)*

Critical Discourse Analysis (CDA) has largely evolved out of CL, regarding itself as being corrective and constructive upon the tenets of CL. CDA criticises CL for its callow appreciation of ideology theory, following criticism from sociologists such as Thompson (1984). The upshot is that, compared to CL, CDA seeks to make much more of an explicit and theoretically rich yoking of sociology, the theory of ideology and linguistics. CDA still retains the Hallidayan component but in contrast to the ‘Whorfian’ base of CL, CDA draws upon Foucauldian Discourse theory, e.g. Fairclough (1992a). Foucault (1972) characterises discourses as systematically organised sets of statements that give expression to the meanings and values of an institution. Discourses are seen as defining and delimiting what it is possible to say and not possible to say (and by extension - what to do or not to do) with respect to the area of concern of that institution. For CDA, (see Fairclough (1992a: chapter 3 and especially page 84), *non-resistant* readers allow texts to position them as subjects such that they draw upon a particular discourse without them necessarily realising that this discourse sets limits on their *interpretation* (‘interpretation-1’ in Fairclough (1996: 50)). One of the purposes of CDA, then, is to expose how discourses can set such limits to the interpretation-1 of text by non-resistant readers. This procedure is known as *explanation* (or ‘interpretation-2’ in Fairclough, 1996: 50):

...interpretation-1 is part of the domain of interpretation-2; one concern of interpretation-2 is to investigate how different practices of interpretation-1 are socially, culturally and ideologically shaped. Fairclough (1996: 50)

Put another way, an aim of interpretation-2 is to highlight how the *macro*-context affects the *micro*-context of interpretation of a non-analytical reader.

### *Why CDA is 'Critical'*

The *critical* aspect of CDA does chime with that of CL but the scope is broader. As indication of this scope, here is Fairclough again:

Relationships between discursive, social and cultural change are typically not transparent for the people involved. Nor is technologization of discourse. 'Critical' implies showing connections and causes which are hidden; it also implies intervention, for example providing resources for those who may be disadvantaged through change. Fairclough (1992a: 9)

There is a strong case to be made for a mode of language education which emphasises critical awareness of ideological processes in discourse, so that people can become more aware of their own practice, and be more critical of the ideologically invested discourses to which they are subjected. Fairclough (1992a: 90)

The focus here is not just a criticism of how language is used to either construct or naturalise a set of attitudes, but an enabling of a critical stance on how socio-historical circumstances inscribe the reader into accepting certain positions over others. So, while CL might be regarded as a branch of stylistics which concentrates on the exposure of ideology within not only literary texts but non-literary ones too [see Simpson (1993: 2-10)], CDA practitioners (such as Fairclough and Kress in his more recent writings) place a greater emphasis than early CL on explaining the social conditions within which texts are read. Although CDA is in many ways corrective of CL, it must be said though that criticism of CL has also emanated from former practitioners (see Kress (1989b); Fowler (1988)). Indeed, Hodge and Kress, two of the progenitors of CL, would happily classify themselves as critical discourse analysts, [see the last chapter from Hodge and Kress (1993), added in the second edition].

In contrast to his CL period, Kress (1989a) now draws explicitly on the work of Foucault, with no allusion to Whorf as a theoretical underpinning. Where some of the original ‘Whorfianism’ is retained in former CL theorists (Fowler, 1991), there are nevertheless explicit references to Foucauldian discourse theory.

*Awareness in CDA of Exegetic Plurality*

Within CDA, there is also a greater awareness, at least (but see 1.3), of exegetic plurality compared to CL:

But texts may be open to different interpretations depending on context and interpreter, which means that the social meanings (including ideologies) of discourse cannot simply be read off from the text without considering patterns and variations in the social distribution, consumption and interpretation of the text.

Fairclough (1992a: 28)

...while some readers may interpret texts compliantly, fitting in with positions set up for readers in texts, other readings may be resistant.

Fairclough (1996: 50)

Signalled here is the greater awareness in CDA of the possibilities of exegetic plurality with regard to a text. There is also a greater awareness in CDA, compared to CL, that readers are capable of being *resistant*, i.e., they do not necessarily comply with the positioning of the text; [see also Fairclough (1992a: 29; 136) and Kress (1989a: 40-43) on non-resistant / resistant readers]. This recognition of exegetic plurality in CDA towards texts, though, does not equate with the exegetic plurality advocated in Derridean deconstruction. This is because CDA seeks *exegetic privilege*. CDA aims to show how non-resistant readers can be inscribed by a discourse or set of discourses which ‘prefer’ a particular interpretation so as to enable ‘liberation’ from the naturalising effects of dominant discourses. So while a Derridean approach might be construed as *radical hermeneutics*, CDA forms part of *critical*

*hermeneutics.*

*The Scope of this Thesis in its Relation to the Scope of CDA*

It should be clear, then, that the overall scope of CDA is much broader than that of CL, in being much more social theoretical. For example, Fairclough contends that a social theory of discourse should encompass the domains of both social production and social transformation, and that a critical discourse analysis is as much a method for studying social change as an investigation into how language use contributes to the reproduction of social structure (as well as seeking to highlight mystifying text). Fairclough (1992a), for instance, has used strategies within CDA as a method for tracing social and economic mutations in the Post-Fordist era and the influence of the market model in education. The boundaries of CDA are with such a developing enterprise difficult to draw, as different directions are sought and its eclecticism becomes even more ravenous. In the nineties, Fairclough (1992a; 1995a) draws upon the work of Bakhtin (1981, 1986) for example in highlighting the meshing of genres and discourses into *discourse types*. Indeed, Fairclough (1995a) reaches for greater scope in trying to provide an analytical ‘tool-kit’ to enhance *Critical Media Literacy*.

As I have said, with regard to CDA, this thesis will only be concerned with its cognitive assumptions, and so does not deal with the sociological / media theory which many practitioners draw upon. To furnish the reader with an idea of the scope of the thesis in its relation to the scope of CDA, below is an abridgement of what Fairclough (1995a: 201-205) calls ‘tentative agenda for teachers’. These ‘agenda’ are based on four questions for students of media and language which relate to any media text:

1. How is the text designed, why is it designed in this way, and how else could it have been designed?
2. How are texts of this sort produced, and in what ways are they likely to be interpreted and used?
3. What does the text indicate about the media order of discourse?
4. What wider sociocultural processes is this text a part of, what are its wider social conditions, and what are its likely effects? [my bold]

This thesis is in the domain of question 1 and the second part of question 2. For question 1, Fairclough summarises the ‘main forms of analysis introduced in the book’:

a) Intertextuality

- What genres, voices and discourses are drawn upon, and how are they articulated together?
- direct and indirect speech, generic structure or ‘staging’, narrative analysis (story, presentation), conjunctions, collocations

b) Language

i) *Representations*

- What presences and absences, foregrounding and backgrounding, characterize the text?
- What process and participant types are there? How are processes and participants categorized and metaphorized?
- What relationships are set up between propositions (clauses) in texts?
- presupposition, process and participant types, nominalisation, agency and voice (active and passive), categorization and wording, metaphor, main and subordinate clauses, theme, local and global coherence relations

ii) *Relations and identities*

What are the participants (voices) in the text, and how are they constructed?

- What relationships are set up between participants - specifically between:
  - media personnel (journalists, presenters) and audiences / readerships
  - ‘others’ (e.g. experts, politicians) and audiences / readerships
  - media personnel and ‘others’
- Are constructions of participants and relationships simple, or complex / ambivalent?
- What relative salience do institutional and personal identities have in the construction of participants?
  - oral delivery, body movement, key (serious or humorous), conversationalization, vocabulary, mood, modality, interactional control features, lists

iii) *Image and text*

- In the case of television, how are visual images constructed, and what relationships (e.g. tension) are set up between language and image?

In the domains of question 1, this thesis only focuses on b) i) (and only then certain aspects).

In relation to the second part of question 2 (*in what ways are texts likely to be interpreted?*) my emphasis is cognitive. As a final point, when I refer to CDA in this thesis, it is often in reference to the work of Norman Fairclough. This is mainly because he is commonly regarded as the principal exponent of CDA (Trask, 1999: 63) and because he has been more explicit than other CDA theorists in alluding to inference generation in text comprehension (see 1.7.3).

I have outlined very generally the theoretical background of CL and CDA. In 1.4 and 1.5, I will outline some examples of how CDA highlights textual *mystification*. But before I do, in 1.3 let me indicate some of the major criticisms made of CDA and CL.

### 1.3 Criticisms of CDA<sup>4</sup>

With regard to CDA, Stubbs (1997: 102) avers that ‘some sharp criticisms have been around for a long time, but remain unanswered’. One such criticism is that CDA lacks appreciation of how readers who are not analysts might interpret texts in different ways to the analyst; see Sharrock and Anderson (1981) and Richardson (1987: 152-3) for this criticism of CL, which is found in more developed form in recent criticisms of CDA by Widdowson (1994; 1995a; 1995b; 1996; 1997; 1998). I accord with this criticism in this thesis. Indeed, the lack of attention paid in CDA to the variation in micro-context interpretation has a certain irony. Fairclough (1992b: 28), for instance, criticises CL for giving little attention ‘to the processes and problems of interpretation, either those of the analyst-interpreter or those of the **participant-interpreter**’ [my bold]; his agenda for critical media literacy reproduced above (1995a: 201-205) includes ‘**how are texts likely to be interpreted?**’ [my bold]. The lack



of attention to micro-context variation leads Stubbs (1997: 106) to assert that CDA fails to meet its own criterion that one should study ‘how texts are produced, distributed and **consumed** [my bold] (Fairclough, 1995b: 1)’.

Widdowson (see above) has probably been the most vocal dissenter from CDA in recent years; [see also Carter (1997: 119-121) for discussion of Widdowson’s criticisms and Fairclough’s reply (Fairclough, 1996)]. Amongst many different censures, Widdowson criticises CDA and in particular Fairclough (1992a) for ignoring the variation of micro-contexts of interpretation. Since CDA does not seek to demonstrate concretely the potential variety of interpretations derived from a particular text by different readers, flagging the notion of interpretative-diversity as CDA does [Fairclough, 1992a: 28 (in 1.2 above) and Fairclough, 1996] is merely lip-service. For Widdowson, any interpretation of a text is ‘partial’, CDA’s interpretations being as partial as any reader’s, since facets of a text are cognised in line with the analyst’s values, motives etc. So, though CDA may regard their analyses as showing *more generally* how a dominant macro-context delimits the micro-context of interpretation of a non-resistant reader, Widdowson argues that in ignoring the details of micro-context interpretation of different readers, a CD analyst merely confirms *their own* political values, i.e., what they offer is merely a *partial* reading. In effect, CDA, then, give too much attention to the macro-context of reading at the expense of the micro-context.

As I have said, it is true that the nature of the micro-context of interpretation of different readers has not been explored in CDA in any detail. We shall see immediately below, in 1.4 and 1.5, that this point certainly applies to non-resistant or non-analytical readers. In 1.4 and 1.5, where I outline some text processing assumptions in CDA for the highlighting of

mystifying text, I flag where CDA explicitly uses the notion of a non-resistant reader, undeveloped as it is. At other times in 1.4 and 1.5, the notion of a non-resistant reader in CDA is merely implicitly understood and so is not given any serious attention.

There are discrepancies between CL and CDA in their socio-philosophical base, as we have seen, as well as a diversity of opinions within ('sociological') CDA. However, on the subject of mental representation, cognition, and textual mystification, *i.e. the concern of this thesis*, the CL and CDA theorists I outline in this thesis converge. This is because they share a set of implicit *symbolic* assumptions (see 1.6.1) that derive from the CL work of Kress and Hodge (1979) and Fowler et al. (1979). Consequently, (most of the time) my use of the contraction CDA is not reductionist but convenient shorthand for notions of mental representation and cognition in CDA. I will highlight in chapter 3 how symbolic ideas of mental representation underpin most of the approaches to text processing in CDA and thus underpin how CDA locates text that is mystifying of subject matter.

Let me now broadly outline some examples of how CDA locates mystifying text. For reasons which will become clear later (see 1.7.3), I also include assumptions of text processing in CDA, which do not, at least directly, relate to the issue of textual mystification.

#### **1.4 CDA and the Highlighting of Textual Mystification: Sentential Structure**

What follows consists of examples of text processing assumptions in CDA and thus of how CDA locates text which is mystifying of subject matter due to particular choices of sentence structure. These examples relate to a) text inference generation, b) nominalisations, c)

confusing semantic transitivity with syntactic transitivity. CDA does not have a detailed model of processing in mind when locating textual mystification. Rather, as we shall see, CDA has a set of ‘principles’ that guide the analyst as to what is highlighted as mystifying.

#### *1.4.1 Text Inference Generation*

##### *Inferences as Weaker Representations: Inferences Downgraded*

In CDA, inferences generated in textual comprehension are often regarded as *weaker* representations than the ‘surface’ sentential structure. Let us now consider some examples. The following is an extract from Trew (1979), now regarded as a classic CL article. This will also be referred to later in the thesis not only as it has attained the status of a classic, but because of its continuous citation, reproduction and endorsement in introductory textbooks.

The following is, for example, reproduced and endorsed in Toolan (1988: 229-30), Lee (1992: 100) and Montgomery (1995: 240). Here now is Trew (1979: 98-9):

##### *The Times*

##### RIOTING BLACKS SHOT DEAD BY POLICE AS ANC LEADERS MEET

Eleven Africans were shot dead and 15 wounded when Rhodesian Police opened fire on a rioting crowd of about 2,000.

‘Not only is it [The Times report] in the passive, but the syntactic agent is deleted (‘11 Africans were shot dead by...’) **and is identified only weakly by implication** through the temporal conjunction with the police opening fire (‘when police opened fire on a rioting crowd of about 2,000’). Looking at this in purely syntactic terms, with the deletion of the agent there is no longer any direct reference to who did the action and there is a separation of the action from whomever did it.’ [my bold]

What is suggested by Trew is that the best representation of an event is one where actor and process are linked directly and adjacently in the active voice. Any inferencing necessary to

link perpetrator and action, something which is not located directly in the ‘surface’ structure, is a *weaker* representation of the actual event. Trew does not mention *explicitly* that his analysis is based on the response of a non-analytical or non-resistant reader. But I assume that he has an *implicit* non-resistant reader in mind who does not notice that the agency of the police is downplayed in the ‘surface’ syntactic structure, partly because of the weak representation they generate. With the agency of the police attenuated, textual mystification occurs. Hodge and Kress (1993: 26) hint at a similar assumption and focus on the ‘surface’ structure, when they contend that in *passives*:

‘the link between actor and process is weakened, that is, the causal connection is syntactically looser.’

A coincident assumption that inferences are weaker representations can be found in Simpson (1993: 171) and in Montgomery (1995: 240) who reproduces and concurs with Trew’s analysis. Another assumption at work in the above is that inferences are *separate* from the processing of the ‘surface’ structure of the text.

In CDA, there is often an over-emphasis on sentential structure at the expense of top-down inferential processes. Consider the following from Fairclough (1989: 50-1):

### **Quarry load-shedding problem**

UNSHEETED lorries from Middlebarrow Quarry were still causing problems by shedding stones on their journey through Warton village, members of the parish council heard at their September meeting.

The council’s observations have been sent to the quarry management and members are hoping to see an improvement.

*Lancaster Guardian, 12 September 1986*

Causality is attributed to *unsheeted lorries from Middlebarrow Quarry*. This itself contains unspecified causality again, for *unsheeted* implies the failure of a process to happen - someone did not put sheets over the loads, when (one assumes) they ought to have done. It is difficult to take literally the notion that the *lorries*

are the cause of the problem, and it is evident that in a different representation it could be this 'someone' - presumably the *quarry management* or people under their control. Yet the quarry management figure only in the second paragraph in this representation as in receipt of the council's *observations*, a term which again avoids attributing any responsibility (it might have been *complaints*).

That Fairclough contends, 'it is hard to take seriously that the 'lorry' is the cause of the problem' suggests that for Fairclough the 'surface' structure *prevails* over more top-down inferences as to causal responsibility. In other words, the 'surface structure' mystifies as to causal responsibility. [This tendency to emphasise 'surface' structure is also seen in discussion of 'transactives' in Kress (1989a) and Hodge and Kress (1993); see chapter 3 below.] Finally, like Trew, Fairclough's non-analytical reader is only implicitly understood.

### *Inferences as Strong Representations*

In CDA, the notion that inferences make for weaker representations is not the only implicit postulate around the issue of inference generation. Consider the following text and Fairclough's (1989: 52-3) commentary upon it:

The Paras' new leader: He'll do his job well says major's wife [+ Photo of Major Keeble]

The wife of the new CO of the 2nd Parachute Battalion spoke last night of her fears for her husband's safety. As she played in the sunshine with her four children, Jenny Keeble said she hoped her husband would not have to go into battle again. She said: 'I pray he and his men have done enough. But if they do go on I know that he is a man who will do his job to the best of his ability and I am certain he and the 2nd Parachute Battalion will succeed.'

Major Christopher Keeble, a 40-year-old devout Roman Catholic, is to succeed Colonel Herbert Jones who died leading his men against an Argentine machine-gun post in the battle for Goose Green.

Yesterday Jenny Keeble's family and friends gathered around in the garden of her old vicarage home - a rambling Tudor building at Maddington on Salisbury Plain - for a picnic afternoon as she tried to maintain an air of normility [sic] for the children's sake.

For Fairclough such a text helps to stereotype 'army wives', thereby placing limits upon the

meanings that readers attach to such an individual as Jenny Keeble. Here is Fairclough again (1989: 52):

Notice that at no point here (or in the rest of the article) is Jenny Keeble explicitly *said* to be ‘a good wife’ or an admirable person; **the process depends entirely on an ‘ideal reader’s capacity to *infer* that from the list of attributes** - she expresses confidence in her husband’s professional abilities, she is concerned for his safety, she ‘prays’ he has ‘done enough’, she tries ‘to maintain an air of normality for the children’s sake’...the process presupposes **an ideal reader** who will indeed make the ‘right’ inference **from the list, i.e. have the ‘right’ ideas about what a ‘good wife’ is.** Texts such as this reproduce sexist, provided that readers generally fall into the subject position of **the ideal reader**, rather than **opposing it.** [my bold]

This time the analysis explicitly mentions a non-resistant or ‘non-oppositional’ reader, termed ‘the ideal reader’ by Fairclough. However, this ‘ideal reader’ is undeveloped from a cognitive point of view. Now, compare the above with the postulate of CDA practitioners in the last section that the inference being generated is a *weaker* mental representation than mental representation of the ‘surface’ structure. We find quite the opposite. Indeed, for Fairclough, the inference itself is *strong* enough to lead to the reproduction of sexist. The two notions of inference generation with regard to mystification that we have met so far are, then, in conceptual tension. As a final point, we can see another conceptual tension over two pages of the same book between Fairclough’s (1989: 52) top-down emphasis on processing (‘Jenny Keeble’ text) and Fairclough’s (1989: 50-1) more bottom-up emphasis on syntactic structure (‘Quarry load-shedding problem’ text).

### *Inference as Work vs Automatic Gap-Filling*

Consider firstly Fairclough (1989: 81):

There is no sharp dividing line between automatic gap-filling and inferencing, both because there is probably

a scale from links which need no working out to links which need a lot of inferential 'work', and because a link which is supplied automatically by one person may need **inferential work** from another (or indeed from the same person on another occasion). Text 4.2 [a problem-page letter from a teenage magazine] would probably not require any inferential work from regular readers of the sort of magazine it comes from, but it might from other people. [my bold]

So, for Fairclough, there are broadly speaking two types of inferencing: 'automatic gap-filling' which require minimum cognitive labour while 'inferences' are those which incur a higher than minimum amount of cognitive labour. The processing assumption above is that readers *not* familiar with the subject matter would *still* work to generate 'inferences'.

A further, more implicit assumption is that the reader will *work* to produce necessary *inferences* to make what Fairclough regards as coherence - as though there is one ideal coherence to which all readers will eventually arrive, automatically or through inferential work (see also Fairclough, 1992a: 177). Along similar lines, consider the following commentary (on problem page advice) from Gough and Talbot (1996: 226), who adopt Fairclough's (1989) position on automatic gap-filling vs inferential work:

Many heterosexual men have a passing curiosity about homosexuality, and that isn't such a bad thing. It compels you to make choices.

...the **causal link** which is needed to coherently combine these two sentences is not cued by any formal element, and this is a point where **the reader's complicity** is required if the two sentences as they stand together are to make sense. **The 'missing link' we need to supply is that heterosexuality and homosexuality are separate sexualities and that interest in homosexuality is useful inasmuch as it reinforces this separate heterosexual identity.** For some readers it may require inferential work...Following Fairclough's approach, this interpretation would be accounted for using the notion of automatic 'gap-filling' between explicit propositions. A reader who is unfamiliar with problem pages...would need to engage in a good deal of inferential work to make this connection. [my bold]

The 'complicit' reader mentioned, in the above, I take to be a non-resistant reader, but like Fairclough's 'ideal reader', the 'complicit reader' above is undeveloped from a cognitive

point of view. The assumption in the text above is that the non-resistant reader who is ‘unfamiliar with problem pages’ will ‘work’ to produce the inference that ‘interest in homosexuality (in this context, by adolescent males) is useful in as much as it reinforces this separate heterosexual identity’. In doing so, the non-resistant reader becomes ideologically positioned. But there is a tension here. While we might suppose that a resistant reader, in necessarily being more critical of the text, engages in more cognitive work, a non-resistant reader is surely someone who is *not* making such an effort.

#### *1.4.2 Nominalisations*

##### *A High Degree of Nominalisation in a Text Requires Extra Processing Effort*

The following involves a discussion by Hodge and Kress (1993: 21) of part of a newspaper editorial on the miner’s overtime ban during the winter of 1972-73 and in particular the sentence from the editorial:

The Government knows that in early 1972 it was caught out by picketing of power stations which curtailed coal deliveries.

Here is the analysis:

##### *Picketing...curtailed coal deliveries*

If we asked speakers of English what the meaning of *picketing* was, they would probably explain it by describing the kinds of things involved: strikers, the action, a factory, or, in this case, a coal-depot. The noun is a contraction of a significant kind. The single word necessarily implies a particular kind of actor and a particular object of action. We might represent the process in this way:

strikers picket a factory ⇒ picketing



...there are two major effects associated with that transformation, which amount to a quite radical changing of the original form. First, although we know that there was an actor and an affected, the specific identities of both have been lost. We can guess about their identity but can never be certain. Second, in the resulting surface form the only thing that meets us is the verbal version of the action which was performed, and in this way our attention is directed to what is present and directed away from what is no longer there. So the focus of the expression has been altered by the speaker, our vision has been channelled and narrowed.'

For Hodge and Kress, 'strikers picket a factory' is the form (i.e., actors-process-patient) that is necessary for a comprehensive understanding of the situation. However, the 'surface' or *nominalised* form here is 'picketing' and so to understand the situation the article is reporting, the reader has to make more processing effort to 'recover' the 'deep' form.

As readers of this editorial we should have to be alert and willing to engage in mental exercise to get beyond the seductive simplicity of the final form, with just three entities, and seemingly precise relations, where everything seems to be there on the surface...we can see **that few commuters on the 8.05 from Brighton would have the energy to perform the mental gymnastics required.**'

Hodge and Kress (1993: 22) [my bold]

This time a non-analytical reader is explicitly referred to ('non-energetic' reader). What is alleged here is that through particular syntactic selections, processing can be made more laborious to the extent that a non-analytical ('non-energetic') reader is less likely to 'recover' the 'deep' meaning. Since the 'deep' meaning is necessary for a 'comprehensive' understanding and appreciation of the events being described, if it is difficult to 'recover', textual mystification transpires since processing for the 'non-energetic' reader is *shallow*.

As a final point in this section, compare the notion of cognitive labour in the above text with Fairclough's / Gough and Talbot's assumption that readers are prepared to work at producing inferences in 1.4.1. There is a conflict. On the one hand, Hodge and Kress (1993: 21) argue that non-analytical readers are *not* prepared to invest much processing labour, but on the other, Fairclough (1989: 52) argues that non-analytical readers *are* prepared to invest more than minimum processing labour.

*Excessive Nominalisation Makes a Text Abstract and Distant from Concrete Events*

For CDA, nominalisations ‘objectify’ the event being described because of their nominal form and in doing so mystify the actualities of the event being described. This argument is also found in most practitioners of CL, especially the CL of Kress and Hodge (1979) and Fowler et al. (1979). Here firstly are Fowler and Kress (1979b: 208): on the ‘objectification’ effects of nominals:

Two further effects of nominalisation may be mentioned briefly. The first is objectification, the rendering of a process as an object: ‘We still need lots of *contributions* to the jumble sale’; ‘our new *development*, the ‘Interference Absorption Circuit’; ‘Now that you’ve had your first *look* at the new Record Saloon’. This in turn affects lexicalization, the provision of words and phrases to code new concepts or consolidate existing ones: ‘strict segregation’, ‘basic approach’, ‘school dinner services’, ‘people’s trial’, ‘illegal detention’. Lexicalization fixes the object-as-process as a single habitualised entity.

Indeed, the tendency to regard all nouns as things, to see ‘thingness’ as a necessary and sufficient condition for nounhood is an assumption often operative in CDA / CL (see, for example, Hodge and Kress (1993). The argument that nominal description of actions ‘objectifies’ action and thus creates distance between the event and its appreciation by the reader is carried forward into latter day CDA (e.g. Kress, 1989a: 58; Martin, 1989: 43; Fowler, 1991: 80; Lee, 1992: 95; Fairclough, 1995a: 112). Lee (1992: 95), flagging Fowler and Kress (1979b), states that: ‘it is arguable that one effect of the nominalised structure...is to *reify* (my italics) the event in question, and thereby to abstract away from the event, to diminish its violent nature’. And in what follows, Martin (1989: 43) specifically relates the effect of distancing in reading from the actualities of events (killings of kangaroos and seals) to the use of incongruent *nominal* linguistic forms. Specifically, this effect is produced because nominals treat ‘killing as a kind of thing’:

In chapter 2 we also looked at the question of congruence, asking, for example, whether actions were being expressed in nouns or verbs. Following this up in Texts 3.1 and 3.2 we can see a big difference in the way in which actions are realised. Overall the CWF (seal) article realises actions as nouns twice as often as the ACF (kangaroo) editorial.

The CWF article uses three types of nominal structure in place of verbs to realise actions. One puts the action into the modifier of an abstract noun: e.g. *sealing operation, killing techniques*. Another makes use of a nominalised form of a verb: *statements, definition, death, coverage, constraints*. A third simply realises the actions as a noun: *the whitecoat harvest, the East Coast seal hunt, the seal hunt*.

Of particular interest is the way in which the two texts refer to the killing of seals and kangaroos. **The ACF tends to refer to the killing congruently, as a process: *the massive level of killing; the favoured killing of bigger, heavier male kangaroos; whose lives will be obliterated; killing 3 million kangaroos a year; when our prime wildlife is killed on this scale*. The CWF text on the other hand tends to refer to the killing indirectly, using incongruent forms: *killing techniques, the whitecoat harvest, the slaughter of animals, the East Coast seal hunt, a slaughtering operation, killing methods, an almost instantaneous death, a humane death, the seal hunt*, and so on. **In this way the ACF text focuses on the process of killing, while the CWF text treats the killing as a kind of thing. This has the effect of immobilising the most unsavoury part of the seal hunt and helps draw attention away to other ‘factual’ considerations.** [my bold]**

For Martin, then, the high propensity of nominal forms are incongruent representations for the actions of the actual event. The use of action categories as modifiers of abstract nouns is also incongruent. And the ‘object-like’ ‘killing as a kind of thing’ diminishes the actions taking place. The implication is that this ‘diminishing’ will transpire for a non-analytical reader, although Martin’s analysis makes no explicit mention of a such a reader.

Recall the ‘Quarry Load-Shedding Problem’ text that I highlighted in 1.4.1. For Fairclough (1989: 50-51):

...the grammatical form in which the headline [‘Quarry load-shedding problem’] is cast is that of a *nominalization*: a process is expressed as a *noun*, as if it were an entity.

Fairclough (1989), in the above, is highlighting how the nominalisation, in the removal of agent and patient, mystifies the causality of the event. However, if we compare Fairclough

above with what I have bolded of Martin, we see there is some conflict. On Fairclough's rationale, 'the massive level of killing' would be a nominalisation and thereby incongruent with the notion of a process. But Martin asserts that 'massive level of killing' *is* congruent with the notion of a process. Another example of conceptual conflict then.

#### 1.4.3 How Readers Can Confuse Semantic Transitivity with Syntactic Transitivity

The final assumption about text processing relates to the possibility that readers can register syntactic transitivity in terms of *semantic* transitivity when the two are *not* equivalent. As an outline of what this entails, consider the following from Kress (1993: 181-2), which is an analysis of the clause 'his parents could not afford a uniform', which features in a newspaper text:

My interest here lies initially in the construction of the concept of poverty in the popular media; and my specific focus is the clause 'his parents could not afford a uniform' in sentence 1...The cited clause in sentence 1 presents a syntactic ambiguity. What syntactic analysis / description do we give to *afford*? Overtly it looks to be a transitive verb, with *a uniform* as direct object, and *his parents* as subject. A syntactically analogous form to that reading of this clause is 'His parents (could not) buy a uniform', which is clearly transitive. However, 'affording a uniform' is not clearly transitive; its subject noun is not clearly agentive: *His parents* is not an unambivalently agentive subject. *Afford* is a state of being, not a process under the control of an agent. And clauses which are not clearly transitive do not passivize easily: 'a uniform was (not) afforded by his parents' (and similarly with further tests, such as prenominalizing of the adjective), 'The afforded uniform...'. **Yet many readers of the *Daily Express* may read across this clause in reading the text, and read it as fully semantically transitive**, which I shall call, following Hodge and Kress (1993), a 'transactive'. In support of that reading, these readers might say: 'We scrimped and saved, and we afforded a uniform for our children, so why can't they?!'

There are thus at least two syntactic readings for this clause, very likely corresponding to the readings of different audiences - a hypothesis which could be tested.... [my bold]

Kress continues. If the clause 'his parents could not afford a uniform' is read as simply '“involving” subjects' it supports a reading of 'poverty as an event in which participants are caught up' (1993: 182-3). However, if agency is assigned to the grammatical subject, as a

result of ‘*reading across*’ a clause (i.e. by a *non-analytical* reader), this can support a ‘politically reactionary view’ where the poor are responsible for their own poverty. Again the influence of the ‘surface’ structure is salient. The ‘surface’ structure of subject-verb-object has the capacity to promote the reading of AGENT-PROCESS-PATIENT, for a non-analytical reader, and thus lead to *poverty as a state* in which people are caught up being obscured. [In Fowler and Kress (1979a: 42 and Fowler and Kress (1979b: 209), we find a similar analysis which relates as much to semantic transitivity as thematization.] In highlighting how a subject-verb-object clause can be ‘read across’, Kress thus indicates how a type of *shallow* processing might occur similar to that highlighted in Hodge and Kress’s (1993: 21) analysis of the ‘picketing strikers’ text.

As a final point compare 1.4.1 with this section. We saw in 1.4.1 a conceptual tension between i) the notion that text inferences are weaker representations and ii) the notion that inferences lead to mystification because of their strength. In the first postulate (Trew, 1979), the ‘surface’ structure of the sentence was regarded as having cognitive salience and that which was ‘added’ (i.e. inferred) was seen as weaker. In Fairclough’s analysis of the ‘Jenny Keeble’ text that which was added inferred *into* the text in top-down processing was seen as potent and leading to the reproduction of sexism. In this section, what is ‘added’ to the sentence - the imparting of semantic transitivity to syntactic transitivity - is regarded as being *more* potent than the *actual* semantic structure. Thus, Kress is in concert with Fairclough but not with Trew.

I now want to turn my attention to how CDA regards the relationship between *metaphor* and mystification.

### 1.5 CDA Highlighting Textual Mystification: Metaphor<sup>5</sup>

The view that metaphor is merely ornamental, transparent and superficial is rejected in CDA.

Instead the following view from Lakoff and Johnson (1980: 3) is thoroughly endorsed in CDA:

...metaphor is pervasive in everyday life, not just in language but in thought and action. Our ordinary conceptual system, in terms of which we both think and act, is fundamentally metaphorical in nature.

From such a position, CDA argues that certain metaphorical choices can mislead, distort or act as a buffer between the reality of the social world and the reader, preventing a 'full' appreciation of the event being described; [see Fairclough (1992a: 194-195) as well as Fairclough (1989); Kress (1989a); Lee (1992) and Shepherd (1994)]. As a concrete example, consider the following from Lee (1992: 91-92). Lee comments upon a newspaper report, an extract of which I also produce below:

The black township of Soweto, which has been simmering with unrest since the riots on June 16 and the shooting of 174 Africans, erupted again today...Police with automatic rifles and in camouflage uniform headed the marchers off after they had swept through a roadblock.

For Lee (1992: 93):

...the metaphorical process...treats the people of Soweto as some kind of natural force, specifically here as a volcano which has been 'simmering' with unrest and then 'erupted'. This is echoed in the later report that the marchers had 'swept through' a roadblock, a river. Note, too that the emotions of individuals and the actions that they give rise to are transferred onto the place where they live. It is 'the township' that has been simmering and that now erupts, rather than the Sowetans experiencing feelings of anger and deciding to march.

**The effect of these processes of metaphor and metonymy is arguably to distance the reader from the subjects of the report.** In speaking of the Sowetans as a natural force and as a place, the emotions of the people involved and the decisions which they make to engage in particular actions are eliminated from the process of interpretation. The situation is seen as resulting from some kind of inevitable set of natural laws

rather than from human feelings and decisions.

[my bold]

This alleged ‘distancing’ or ‘buffering’ effect of the metaphors, I treat as a kind of mystification of the actual events which took place. Similar ideas on the metaphor of ‘eruptions’ are expressed in Fairclough (1995a: 114). Finally, although it is not mentioned explicitly, I take the ‘reader’ posited by Lee to be a non-resistant one.

In 1.4 and 1.5, I have laid out some examples of how CDA highlights mystifying text for a non-analytical or non-resistant reader. [In later chapters, more examples of how CDA isolates mystifying text will be provided]. However, we have seen that there is no stable, consistent and detailed conception of a non-analytical reader in CDA, often the concept of a non-analytical reader is only implicitly understood, and several processing assumptions in CDA are in conflict with one another. In 1.7, I shall give an indication of an alternative framework I create, in section C of this thesis, for the analysis of mystifying discourse. In contrast to CDA, this framework will be based on a *detailed* and *consistent* notion of a non-analytical reader. But before I begin to discuss this alternative framework, I outline in 1.6 the cognitive theoretical positions found in this thesis.

## **1.6 Cognitive Theoretical Positions found in the Thesis**

### *1.6.1 Symbolicism*

I argue in this thesis that a particular notion of mental representation has been inadvertently absorbed into CDA via CL, especially via *Chomskyan* ideas. This is known as the *symbolic* view of mental representation, or *symbolicism* for short. I argue also that symbolic notions influence how CDA isolates mystifying text. Symbolic modelling of the mind is based on

the idea that mental processing consists of the manipulation of symbols in accordance with a rule-governed system analogous to a syntax. These symbols are held to be storable and retainable and thus an *enduring* set of entities. As we shall also see, the tradition of *logical empiricist* philosophy with regard to the philosophy of language and the philosophy of mind also permeates cognitive assumptions of CDA.

### *1.6.2 Connectionism*

The work in CL on linguistic representation and ideology was produced in the seventies, in the shadow of Chomskyan ideas produced in the late 1950s and 1960s, and when classical cognitive science dominated. However, an alternative view of cognition emerged in the eighties and has flourished in the nineties. This view is known as connectionism and challenges many of the precepts of classical cognitive science. In contrast to symbolicism, connectionism takes into account the neurophysiological ‘wetware’ of the brain, deriving its chief inspiration from neural networks. The currency of connectionism is not symbols but excitation and inhibition over networks of neuron-like structures. Again, in contrast to symbolicism, thought is not regarded as possessing a ‘syntax’ since explicit linguistic rules need not be mentally represented. Symbolic rules merely approximate the more detailed account of data that is provided by connectionist models. In this thesis, I use elements of connectionism to problematise the symbolic assumptions of CDA, and in turn the symbolic assumptions that affect the criteria for CDA’s isolations of mystifying text. Another enterprise which came to prominence in the eighties and like connectionism continues to flourish is cognitive linguistics.



### 1.6.3 Cognitive Linguistics

Chomsky's concept of linguistic *competence* has been criticised by a number of linguists who claim that we cannot understand the syntax of a language without appreciating how it fits in overall in linguistic *performance*. Cognitive linguistics is a branch of linguistics that denies the autonomy and primacy of syntactic analysis as promoted within Chomskyanism.

In contrast, semantics is seen as interactive with syntax. *Objectivist* semantic analysis that begins with propositions and assesses meaning in terms of truth conditions is repudiated.

Cognitive linguists recommend an analysis of language which aims to provide an account of how both the grammar and meaning are grounded in such factors as the knowledge that the speakers possess, the cognitive models that they construct, and the mappings they make between these models. Furthermore, there are many points of convergence between connectionism and cognitive linguistics which contrast with symbolicism. Both enterprises, unlike symbolicism, do not regard logical reasoning and rule application as fundamental to understanding language and they both relate accounts of mental behaviour to our neurophysiology. In cognitive linguistics this manifests itself in the way in which *basic-level* categories have been studied (see chapter 5). Such categories are seen as being intimately associated with the human capacity for motor-interaction and gestalt perception, and are the easiest type of category to generate mental imagery from.

It is not the overall intention of the thesis to offer connectionism and cognitive linguistics as the absolute 'truth' of language processing although at times I will indicate the superiority of certain aspects of these perspectives over symbolicism. Instead, the purpose of introducing these enterprises, in the early parts of the thesis at least, is to *problematise* the taken-for-granted symbolic assumptions of mental representation in CDA, and thus what CDA highlights as being mystifying text. At other times in the thesis, I provide a more

explicit *critique* of CDA in relation to other attitudes to text processing and its strategies for the highlighting of mystifying text. The critique will sometimes extend to assumptions of text processing which do not necessarily have a direct bearing on how CDA highlights mystifying text. As a final point, I should stress that the enterprises of connectionism and cognitive linguistics are not *wholly* complementary. The Lakoff and Johnson (1980) view of metaphor, I referred to in 1.5, forms part of cognitive linguistic theory. I will indicate, however, theoretical tensions between connectionism and cognitive linguistics which problematise Lakoff and Johnson's approach to metaphor, and then tease out the implications of this for CDA's use of this approach to metaphor.

In 1.3, I indicated that CDA had paid too much attention to the macro-context of interpretation (Foucauldian 'discourse') and not enough to the micro-context of interpretation (Widdowsonian 'discourse' - see below 1.7.2). In this thesis, I want to redress this balance. After problematising in *section B* how CDA highlights mystifying text, in *section C*, I create the following: an alternative framework for the analysis of mystifying discourse in the *micro-context* of interpretation by a non-analytical reader. I discuss the contents of *section C* in 1.7 below.

## **1.7 An Alternative Framework: The Analysis of Mystifying Discourse Produced in the Micro-Context of Interpretation by a Non-Analytical Reader**

### *1.7.1 Developing the Non-Analytical Reader*

In 1.4 and 1.5, I indicated how CDA, either explicitly or implicitly, tries to gauge how non-analytical or non-resistant readers might process text and how portions of a text might be mystifying for them. However, as we saw, the notion of a non-analytical reader in CDA is

undeveloped and inconsistent. In contrast to CDA, section C is centrally concerned with developing a *detailed* and *consistent* processing framework for a non-analytical reader. This non-analytical reader is, broadly speaking, one who has little vested interest in the text and is largely unfamiliar with the text's subject matter. I term this reader, the *idealised reader* (IR). The framework of non-analytical IR draws upon work in connectionism, cognitive linguistics and recent psycholinguistic research on inference generation, the framework consisting of compatible 'principles' from these areas. Like CDA, I do not offer a model of text processing but a set of processing 'principles'. *I use these framework principles to indicate how certain text can lead to mystification in IR's reading.* Or put another way, I use the framework principles to guide my analysis of the *mystifying discourse* (see 1.7.2) derived by non-analytical IR in the reading of certain text. The kind of text I focus upon is *news text*. As I have said, connectionism and cognitive linguistics are diametrically opposed to the tenets of symbolism. Owing to their completely different assumptions about mental representation and cognition, where *my* 'framework' detects text that can lead to mystification in reading is not necessarily coincident with CDA detections. This fact is further reinforced when I show that the recent psycholinguistic evidence I draw upon in my framework conflicts with the *symbolic* assumptions which CDA operates upon in its highlighting of mystifying text.

### *1.7.2 Discourse vs Text*

As I have said, my framework will be concerned with the analysis of *mystifying discourse* (as opposed to *mystifying text*). So, before I go any further, I need to provide definitions for *text* and *discourse*. Following Cook (1994: 24), I define *text* as 'linguistic forms in a stretch of language, and those interpretations of them which do not vary with context.' In reading, *discourse* is created through the interaction of text and context (which can include: co-text,

paralinguistics features, intertext, the physical situation, the social and cultural situation, interlocuters and their schemata). Again, I follow Cook's definition:

'Discourse', as opposed to text, is a stretch of language in use, taking on meaning in context for its users, and perceived by them as purposeful, meaningful, and connected. This quality of perceived purpose, meaning and connection is known as 'coherence'. 'Discourse analysis' is the study and the explanation of this quality of coherence. A discourse *is* a coherent stretch of language.

Defined in this way as a 'perceived' quality, the coherence of a given stretch of language will vary both with its perceiver and with its context. Cook (1994: 25)

These definitions are in keeping with definitions provided by Widdowson (1995a et alibi).

For Widdowson (1995a), addressee discourse is defined as the pragmatic process of meaning negotiation - what is derived from the text through reading in accordance with the purpose and situation of reading. Just to stress this point, for both Cook and Widdowson, discourse varies with perceiver and context so that different discourses can be derived from the same text.<sup>6</sup> To reiterate, my framework is largely concerned with how certain news text leads to mystification in a *particular reading context* - that of a non-analytical reader with little vested interest in a text and one who is unfamiliar with its subject matter. So this is why (following Cook and Widdowson) I am concerned not so much with mystifying text but mystifying *discourse*, or more specifically, the analysis of the *mystifying discourse* that the idealised reader derives from certain news text.

### *1.7.3 Inference Generation as a Part of Discourse*

We saw in 1.4 how appreciation of inference generation amongst different CDA authors was inconsistent. It is also, as I will show in this thesis, an impoverished appreciation. There is, for example, in CDA, little awareness of inference typology in psycholinguistics (e.g. what types of inference are usually automatic) and there is no attempt to indicate how certain

types of inference are less probable than others for readers who have little vested interest in a text. When inference theory *is* drawn upon in CDA (e.g. Fairclough, 1989), the age of the sources in discourse analysis (e.g. Brown and Yule, 1983) means important experimental psycholinguistic work on inference generation which has transpired in the late eighties and early nineties is absent. While this more recent experimental work is absent from reference books in discourse analysis, it is readily available in recent standard reference books in psycholinguistics and cognitive psychology, (e.g. Gernsbacher, 1994; Eysenck and Keane, 1995; Harley, 1995). Another aim of this thesis, then, is to plug this gap in discourse analysis with regard to contemporary experimental research on inferencing in reading.

The anaphoric inference in, for example, ‘*The man* walked into a bar. *He* ordered a drink’ can be taken to be ‘an interpretation which does not vary with context’ and so fits Cook’s (1994: 24) definition of *text* that I gave earlier. However, the inferences I focus on in this thesis are ones whose efficacy are dependent on the degree of vested interest in a text and a reader’s familiarity of subject matter. Because of this variability, in contrast to the anaphoric inference above, the inferences I am interested in are more *discourse*-based rather than *text*-based. I try to show in section C how certain news text would lead a particular reader (IR) to derive a mystifying *discourse* because of the type of inferences this reader would or would not produce from such text. The breadth and detail of the IR framework, I hope, enables not only a much clearer, consistent and more comprehensive perspective on the issue of inference generation / mystification than found in CDA, but also the separation of the occasionally sound CDA from unsound CDA on this issue. As a final point in this paragraph, in 1.4, I outlined some CDA processing assumptions with regard to inference generation which were not directly related to mystification of events. I indicated that reasons for doing so would become clear later. Since a concern of my framework, in section C,

is with recent psycholinguistic research in inference generation, we shall see that this will reveal CDA's assumptions about inference generation, not necessarily related to mystification, to be dubious also.

#### *1.7.4 Using IR to Highlight 'Over-Interpretation' in CDA on Behalf of a Non-Resistant Reader*

In one of his critiques of critical discourse analysis, Widdowson (1995a: 169) argues a basis for discourse *analysis*:

To the extent that critical discourse analysis is committed, it cannot provide analysis but only partial interpretation. What analysis would involve would be the demonstration of different interpretations and what language data might be adduced as evidence in each case. It would seek to explain just how different discourses can be derived from the same text, and indeed how the very definition of discourse as the pragmatic achievement of social action necessarily leads to the recognition of such plurality.

This basis for discourse analysis is similar to what Cook (1994: 25) outlines (see 1.7.2 above). I accord with this basis for discourse *analysis* in section C by showing that different levels of cognitive effort lead readers to produce *inferences* of a different cast from the same text and thus different discourses. I make the assumption that a 'resistant' reader such as a critical discourse analyst makes a good deal of cognitive effort. For CDA, 'non-resistant' readers allow news texts to position them into particular interpretations. I assume, then, that 'non-resistant' readers for CDA make only minimum cognitive effort. Since IR invests minimum cognitive effort, IR could be construed as a non-resistant (or non-analytical) reader. I show the following:

- i) the inferences IR makes from a news text, in line with minimum cognitive effort, are not the same as the inferences a CD analyst makes by proxy for their non-resistant

reader since the latter are more in line with a greater amount of cognitive effort.

- ii) since non-resistant IR produces a different discourse from the CDA non-resistant reader, it does not necessarily follow that a non-resistant reader would be ‘interpellated’ by a news text into the subject position deemed by a CD analyst. That is, with the reading of news text, it cannot be assumed, as CDA does, that the macro- or socio-cultural context weighs so heavily on the micro-context of interpretation of a non-analytical reader.

Overall, from i) and ii), I indicate that CDA is characterised by interpretative gratuitousness from the point of view of ‘non-resistant’ IR and that this ‘over-interpretation’ stems from an impoverished appreciation of the micro-context of interpretation of a non-analytical reader.

Because of the constraints of my thesis, my focus on inferencing related to a *particular reader* and my aim to counter the paucity of appreciation of cognitive matters in CDA, I am unable to devote as much space as I would like to work in text linguistic analysis, i.e., analysis concerned with the formal properties of text. For similar reasons, I am not able to pay enough attention to socio-cultural considerations in reading. However, since IR is a reader with little vested interest in a text, their reading is not then *engaged* in the *social* or *political* sense, mitigating to some extent the absence of attention to socio-cultural considerations in this thesis. Indeed, it is in eschewing a socio-cultural approach to reading, in favour of a cognitive one, that I am able to show that the macro- or socio-cultural context does *not* in fact weigh heavily on the micro-context of interpretation of a non-analytical reader of news text as CDA supposes! As I indicated above, one criticism predominant in Widdowson (1994; 1995a; 1995b; 1996; 1997; 1998) is that CD analysts privilege their own

interpretations from a particular socio-political view<sup>7</sup>. In contrast to CDA, I do not privilege text interpretations in this thesis because of socio-political views. Rather, I *openly* privilege a set of *principles of text cognition*, for my non-analytical reader, these principles guiding my analysis of mystifying discourse.

### *1.7.5 Using Empirical Studies to Support My Framework*

In 1.4.3, we saw how, for Kress (1993: 182), syntactic form could be read in two ways and that this was ‘a hypothesis which could be tested’. Kress did not seek empirical substantiation for this - either his own or others. Other CD analysts I examine in this thesis do not offer empirical substantiation for their claims either. Indeed, one of the difficulties with CDA is that it is adrift from an overt and credible *cognitive performance* theory. Since CDA is sociologically-orientated, there has been little appreciation of matters of cognition and mental representation around the issue of mystifying text or mystifying discourse. In contrast to CDA, the thesis draws upon *current psycholinguistic* empirical evidence, and upon models of *performance* in connectionist and cognitive linguistics, in an effort to check what are often sketchy, unrigorous and speculative notions of cognition in CDA. Drawing upon empirical evidence, albeit not my own, counters the paucity of appreciation of cognitive matters in CDA with regard especially to the issue of mystification in reading, probing more comprehensively than CDA into the relationship between inference generation and mystification in reading. I imagine, though, that some CD analysts may feel that my drawing upon recent psycholinguistic evidence for inference generation in text comprehension, in other words my cognitive emphasis, is irrelevant to their more sociological or *socio-cognitive* concerns.<sup>8</sup> But this would be in tension with CDA’s processing assumptions derived, albeit with modification, from explicitly cognitive (and thus non-sociological sources), e.g Chomsky (see chapter 3) and explicit use of research



developed in artificial intelligence, e.g (Schank and Abelson, 1977) (see chapter 3). Stubbs (1997: 106) criticises CDA for not making use of independent empirical evidence on the relationship between language and thought:

If language and thought are to be related, then one needs data and theory pertinent to both. If we have no independent evidence, but infer beliefs from language use, then the theory is circular.

Unlike CDA, my drawing upon of recent psycholinguistic work on inference generation tries to go some way to satisfying this appeal for independent evidence on the relationship between thought and language. Discourse analysis, of course, is not and never will be ‘prototypical science’. But discourse analysis done by proxy for a non-analytical reader can still be *more* ‘scientific’ if it adheres to the scientific principle that claims be based on empirical evidence as opposed to a brand of discourse analysis based solely on the intuitions of the analyst.

## 1.8 Chapter Outline

The substantive part of the thesis, as I have said, is in three sections. In terms of word length, section A and B form one *half* of the substantive thesis and section C the other *half*:

### SECTION A: *The Relationship Between Symbolicism, Logical Empiricism, the Classical Approach to Categories and CDA*

In *chapter 2*, I outline symbolism, as well as the highly interrelated areas of logical empiricism and the classical approach to categories. In *chapter 3*, I demonstrate how many of CDA’s procedures for locating mystifying text are actually underpinned by symbolic / logical empiricist notions. I indicate how many of these notions have been absorbed



inadvertently into CDA via their ‘borrowing’ of Chomskyan ideas.

*SECTION B: Connectionism and Cognitive Linguistics: Problematizing Symbolicism etc and Implications for CDA*

*Chapter 4* outlines some connectionist models. Discussion in this chapter is then applied to CDA analyses, problematising the symbolic / logical empiricist notion of mental representation on which they rest, and in turn problematising what CDA highlights as being mystifying text. *Chapter 5* outlines the related perspective of cognitive linguistics and how it is in tension with symbolism / classical approach to categories, and consequently also with the ways in which CDA highlight mystifying text.

*SECTION C: Creating an Alternative Framework for the Analysis of Mystifying Discourse*

The chapters in section C build towards an alternative framework for the highlighting of text which is mystifying in reading for the idealised reader. In chapter 6, I refer to recent psycholinguistic inference typologies and indicate the ‘depth’ and ‘shallowness’ of certain types of inference generation for readers with little vested interest in a text. I show how this psycholinguistic evidence also conflicts with the symbolic / logical empiricist assumptions in CDA which govern how mystifying text is highlighted. In chapter 7, I highlight the compatibilities between elements from cognitive linguistics, connectionism and the psycholinguistic evidence that processing with regard to certain types of inference is shallow. In this chapter, I also explore the tensions between connectionism and cognitive linguistics. The tensions have ramifications for Lakoff and Johnson’s view of metaphor which I demonstrate is, ironically, based partially on symbolic assumptions. In problematising the Lakoff and Johnson view of metaphor, in turn, I problematise its use by CDA. In the final sections of chapters 6 and 7, I demonstrate the ‘partiality of interpretation’

principle, showing how different discourses can be derived from the same text. To do so, I use texts examined in CDA. I show how a CD analysis is a discourse based on a high amount of cognitive effort and is one not coincident with the discourse derived from the same text by IR who expends minimum cognitive effort. In so doing, I challenge CDA's claim to exegetic privilege as a form of critical hermeneutics.

In chapter 8, I assemble systematically my alternative framework for the highlighting of mystifying discourse, drawing upon what is highlighted in chapter 7 - compatible elements from connectionism, cognitive linguistics and psycholinguistic evidence for shallow processing. In chapters 4, 5, 6 and 7, in problematising what CDA regard as being mystifying text, I set limits on how far an analyst can make interpretations by proxy for the non-analyst. The opposite occurs, however, in chapter 8. Because my framework does not have the symbolic basis of CDA, and it places a greater emphasis on different types of inference generation, the text my framework highlights as leading to mystification in *reading* would not be highlighted, in the main, by CDA.

## Notes

1. When the 'hypothesis' reading of Whorf was first suggested in the forties and fifties, positivism prevailed. A characteristic of positivism is naturalism - the idea that the methods of the natural sciences can be used in the social sciences. A belief in naturalism, then, helped to demand 'scientific' testing of Whorf's views (see Lucy 1992). Distinct theoretical statements would have been required to facilitate testing and so a sentential approach to Whorf's writings, rather than a co-textual one, would have seemed more appropriate. This would have reduced Whorf's subtlety to some extent. To further facilitate testing, distinct variables would have been necessary too. Given the pervasiveness of the separability of language and thought in positivism and a sentential approach to Whorf's articles, it was without much resistance that 'language' in Whorf's writings became segregated from 'thought' in order to satisfy this need for discrete variables. With language and thought separated, irrespective of the global meaning of Whorf's writings, the background naturalism in positivism also meant that the concept of causality was imposed upon Whorf's work. That is, Whorf was read as positing a unicausal link between language and thought, even though there is no indication of such causality

in his work.

2. Trew (1979: 111-2) acknowledges that reading of significance from a metalanguage is not a straightforward task:

‘No simple one-to-one correspondence can be set up between the linguistic and theoretical processes, because the latter are structured sequences of the former, and can occur in various forms and because individual linguistic changes can occur in different kinds of sequence. A single transformation - like passivisation - does not have a fully determinate theoretical significance. But if it stands as the first in a sequence of changes that include deletion of agents, selective rewording, nominalisation and embedding...then that single linguistic change belongs to a structured sequence of changes, which as a whole has determinate theoretical or ideological significance.’

3. Another enterprise which has affiliations with CDA is Critical Applied Linguistics. This enterprise was initiated by Pennycook (e.g. 1990 and 1994) and has many similar aims to CDA. Differences, though, lie in philosophical leaning. While taking his bearings from CDA, Pennycook argues for a more explicitly postmodern underpinning and distances himself from Fairclough’s (1992a) use of the concept of ideology despite Foucault’s deliberate neglect of the term. With regard to textual mystification, Critical Applied Linguistics has so far little to add and so does not feature in this thesis.

4. For criticisms of CDA’s relationship to ‘critical theory’, see Hammersley (1996).

5. Metaphor is not the only semantic phenomenon highlighted in CDA as leading to mystification. Chilton (1988), for example, makes common ground with Orwell’s ‘Newspeak’ and his famous essay ‘Politics and the English Language’ (1946), highlighting how *euphemism* may prevent readers from actualising the reality of an event. Here is Chilton (1988: 80):

‘The Americans were particularly adept at ‘pacification’ during the Vietnam war, and at many other related activities such as ‘protective reactions’ (bombing raids) and ‘urbanization’ (the destruction of peasant villages).’

Chilton (1988) highlights the mystifying properties of ‘Nukespeak’ [an obvious echo of Newspeak], a generic term for a vocabulary set employed by those with a vested interest in nuclear weapons. Chilton (1988: 80) again in referring to Nukespeak writes that:

‘Military jargon of this type amounts to a semi-secret language: it has a precise meaning for the initiated but constitutes a misleading smoke-screen for the general public. In addition, such terms are basically euphemisms: one may or may not know precisely what they refer to, but either way one is desensitised to the reality of means of distracting associations...And then there is a missile that can carry up to ten times the explosive power of the ‘Little Boy’ [the bomb which destroyed Hiroshima] and is known as a ‘tomahawk’. The effect...is to

minimise the horrific destructiveness of the thing - Indian tomahawks are, after all, scarcely more than playthings.’

The power, though, of euphemisms works of course according to the degree of encyclopaedic knowledge people have. If one knows the firing of ‘tomahawks’ can lead to devastation, they are not then ‘playthings’.

With this knowledge, ‘tomahawk’ - the axe and ‘tomahawk’- the missile lose any polysemic connection they may have had, becoming distinct homonyms.

6. For the sake of clarity, I think it is worth quoting Widdowson (1996: 58) at length on his definition of discourse and how this contrasts with the definition of discourse in CDA, which I referred to in 1.2:

‘Norman Fairclough complains that discourse analysis, on my account, ‘is reduced to pragmatics’. Why *reduced*? Discourse, in my conception of it, cannot be *reduced* to pragmatics. It is, for me, crucially, a function of pragmatics: the process whereby different interpretations are drawn from the textual data. Of course this process implicates all manner of social factors: assumptions, beliefs, values, ideologies, which would fall within a Foucault concept of discourse. So it would seem that we need to distinguish two senses of discourse (see Widdowson, 1990). The discourse process in this pragmatic sense (Discourse 1, we might call it) is influenced by the different discourses that participants have been socialised in (Discourse 2). But it is not *determined* by them. To suppose that it is *is* indeed to reduce discourse. And this is probably where my position differs most radically from that of Norman Fairclough and his colleagues. So let me make it as explicit as I can.

I do not believe that individuals simply act out social roles. There are socially constituted Discourses 2 - conventions of belief, established values which constrain the way people think and use their language to achieve meaning. But people’s activities are not determined by their ideological allegiances. They are not bound by them. You can of course ascribe social roles to individuals and part of their individuality can obviously be associated with this group identity. Pragmatics would take account of this in its consideration of contextual conditions. But to think of individuals as if they were *representative* of such groups, as tokens of the type, is to deal in stereotypical constructs, well defined social categories.’

And Widdowson (1996: 59):

‘...individuals are constrained, but they remain individuals none the less: they are not just ‘subjects discursively constituted’. This is not to deny the existence of discourses in the Foucault sense as conventionalised modes of knowledge (i.e. Discourse 2), nor the importance of studying the discursive construction of social subjects at an appropriate level of idealisation. But these discourses are abstract constructs. They can only be actualised through discourse as I have defined it, as the pragmatic process of meaning negotiation (i.e. Discourse 1).’

7. The privileging of interpretation can be seen in Fairclough’s (1989: 52) analysis of the Jenny Keeble text in 1.4.1. Fairclough conflates the abstract qualified category ‘good wife’, and its specific meaning in the text. The abstract and general ‘good wife’ has often been used to refer to wives who consciously or sub-consciously

agreed to: be house-bound, be (mostly) responsible for raising children, have no full-time job, be completely responsible for domestic chores etc - as such a willing thwarter of her own potential for the sake of her husband. But there is little textual warrant here that Jenny Keeble is that kind of 'good wife' and thus it is questionable whether this text would reproduce *sexist discourse* (i.e. discourse 2 - see note 6) in the way in which Fairclough argues. The general category of 'good wife' is sufficiently reduced here to be considered reasonably appropriate for any 'good partner' if their 'significant other' were at war. It is because Fairclough adopts the exegetic practice of reading a general category into text specificities - conflating the macro-context with the micro-context of interpretation - that he reaches the conclusion he does. A need to read the macro-context into the micro-context becomes patent when we see that some of the 'attributes' of the 'good wife' he refers to come from *Jenny Keeble's lips* rather than a 'sexist' journalist producing a sexist text.'

8. Occasionally, statements delimiting psychological considerations are made in CDA. Here, for example, are Chouliaraki and Fairclough (1999: 68):

'We do not see CDA as a theory specifically of the relation between cognition and text.'

## **Section A: The Relationship between Symbolicism, Logical Empiricism, the Classical Approach to Categories and CDA**

This section comprises two chapters. In chapter 2, I outline symbolism and areas related to symbolism - logical empiricism and the classical approach to categories. In chapter 3, I show how much of CDA, and especially how CDA highlights mystifying text, is underpinned by symbolic, logical empiricist and classical assumptions.

## CHAPTER 2: OUTLINE OF SYMBOLICISM, LOGICAL EMPIRICISM AND THE CLASSICAL APPROACH TO CATEGORIES

### 2.1 Introduction

Most of this chapter is given over to an outline of the *symbolic* architecture of the mind or *symbolicism* for short. I also deal with highly related areas including: logical empiricist philosophy of language, the ‘classical’ theory of categories and ‘syntax-first’ approaches to processing. I show how many of Chomsky’s ideas are consonant with the above areas. I shall then go on in chapter 3 to show how:

i) CDA has inadvertently absorbed many symbolic postulates into their perspective on language processing through their ‘borrowings’ from Chomsky.

ii) postulates within logical empiricism can be found to underwrite some assumptions of language processing for some CDA authors. Some of these logical empiricist postulates have been absorbed through ‘borrowings’ from Chomsky.

iii) the classical theory of categories can be found to underwrite some assumptions of processing for some CDA authors.

In showing i), ii) and iii), I highlight in chapter 3 how CDA’s *isolations of mystifying text* are influenced by symbolic / logical empiricist / ‘classical’ assumptions.



## 2.2 Symbolic Architecture

### 2.2.1 Turing Machines and von Neumann Machines

Symbolic modelling of the mind within cognitive science is based on the idea that mental processing consists of the manipulation of symbols that can be transformed according to rules. These symbols are held to be *retainable* and thus an *enduring* set of entities. In symbolic architectures of the mind, mental representations are viewed as semantically interpretable, structured objects consisting of symbols which have parts. As exemplification of this notion, here is Cooper (1996: 28):

The symbol '34' for example, has parts (the symbols '3' and '4'), and the meaning of '34' is a function of the meaning of '3' in the tens position and '4' in the units position. The arabic representation of numbers, then, is a structured, semantically interpretable representation.

And for *linguistic* cognition more specifically, this set of unequivocal algorithms constitutes a *syntax* - i.e. *a set of linguistic rules which formally specify a set of operations on linguistic symbols*.

The provenance of symbolic modelling lies in the computational theory developed by John von Neumann (1947) and Alan Turing (1950) and has been incorporated into nearly all existing electronic computers. Sketching out the nature of what has become known as a *Turing machine* will help understanding of the essence of symbolic modelling. A Turing machine is not in fact an actual machine at all but an abstract model. It has an infinite memory in the form of an infinite strip of tape. It also consists of a type-writing device which can type a symbol on the tape, erase such a symbol, and move left and right along the tape. By specifying exactly what symbols the typewriter uses, and how it should respond

to them as it passes along the tape, the Turing machine converts one set of symbols and spaces ('input') into another ('output'). For Turing, the machine was seen as extremely powerful since theoretically it could transform any input into any output given some computable relation between them. That is, a Turing machine can execute any algorithm (a set of computable instructions). However, the power of the Turing machine lies in the abstract. The blueprint for most standard modern computers was in fact established by John von Neumann. Processing in a von Neumann machine is serially-based on a fetch-instruction / execute-instruction cycle. The machine fetches the current instruction and then moves the instruction indicator to the next directive. It executes the instruction it has just retrieved, fetches the next directive and so on. In a strict sense, von Neumann machines are less powerful than a Turing machine because they have finite memory. However, in practice and particularly with modern computers this is rarely a problem.

### 2.2.2 Turing Machines, Symbolicism and Philosophy: General

The symbolic notion of a Turing machine has been influential in the philosophy of mind and cognitive science. One philosopher influenced by the concept of a Turing machine is Putnam (1975). For Putnam (1975: chpts. 18, 20, 21) mental cognition is understood in terms of the manipulation of symbols via the computation of a set of unequivocal algorithms and without recourse to the neurophysiology or 'wetware' of the brain. But in many ways, the Turing machine is *itself* a distillation of well-known perspectives in logic and philosophy which view the mind most prominently in its capacity to reason and where reason is characterised as the algorithmic manipulation of abstract symbols (Bechtel and Abrahamsen 1991: 8). An example of an algorithm within *logic* is the simple inference rule *modus tollens*: from one proposition of the form, if  $p$  then  $q$ , and another of the form  $\sim q$ , we can

infer the proposition  $\sim p$ . Within *philosophy* also, the idea that human cognition is based upon symbolic manipulation has been a recurrent one that cuts across the old division between empiricism and rationalism (Lakoff, 1987a: 164). It is found in Hobbes ([1651] 1962: 41) who regarded reason as analogous to mathematical computation. Rationalists such as Descartes and Leibniz and empiricists such as Locke and Hume helped to further establish this view of cognition as consisting of rule-governed logical manipulation (Bechtel and Abrahamsen, 1991: 10). (See also 2.2.4 on logical empiricism below for a more specific outline of the relationship between symbolicism and a philosophical position).

### 2.2.3 Assumptions about the Mind-Brain in Symbolicism

One of the ancient ‘problems’ in philosophy is what is known as the *mind-brain problem*. Reduced to its essence, the mind-body problem revolves around a dichotomy. Is the mind a ‘ghost in the machine’, composed of non-corporeal material, or do mental phenomena (beliefs, intentions, etc) have physical instantiations? The positions that answer these questions in the affirmative are known as *dualism* and *materialism* respectively. Well-known dualists have included Plato and Descartes and more recently Eccles (1977) and Swinburne (1986). Prominent materialists have included Hobbes and more recently, Skinner (1957, 1976), Smart, (1959), Quine (1960), Dennett (1978) and Churchland, P.M. (1988). As a consequence of a rigid separation between mind and body, dualists regard psychology as independent of neurobiology since the latter is concerned much more with mental ‘wetware’ and much less with the mental phenomena it ‘supports’. Indeed, dualism is a *commonplace* assumption about the mind, supported by a ‘folk-psychology’ vocabulary such as ‘beliefs’, ‘attitude’, ‘hopes’ etc. Since dualism is such a commonplace notion it makes the study of mind, in the absence of biological considerations or widely accepted

evolutionary principles of selection or descent, appear so natural.

Now for Putnam (1967) (as well as Fodor (1975) and Pylyshyn (1984)) mental states are algorithmic. They need not be implemented in brains, thereby reducing the importance of locating mental states with brain states. Thus symbolicism ties in neatly with philosophical dualism and ‘common-sense’ dualism, naturalising the study of the mind independent of the brain’s wetware. Indeed, computational machines based on the logical foundations of Turing machines or on von Neumann architecture for a digital computer have come to be regarded for many as sound models for the understanding of mind-brain functions. In chapter 4, I outline an approach to cognitive modelling known as connectionism. In contrast to symbolicism, connectionism explicitly regards mental states as being dependent on neurophysiological ‘wetware’.

I have given a brief sketch of the relationship between symbolicism and philosophy, and how symbolicism entails dualism and thus sanctions the study of mind independent of brain ‘wetware’. I mentioned briefly that, in symbolicism, linguistic cognition involves a set of unequivocal algorithms which constitutes a *syntax* - i.e. a set of linguistic rules which formally specify a set of operations on linguistic symbols. Section 2.2.5 will give more attention to linguistic issues in symbolicism. But firstly in 2.2.4 below, I examine the *linguistic* philosophical position known as *logical empiricism*. I then go on (in 2.2.5) to show how logical empiricism has exerted influence on how *linguistic* mental representation has been cast within *symbolic* cognitive science. Following this, in 2.3, I show the symbolic / logical empiricist influences on the Chomskyan paradigm upon which CDA borrows. As we shall see in chapter 3, in adapting Chomsky, CDA has absorbed many symbolic / logical empiricist postulates, these postulates influencing how CDA highlights mystifying text.

### *2.2.4 Logical Empiricism<sup>1</sup>*

#### *Orientation*

Logical empiricism was a branch of philosophy which is mostly associated with members of the Vienna Circle of the 1920s and 30s. Such members included Rudolph Carnap, Herbert Feigl, Otto Neurath and Moritz Schlick. It was so called since its practitioners sought to apply logical methods to the world of empirical experience. Thus, logical empiricism can be viewed as something of an amalgam of the traditional poles of empiricism and rationalism. Their demarcation criterion of verificationism, whether propositions could be empirically tested or not, adjudicated between scientific knowledge and metaphysics or 'meaninglessness'. Logical empiricism espoused physical reductionism, that the propositions of sociology, for example, could be analysed into those of physics. Logical empiricism was also nominalist and as such opposed to theoretical entities postulated to account for the physical world. This nominalism, when applied to the social world, emphasised methodological individualism over holism, and thus avoided any postulating of a social ontology behind abstract social categories. Not only this but any lapse into a hermeneutics of social action was regarded as metaphysical also.

#### *Russell*

The heritage of logical empiricism lay in work on the interface of logic, mathematics and philosophy by Gottlob Frege, Bertrand Russell, Alfred North Whitehead and the early work of Ludwig Wittgenstein in the initial part of this century. Russell and Whitehead (1910) aimed to derive mathematics from the fundamentals of logic. Their end-product, the 'Principia Mathematica', was to exercise a large influence on Anglo-American-Austrian philosophy in the first few decades of this century. For Russell, the real world was made up

of ‘facts’ which are essentially *atomic* in nature, i.e. cannot be reduced, and the ‘Principia Mathematica’ provided the basis of a perfect logical language because it mirrored the structure of the real world. Russell’s perspective became known as *Logical Atomism*. Atomic propositions by definition had a subject-predicate structure. Russell inherited the idea of a sentence being a fundamental unit from Frege, and indeed in logical empiricism, as a whole, emphasis was placed upon *syntactic* structure. For Russell, atomic propositions were largely analysed in terms of truth-functionality. But for an atomic proposition to be *meaningful*, i.e., to be true or false, the *subject* should denote an individual entity and the *predicate* refer to some quality of that entity. By these criteria, ‘Tony Blair is Prime-Minister of the UK in 1999’ is a *valid atomic proposition*. But, what of a sentence such as ‘the present king of France is bald’ uttered in 1999? The structure may be that of subject-predicate, but the subject does not have denotation and so the sentence cannot be treated in a truth-functional manner. Such a sentence poses problems for Russell’s logical atomism since by his criteria it is not meaningful.

For Russell, mathematical logic could provide philosophy with a set of techniques for clarifying such a sentence. Russell held that there was an *underlying* logical structure to the sentence ‘the present king of France is bald’ that, once revealed, would show that the sentence is meaningful after all. Russell contended that the underlying logical form of ‘the present king of France is bald’ is actually quite complex and contains the following assertions:

There is an entity  $x$ , such that:

- a)  $x$  has property K [K = ‘king of France’]
- b) there is no other entity  $y$  which is distinct from  $x$  and has property K
- c)  $x$  has property B [W = ‘bald’]

In logical form this becomes:

$$\exists x [ Kx \quad \& \quad \sim \exists y ( (y \neq x) \& Ky ) \quad \& \quad Bx ]$$

This can be paraphrased as the conjunction of three propositions, or in Russell's parlance, a molecular proposition consisting of three atomic propositions:

There is a king of France	<i>and</i>
There is no one else who is king of France	<i>and</i>
The king of France is bald	

In expanding 'the king of France' to include the existential clause, 'there is a king of France', in predicate calculus, Russell was able to claim that the sentence 'the king of France is bald' was now meaningful. Why did Russell think this? In a truth table for a conjunction, if one of the conjuncts is false, then the truth value for the conjunction is also false.<sup>2</sup> On Russell's analysis, since the first conjunct ( $\exists x Kx$ ) is false, the *whole* of the logical form of the 'king of France is bald' is also *false*. Because Russell viewed the meaning of a sentence to a large extent in terms of whether truth conditions could be ascribed to a sentence, he could then regard 'king of France is bald' as meaningful; (see Levinson (1983: 171) for more details).

### *Wittgenstein*

Russell's student, Ludwig Wittgenstein, sought to develop his tutor's perspective in the 'Tractatus Logico-Philosophicus' (1921). He too inherits from Frege the notion that the fundamental unit of meaning is not the word but the sentence. The basis of the 'Tractatus' was a logical formalising of a tri-partite correspondence between thoughts, sentences, and object relations in the world. Like Russell, there is the desire in Wittgenstein to clarify

*linguistic* descriptions of the world by reducing them to their underlying propositions, simples with a rudimentary logical syntax. For Wittgenstein, the logical structure of simples was *concealed* by the ordinary language form of a sentence. Another feature of these ‘simples’ was that they were logically independent of one another. In the *Tractatus*, Wittgenstein expressed the belief that if meaningful discussion was to transpire then a sentence would not only have to represent reality, but the sentence and the state of affairs it represents must have a common structure. The corollary of this, for Wittgenstein, is that since the structure of the language must mirror the structure of the world, then it is possible to discern the structure of the world by analysing the structure of sentences. From this, Wittgenstein derived his ‘picture theory of meaning’, where language was a picture of how facts about the world were structured and so also a ‘picture’ of how the world itself was structured.

### *Carnap*

In turn, the philosophers of the ‘Vienna Circle’ were heavily influenced by the ‘*Tractatus*’, one member, Rudolph Carnap, believing also that much of philosophy was reducible to concerns of logical syntax. Indeed, for Carnap, by refraining from syntactic errors through the employment of logical analysis, a philosophical conundrum could be solved or shown to be insoluble (Carnap, [1928 ] 1967). For Carnap, philosophy had to consist in the logic of science, which was itself identified with the logical syntax of a scientific language. Carnap (1967) distinguished three kinds of sentences: *object*, *pseudo-object*, and *syntactical*. Examples of *object* sentences are ‘5 is a prime number’ and ‘Babylon was a big town’. Examples of *pseudo-object* sentences are ‘Five is not a thing but a number’ and ‘Babylon was treated of in yesterday’s lecture’. Examples of *syntactical* sentences are ‘Five is not a thing-word, but a number word’ and ‘The word ‘Babylon’ occurred in yesterday’s lecture’.



Carnap also made a distinction between *material* and *formal* modes of speech. *Object and pseudo-object* sentences were regarded as belonging to the *material* mode of speech and *syntactical* sentences to the *formal* mode. The point I want to highlight here is that for Carnap, pseudo-object sentences were so called because they were held to be syntactical sentences masquerading as object-sentences. As such, they were said to be syntactical sentences expressed in the *material* mode. Translating them into the *formal* mode revealed their syntactical character. It was Carnap's ambition to show that the respectable propositions of philosophy, as commonly formulated, were *syntactical* propositions of philosophy misleadingly expressed in the *material* mode; see Ayer (1982) for further discussion.

### *Compositionality*

One prominent notion within logical empiricism that is a given with the above thinkers is that meaning is *compositional*. Essentially, this means that logical empiricism operates on a 'building-block' attitude to meaning. Within logical empiricism, the world is regarded as being constituted by objects with *well-defined* inherent properties and there are *fixed relations* between objects at any given time. A world constituted by *well-defined objects* can be ascribed *discrete* names. Since these objects have well-defined inherent properties, then each of these properties can be ascribed a one-place predicate corresponding to each of those properties. And since the objects stand in *fixed relations* to one another, then a series of many-place predicates can be ascribed so as to correspond to each relation. The natural consequence of such a perspective is to regard meaning atomistically - 'in building blocks'. From the 'building block' view of meaning within logical empiricism arises a *correspondence theory of truth*. Here are Lakoff and Johnson on the logical empiricist

perspective on objects, properties and relations I have just outlined (1980: 202-3):

Assuming that the world is this way and that we have such a language, we can, using the syntax of this language, construct sentences that can correspond directly to any situation in the world. The meaning of the whole sentence will be its truth conditions, that is, the conditions under which the sentence can be fitted to some situation.

*The Removal of the Human Understander in Logical Empiricism*

However, for a sentence to *mirror* or *successfully correspond* to a situation in the world, it must *fully* represent objects, properties and relations, independent of the contribution to processing of the human ‘understander’. Lakoff and Johnson once more on logical empiricism (1980: 202):

The meaning of the whole sentence will depend entirely on the meanings of its parts and how they fit together...every sentence of the language **must contain all of the necessary building blocks** so that, together with the syntax, nothing more is needed to provide the truth conditions of the sentence. The ‘something more’ that is ruled out is any kind of human understanding. [my bold]

Indeed, Carnap’s attempt to produce a logical syntax that could be applied to the problems of philosophy consisted of these building-block properties with little consideration given to the contribution of the human understander.

I now go on to show how the thread of ideas that runs from Russell through to the Vienna Circle has exerted an influence upon the *symbolic* paradigm in cognitive science.

*2.2.5 The Influence of Logical Empiricism upon the Symbolic Paradigm*

As support for the view that the ideas of logical empiricism have exerted an influence upon

the symbolic paradigm, here first is Harder (1997: 52):

The basic problem with the classical computational approach was that it smuggled a number of assumptions associated with logical positivism into the supposedly new mental framework, instead of facing the challenge of actual mental phenomenon.

Harder goes on to say that this challenge is met by cognitive linguistics, an enterprise to which chapter 5 is devoted. Now consider the following from Gardner (1987: 64-5):

...a major ingredient in ongoing work in the cognitive sciences has been cast in the image of logical empiricism: that is, the vision of *syntax* - a set of symbols and the rules for their concatenation - that might underline the operations of the mind (and a correlative discomfort with issues of mental *content*)... Thus, when Noam Chomsky (1965) posits the basic operations of grammar, when Richard Montague (Thomason 1974) examines the logic of semantics, when Allen Newell and Herbert Simon (1972) simulate human reasoning on a computer, or when Jerome Bruner (1973) and George Miller (1956) seek to decipher the rules of classification, or 'chunking', they are trying to decipher a logic - perhaps *the* logic - of the mind. This vision comes through even more clearly in the writings of Jerry Fodor, who explicitly searches for a 'language of thought' and even appropriates certain of Carnap's methods. Thus, a model that proved inadequate for the scientific enterprise as a whole still motivates research in circumscribed cognitive domains.

Let me pick up on Gardner's point that logical empiricism is implicit within Fodor's 'language of thought' hypothesis. Fodor is a thinker firmly entrenched within the *symbolic* camp. Fodor (1975) offers the view that mental representation occurs within a 'language of thought' (often known as *mentalese*) where mental representations encode *propositional information via a language-like syntactic structure*. The connection between Fodor's *mentalese* and logical empiricism should be apparent. This connection is also evident in a more recent articulation of Fodor's position in Fodor and Pylyshyn (1988). Both Fodor and Pylyshyn (1988) stress the 'syntactic' nature of thought, maintaining that structured symbols and computation which is structure-sensitive are key elements in mental cognition. Human cognition is wired to assemble complex symbols from rudimentary ones. On the basis of

this notion, they hold, similar to the *logical empiricists*, that thought is *compositional*. So, for Fodor and Pylyshyn, a thought's meaning is regarded as a function of the meaning of its parts, molecular representations being formed out of its constituents. The thought that 'David is tall' is, for instance, a function of the meaning of the thought 'David' and the meaning of the thought 'is tall'. Patently, another logical empiricist echo here is of Russell's logical atomism where atomic propositions are more primitive than molecular ones.

Another characteristic of human cognition that Fodor and Pylyshyn highlight is the idea that mental representations are *systematic*; systematic exchange can occur in mental representations to yield new representations. So if a system contains a representation  $xRy$ , where R is a relation and  $x$  and  $y$  are variables of the same syntactic type, then the system has the ability to produce the representation  $yRx$ . For example, if we can have the thoughts that 'David loves Mary' then we can also have the thought that 'Mary loves David'. For this to take place, 'the two mental representations, like the two sentences, must be made of the same parts' (Fodor and Pylyshyn, 1988: 39). Again, like Russell's conception, the referents are handled via syntactic concerns rather than semantic ones. Thirdly, Fodor and Pylyshyn indicate that mental representations are characterised by *productivity*. What this means is that, despite a finite number of simple concepts, humans can produce an infinite set of propositions. Fodor and Pylyshyn reason that a set of infinite propositions which can nevertheless be systematically related, as with the above example, can only be explained on the basis that they are constructed via repeatable units. Finally, Fodor and Pylyshyn also hold that syntax is autonomous and effectively operates an apartheid between semantics and syntax. Algorithmic operations are viewed as syntactic and are applied without recourse to the semantics of the symbols. *This emphasis on syntactic operations is very much a characteristic of symbolism.*

As I intend to show in chapter 3 how CDA notions of mental representation are essentially symbolic / logical empiricist, partly at least because they were derived via Chomskyan ‘borrowings’, I need now to manifest the essentially symbolic / logical empiricist nature of Chomsky’s work. This was already hinted at in the quote from Gardner (1987: 65). Since ‘borrowings’ of Chomsky in CDA do not go beyond 1969, the scope of this examination is restricted to Chomsky (1957) and (1965), the two most cited works, for example, in Hodge and Kress (1993: 35-7).

## **2.3 Chomsky and Symbolicism / Logical Empiricism**

### *2.3.1 Synopsis of Chomsky’s Early Positions*

In this section, I sketch the development of Chomsky’s thought so as to make evident affinities between Chomsky and the symbolic paradigm. I start with Chomsky’s famous review of B.F. Skinner’s *Verbal Behaviour*. In this review, Chomsky (1959) argued that a behaviouristic explanation could not explain the human capacity to learn a language. Chomsky highlighted (i) how any natural language has an infinite number of syntactically well-formed sentences, and (ii) the fact that speakers can understand and produce sentences that they had not previously encountered (Chomsky 1957). Chomsky maintained that the output of an infinite set of well-formed sentences could not be stimulated by a set of finite learned associations between environmental stimuli and linguistic responses (the ‘poverty of stimulus’ argument). Instead, Chomsky proposed that the potential for producing an infinite set of well-formed sentences exists because of an innate capacity rather than one which relied upon environmental stimulation. The approach that Chomsky offered to replace behaviourism became known as generative grammar. In trying to establish the

notion of a generative grammar, Chomsky aimed to show that the current models for explaining the generation of an infinite set of well-formed sentences were inadequate for the task. The two models which Chomsky isolated as being prominent and at the same time problematic were *finite-state grammar* and *phrase-structure grammar*. I shall consider each in turn.

Chomsky turned his attention to *finite-state grammars* since language had been considered from this perspective 'in connection with the design of efficient channels of communication during the Second World War; and the highly sophisticated mathematical theory of communication that resulted ('information theory') was extended to many fields, including psychology and linguistics, after the war' (Lyons 1991: 55). A finite-state grammar is predicated upon the view that sentences are generated via a selection procedure made 'from left to right'. After the first linguistic element has been chosen, subsequent selections are made on the basis of previous elements. Take the sentence 'this man is running'. The word 'this' would be selected from the set of all words occurring at the beginning of a sentence, 'man' would be chosen as a possible word to follow 'this' (rather than 'men' say) and so on. In *Syntactic Structures* (1957), Chomsky disregards finite-state grammars since they cannot capture recursion. Consider the sentence 'The man who was wearing the trilby started to wave at me'. For Chomsky, a finite-state grammar cannot capture the syntactic dependency between 'man' and 'started to wave' which endures across the intervening defining relative clause. Furthermore, a finite-state grammar cannot cope with clausal embeddings that can recur interminably (there was an old woman who swallowed a fly, that was eaten by a toad, that was eaten by dog etc...). Even though such sentences become more and more difficult to cognise, for Chomsky they are technically grammatical; that is, any grammar must be able to account for or generate them.

I turn now to the other grammar that Chomsky considered problematic - *phrase-structure grammar*. This grammar consists of a fundamental set of strings (a sequence of symbols) and a finite set of phrase-structure rules which operate upon the fundamental set to 'rewrite' it into another permissible form. Chomsky argued that such a grammar was not only maladroit but failed to account for speakers' intuitions that strings like 'David hit Bill' and 'Bill was hit by David' bear a close relationship to one another. With its roots lying in the work of his former tutor, Zellig Harris (1952), in *Syntactic Structures*, Chomsky laid the beginnings of a grammar which aimed to both capture recursion and account for speaker intuitions regarding sentential relationships such as that between the active and passive voice. This grammar became known as 'transformational grammar'. In a transformational grammar, Chomsky contended, a matrix of rules is established so that separate strings can be related, and where one sentence, or to be more accurate - the sentence's abstract mental representation - can be transformed to generate sentences. Such a transformational or generative grammar is a set of algorithms of mathematical exactitude which generates sentences in a language without utilising any information extraneous to the system.

In *Syntactic Structures*, Chomsky begins with phrase-structure rules which enable the generation of the core sentences, or kernel sentences. The kernel sentences are generated via a set of algorithms such as the following (which are not intended to be exhaustive):

- |    |          |   |                                     |
|----|----------|---|-------------------------------------|
| 1. | Sentence | → | NP + VP                             |
| 2. | NP       | → | (Det) + (A) + N                     |
| 3. | VP       | → | V + NP                              |
| 4. | N        | → | e.g. boy, girl, dog, cat, ice cream |
| 5. | V        | → | e.g. eats, likes, bites             |
| 6. | A        | → | e.g. happy, lucky, tall             |
| 7. | Det      | → | e.g. a, the, one                    |

So taking the symbol S as the starting point, a kernel sentence such as ‘The boy eats the ice cream’ can be generated. From this platform, the other grammatical sentences of the language can be generated not via the transformation of these kernel sentences, but via the transformation of a common underlying string. Since transformations are algorithmic in nature and operate serially, conversion from one underlying string into another follows an established sequence. Transformations were also believed to resolve *structural* or *formal ambiguities* such as are extant within ‘Jimmy saw the girl with the binoculars’ through the exhibition of different transformational procedures - one where the PP ‘with the binoculars’ is a post-modifier in the NP, ‘the girl with the binoculars’, and another where ‘with the binoculars’ is a PP constituent of the VP ‘saw the girl with the binoculars’ and not of the NP ‘the girl’. In 1965, Chomsky tendered a new expression of transformational grammar in the form of the ‘Standard Theory’ in *Aspects of the Theory of Syntax*. The notion of kernel sentences was discarded. Instead, the concept of ‘deep structure’ was introduced where a transformational element mutates into other structures, the ultimate being the ‘surface structure’.

### *2.3.2 Affinities Between Chomsky’s Early Position and Symbolicism / Logical Empiricism.*

Clearly, there are affinities between Chomsky’s approach to language and that of the logical empiricists: the concentration on syntax, the exhibition of underlying structure in natural language, kernel sentences as simples (akin to Russell’s atomic propositions in their subject-predicate structure, declarative rather than interrogative or imperative, active rather than passive etc) the application of formal mathematics to a natural language like English (though in Chomsky’s case this was ‘finite automata theory’ and ‘recursive function theory’) etc. Indeed, Hacking (1975: 91) goes as far as to say that:



Russell's idea of logical form as opposed to grammatical form is strikingly like Chomsky's idea of depth grammar as opposed to surface grammar.

And Lakoff (1987a: 226), referring to the Chomskyan paradigm but without naming it specifically, summarises the influence of mathematics and logic on logical empiricism in turn influencing Chomsky's view that languages are viewed principally in terms of formal syntax separate from semantics:

How did it come about that philosophers, linguists, and even many cognitive psychologists have come to view natural human languages in terms of formal syntax and formal semantics?

The principal reason was the rise of mathematical logic, the enormous prestige that it acquired, and the fact that it was taught in European and American universities by objectivist philosophers, who viewed it as the study of reason. When logic was turned into a form of mathematics by Frege, Russell, Hilbert, and others, the axiomatic method was adopted into logic itself..

The formalist program of separating syntax from semantics accompanied the mathematicization of logic and the unification of logic with mathematics. The separation was needed in order to make sense of axiom systems.

Through the influence of Bertrand Russell, British and American philosophers eventually adopted the objectivist equation of reason with mathematical logic. Along with that development came the idea that natural languages also had a division between syntax and semantics, with syntax being a matter of uninterpreted symbols and semantics providing a separate interpretation. To objectivist philosophers trained in mathematical logic, the division came to be seen as natural.

Ties between Chomsky and the symbolic paradigm of cognitive science include the algorithmic nature of transformations, the serial manner of their operation. As Gardner argues (1987: 188) 'clearly, Chomsky was a child of the new era of Wiener, von Neumann, Turing and Shannon...'<sup>3</sup> The cognitive linguist Langacker (1987b: 6) concurs in regarding Chomsky as a progeny of the first generation of artificial intelligence:

...linguistic theory in the generative tradition presupposes the von Neumann architecture, accepting without question the need for discrete and explicit rules couched in some 'propositional' format, and which constitute an algorithm specifying the sequential manipulation of abstract strings of symbols.

and the symbolic grounding of Chomskyan thought is attested to by Brown (1991: 491):

... although in *Syntactic Structures* Chomsky was very concerned to explore the mathematical properties of PS [phrase-structure] rules, little attention was devoted to the mathematical power of transformations. Once the mathematical properties of this kind of rule were explored, *it became clear that a grammar with transformations has the formal properties of a universal Turing machine*: in other words, they are such a powerful tool that they can explain nothing except that language can be described in terms of some set of rules. [my italics]

I have established that Chomsky's perspective has been influenced by symbolic cognitive science and logical empiricism - both highly interrelated areas. Both these areas are so interrelated that often when I use the term *symbolic* in this thesis, it should be obvious that I am referring to logical empiricism as well. In chapter 3, I show how many of the language processing assumptions of CDA and in particular how CDA highlights mystifying text are predicated upon symbolic / logical empiricist postulates. I indicate how such postulates have been absorbed, in part at least, via CDA's use of *Chomsky*.

In the next section, I want to outline the classical approach to categorisation which is related to both symbolism and logical empiricism. For instance, as we shall see, Chomsky's thinking is bound up with it. Lakoff (1987a) refers to the classical approach to categorisation as *objectivism*. This is because it does not include human understanding, being *objectively* apart from it. In this respect, the classical approach to categorisation dovetails also with logical empiricism (2.2.4). In chapter 3, I will indicate how the classical approach to categorisation informs some aspects of how CDA highlights mystifying text.

## 2.4 The Classical Theory of Categories

### 2.4.1 Aristotle

What is known as the classical approach to categories is synonymous with Aristotle's ideas and it is from its roots in antiquity that this theory of categories derives its name (Taylor, 1995: 22). In 'Metaphysics', Aristotle draws a distinction between the *essence* of a thing and its *accidents*. The *essence* is 'all parts immanent in things which define and indicate their individuality, and whose destruction causes the destruction of the whole' (Metaphysics 5.8.3); that is, the essence is what constitutes and thereby defines the thing. *Accidents* are on the other hand what may be called incidental properties and are not instrumental in divining what a thing is; 'Accident' means that which applies to something and is truly stated, but neither necessarily nor usually' (5.30.1). As illustration, here is one of Aristotle's examples: man's essence is 'two-footed animal'. The properties of whiteness, being cultured etc are merely accidental since they do not determine whether a thing is a man. For Aristotle, both the concept MAN and the meaning of the word *man* are defined by a 'formula ('logos') of the essence' (7.5.7):

If 'man' has one meaning, let this be 'two-footed animal'. By 'has one meaning', I mean this: if X means 'man', then if anything is a man, its humanity will consist in being X (4.4.8).

To attach the predicate 'is a man' to an entity, the meaning of the word man must be divined, that is, the essence of man must be known:

If anything can be truly said to be 'man', it must be 'two-footed animal'; for this is what 'man' is intended to mean (4.4. 14-15).

In modern parlance, Aristotle regards ‘two-footed’ and ‘animal’ as *necessary* features and jointly the two features are deemed to be *sufficient*. Indeed, the notion that categories can be defined in terms of discrete necessary features which, taken as a set, are sufficient is an assumption of the classical approach. This assumption entails an absoluteness to the issue of category membership. Any entity which manifests the number of necessary features required for sufficiency is *ipso facto* a member of the category. Vice-versa, if any one of the necessary features for sufficiency is absent then category membership is denied. There are other aspects of Aristotelian theory which conjoin with the above assumption. These derive from the ‘law of excluded middle’ and the ‘law of contradiction’ (Metaphysics 4.4). The law of contradiction states that a thing cannot be and not be, and the law of excluded middle states that a thing must either be or not be. What follows from these ‘laws’ are that features are binary, indicating presence [+] or absence [-] since categories have clear unambiguous boundaries. One other aspect that follows from this position is that all members of a category have equal status. In other words, if three necessary features are required for sufficiency in X, and only two are present in Y then Y cannot be less of an exemplar instance of X but rather is barred from category membership.

#### 2.4.2 Chomsky / Twentieth-Century Linguistics

Aristotle’s model of categorisation has exerted an enormous influence in twentieth-century linguistics. Formalist approaches to phonology, syntax, semantics rest on the assumptions of the Aristotelian model. These are apparent in the status of grammatical categories in Chomsky’s transformational paradigm where there is the requirement that membership is a clear-cut matter. This is because transformational rules operate on underlying phrase markers, independently of the semantics of the lexical items that fill the category slots. Because of this, grammatical categories are ‘necessarily’ clear-cut entities (Taylor 1995:

186-7). The ‘obviousness’ of Aristotelianism can also be seen in phonology and is, for example, a basic assumption of Chomsky and Halle (1968: 297):

In view of the fact that phonological features are classificatory devices, they are binary...for the natural way of indicating whether or not an item belongs to a particular category is by means of binary features.’

Phonemes are classified in terms of binary features which can be present or absent e.g. [+HIGH] or [-HIGH]. Another aspect of features is the lack of *decomposition*, these features being said to be *primitives*. An analysis of semantic categories along the lines of phonology has been pursued within the transformational-generative paradigm by Katz (e.g. Katz and Fodor 1963; Katz and Postal 1964) and Bierwisch (1967, 1970); it has also been explored under the name ‘componential analysis’ by Nida (1975) and Leech (1981). As illustration, consider the category of ‘bachelor’. Katz and Postal (1964) represent the meaning of this word in terms of four semantic features, namely [HUMAN], [MALE], [ADULT], and [NEVER MARRIED]. Together these are a sufficient set of features for capturing the essence of bachelor and likewise if any of these necessary features is absent (e.g. [- MALE] or [-ADULT], the predicate of BACHELOR cannot be ascribed. Predicate calculus translations of sense relations between predicates reflect this ‘all or nothing’ attitude to category membership:

<i>Hyponymy:</i>	$\forall x [ C (x) \rightarrow D (x) ]$
<i>Synonymy:</i>	$\forall x [ C (x) \rightarrow D (x) ]$
<i>Converses:</i>	$\forall x \forall y [ C (x, y) \rightarrow D (y, x) ]$
<i>Binary Antonymy:</i>	$\forall x [ C (x) \rightarrow \sim D (x) ]$

Chomsky has made claims for the universality of semantic features. Just as the set of universal phonological features defines the sound-producing capabilities of man, so the set of universal semantic features defines his cognitive capabilities:

It is important to determine the universal, language-independent constraints on semantic features - in traditional terms, the system of possible concepts. The very notion 'lexical entry' presupposes some sort of fixed, universal vocabulary in terms of which these objects are characterized, just as the notion 'phonetic representation' presupposes some sort of universal phonetic theory. It is surely our ignorance of the relevant psychological and physiological facts that makes possible the widely held belief that there is little or no a priori structure to the system of 'attainable concepts'.  
Chomsky (1965: 160)

The postulation of universal semantic primitives is not an innovation of generative linguists. Leibniz, in the seventeenth century, set himself the task of discovering the 'alphabet of human thought' - a set of basic conceptual building blocks, not susceptible to further decomposition, whose combination might underlie all possible concepts, (Eco, 1995: 270).

Chomsky only ever claimed to have produced a model of linguistic *competence* rather than linguistic *performance*. Despite this, the relevance of transformational grammar for sentential production and comprehension was investigated in a series of fairly well-known experiments. In the next section, I outline an attempt to implement transformation grammar as a model of performance known as the Derivational Theory of Complexity. I then go on to outline other performance models which, while not Chomskyan, evoke the symbolic / logical empiricist 'spirit' in placing a large emphasis on syntactic considerations in processing. Again, similar to previous sections, in chapter 3, when we examine some CDA text commentaries, we shall see that much of what is outlined below is reflected in how CDA highlights mystifying text.

## **2.5 Approaches to Language Processing which Place Emphasis on Syntax**

### *2.5.1 Chomsky as a Model of Processing - Derivational Theory of Complexity [DTC]*

The main assertion of the Derivational Theory of Complexity was that the more involved

the transformational history of a sentence, the more processing labour would be required. Initial evidence (Miller 1962; Miller and McKean 1964; Savin and Perchonock 1965) appeared to support the above hypotheses. Miller reasoned that a passive sentence should be more difficult to handle than a simple, active, affirmative and declarative sentence. He tested this hypothesis by giving subjects two columns of shuffled sentences, and asked them to find consonant pairs. In one section of the experiment, actives and passives were shuffled so that a passive like 'the rioters were shot by the police' had to be twined with its intended 'companion' sentence 'the police shot the rioters'. In timing his subjects, Miller found that it took nearly twice as long to match sentences which differed by two transformations. On this experimental evidence, Miller concluded that transformations were psychologically real.

Belief in DTC, though, did not last long. Both theories withered under the attack of Fodor and Garrett (1966; 1967). Their arguments were as follows. Fodor and Garrett highlighted that there could be other reasons for the difference in processing time other than processing histories. For example, with the passive there is actor displacement away from the beginning of the sentence and so extra cognition time is necessary to locate and then link actor to process. Moreover, compared with verb forms in present and past simple tenses which are also *active*, affirmative and declarative, parallel *passives* introduce auxiliaries, elongating both syntax and thus cognition time. That is, the extra processing time needed for passives may be unconnected to the intricacy of a transformational history. They also found that there was no detectable time difference in comprehension between 'John phones up the girl' and 'John phones the girl up'. If DTC was correct, then the second should involve more processing labour since a 'particle separation' transformation is entailed. After the initial hope and excitement, then, the idea that transformational grammar was a model of language production and comprehension was abandoned. For Chomsky, it should be

made clear, such refutation would be immaterial since his model was one of competence and not performance. In other words, transformational grammar is a description of the knowledge of our language rather than an account of parsing procedures.

### *2.5.2 Canonical Sentence Structure (CSS)*

I have indicated how Fodor and Garrett (1966; 1967) denied the possibility of there being psychologically real transformations. Fodor and Garrett nevertheless claimed that the end-product of syntactic processing was a syntactic representation equivalent to the deep structure as posited by Chomsky (1965). If syntactic parsing did not operate via transformations, how was it done? Fodor and Garrett (1967) argued that instead parsing was performed by perceptual heuristics or surface structure cues. They detailed a number of parsing strategies that used only 'surface syntax' as a cue. One strategy became known as *canonical sentence structure* (CSS). CSS in English is essentially subject-verb-object and so, given this, it is reasonable for a comprehender to initially assume that many sentences conform to this structure. On such a heuristics, comprehenders intuitively try a simple strategy like CSS first, and failing this, move on to other more complex ones. Perceptual heuristics as an approach to sentential parsing was further developed in Fodor, Bever and Garrett (1974). It is easy to trace echoes here of the symbolic separation of syntax from semantics and subsequent emphasis on syntax.

### *2.5.3 Modularity Hypothesis*

The general principle that when readers parse they employ 'surface syntactic' cues has remained influential in psycholinguistic research and has gradually become more and more systematised. It tends to go hand in hand with an espousal of syntactic autonomy - that there



exists an autonomous syntactic processor. This assertion is incarnate within the *symbolicist*, Fodor's (1983) *modularity* hypothesis. For Fodor, modules are discrete centres of cognition dedicated to specific cognitive tasks. For example, Fodor posits that there exist different modules in the visual system devoted to stereoscopic vision, colour vision etc. Language is also regarded as being modular, the syntactic processor being one module amongst others in the linguistic system. Input modules are separate from the central processor whose remit is to infer inductively from what the modules despatch. Modules are differentiated from the central processor in certain respects:

- i) Domain specificity of modules - processing in the syntactic module transpires only in the language domain
- ii) Mandatory functioning of modules - processing not under conscious control
- iii) Modules have restricted access to the central processor - only the final yield of processing is open to higher-level (top-down) processors which cannot ingress the module and affect performance.
- iv) Processing in modules is encapsulated - only following an entire operation is processing yield made available to other systems.

For Fodor, it is the modularity of syntactic processing which sanctions its speed and involuntariness; (see Lakoff (1987a: 225) for a discussion of the notion of autonomous syntax).

#### *2.5.4 Late Closure and Minimal Attachment Strategies*

The idea of reader-employment of surface-structure cues was cultivated in different directions by Kimball (1973) and by Frazier and Fodor (1978) and Frazier (1987). In Frazier

(1987), a somewhat strong version of the autonomy of syntactic parsing prior to semantic processing is postulated. Frazier puts forth two strategies for retrieving the underlying syntax:

i) *late closure strategy* - if grammatically feasible, each new item is hooked to the clause or phrase currently being processed.

ii) *minimal attachment strategy* - the reader mentally constructs the phrase structure such that the number of nodes is kept to a minimum.

### *Late Closure*

Frazier appeals to the phenomenon of 'garden-pathing' in substantiating her claims. 'Garden-path' sentences are those where the smoothness of processing is interrupted midstream. The reader is forced to backtrack and use an alternative parsing strategy because they have been 'led up the garden-path'. Consider the following sentence:

Since Jay always jogs a mile and a half seems like a short distance

For Frazier, this sentence institutes a 'garden-path' effect because readers typically employ the strategy of 'late closure'. The reasoning is as follows. *Jogs* is a constituent of a VP and the next input is an NP (*a mile and a half*). On the strategy of late closure, the new input would be fastened to the VP currently being processed. The result is the erroneous interpretation where the NP (*a mile and a half*) is the object of the verb *jogs*.

### *Minimal Attachment Principle*

In an earlier set of experiments and along with two colleagues, Frazier aimed to highlight how readers operate on a syntax-first strategy despite pragmatic implausibility. Rayner, Carlson and Frazier (1983) examined subject's eye movements while reading sentences like the two below:

The spy saw the cop with binoculars but the cop didn't see him

The spy saw the cop with a revolver but the cop didn't see him

The syntax-first strategy of minimal attachment should lead the parser to attach the PPs 'with a revolver' or 'with binoculars' to the VP rather than to the NP 'the cop' so as to keep the number of nodes to a minimum. This would lead to a pragmatically plausible reading in the first sentence *but* an implausible one in the second. Rayner et al. (1983) found that subjects hesitated on 'revolver' in the second sentence but not on 'binoculars' in the first.

It was concluded that when the minimal attachment principle is consistent with pragmatic plausibility, smooth comprehension ensues. Conversely, when the minimal attachment principle is incongruous with pragmatic feasibility, comprehension is frustrated. The larger conclusion drawn by Rayner et al. (1983) was that syntactic parsing transpires independent of and prior to the construction of semantic or pragmatic representations.

In the last section of this chapter, I want to briefly indicate how the emphasis on syntax in accounts of language processing such as outlined above has meant that inferential processes have been downgraded in emphasis and importance. Again in the next chapter, we shall see that this perspective underlies many of CDA's text commentaries.

## 2.6 A By-Product of Syntax First Processing: Inference Last

One of the by-products of a syntax first approach to processing in certain psycholinguistic quarters and an emphasis on sentential form in general is that inference generation is given much less prominence. This can be seen in the stipulation of Fodor's modularity hypothesis concerning higher-level (top-down) processing. In other words, the generation of discourse inferences would be separate from and only *succeed* modular syntactic processes. The lack of emphasis upon inference generation as a mandatory process is encapsulated in the following from Fodor, Fodor and Garrett (1975: 526):

...the distinction between processes that are involved in understanding a sentence and processes that are involved in drawing inferences from it corresponds to a distinction between mandatory, on-line psychological processes and optional, long-term psychological processes. For, by hypothesis, the output of the sentence comprehension system is that representation of the sentence which must be recovered by anyone who understands it. But the application of principles of inference is presumably largely context-determined. What inference we draw from what we hear must be a question of what we take to be relevant to the task at hand.

Thinkers such as Fodor are firmly entrenched in Chomskyan ideas, and so we might regard a perspective which regards inferences as being *separate* from the comprehension process proper as being a part of, or an implicit part, at least, of the 'symbolic vision' which awards primary status to syntactic representation.

In this chapter, I have outlined *symbolicist / logical empiricist* assumptions, the classical theory of categories and how the emphasis on syntax has informed some psycholinguistic theory. In the next chapter, we shall see that many of the above assumptions inform how CDA isolates mystifying text. In chapter 4 and 5, I outline how symbolic / logical empiricist / classical assumptions of processing are problematised by the enterprises of connectionism and cognitive linguistics and thus, in turn, CDA strategies for the location of mystifying text.

In contrast to what I have outlined in this chapter, we shall see, in chapter 4's discussion of connectionism, that far from being downplayed, the issue of inference generation in connectionist networks is a prominent one. This is echoed in work on inference generation in psycholinguistics in the late eighties and early nineties, which I outline in chapter 6.

## Notes

1. Logical empiricism is commonly thought of as the second variant of *logical positivism* (Outhwaite, 1987: 6). The first variant is associated with Comte in the early nineteenth century. Comte submitted a hierarchical conception of science, based on causal laws of phenomena, derived from observation. His main concern was with *positive* knowledge, based directly on experience as opposed to theology and metaphysics. His was also a *naturalist* perspective as are all variants of positivism. For Comte, knowledge progresses when each science attains *positive* knowledge. 'Sociology' (Comte's coinage) is at the bottom of the scientific hierarchy since for Comte it is the last to attain positive status. The logical empiricists of the Vienna Circle preferred to avoid the term positivism since they regarded Comte's philosophy of history to be itself metaphysical. They differed from Comte also in espousing physical reductionism. Logical empiricism became eventually modified into the third variant of positivism. This was sometimes referred to as the 'Standard View' in the philosophy of science and was most influential for twenty or thirty years during the middle of the century. Philosophers associated with this view included Carl Hempel, Ernest Nagel and Karl Popper together with the later work of Rudolph Carnap. One of the central elements in the 'Standard View' was a concern with scientific explanation. The physical and social sciences were regarded as being devoted to the pursuit of explanations which are formulated as general laws or *covering laws*. To explain an event is to relate it to a general law, analysed as a universal generalisation. The freezing of a car radiator, for example, is explained by the general laws governing the behaviour of water together with the low temperature (initial conditions). The provenance of this conception of explanation lie in Hume's theory of causation where what can be observed are the 'constant conjunction' of events such as freezing temperatures and burst radiators rather than the actual causes. In the 'Standard View' a case for a strong physical reductionism is rejected but physics still remained the ideal.

2. Truth values for the logical connector &:

<i>A</i>	<i>B</i>	<i>A &amp; B</i>
<b>T</b>	<b>T</b>	<b>T</b>
<b>T</b>	<b>F</b>	<b>F</b>
<b>F</b>	<b>T</b>	<b>F</b>
<b>F</b>	<b>F</b>	<b>F</b>

3. Gardner (1987: 188) does indicate, all the same, that ‘some of his [Chomsky’s] specific ideas about how language works ran directly counter to information-theory notions.’ By this he is referring to Chomsky’s criticisms of finite-state grammar whose provenance, as I have said, lay within ‘Information Theory’ (Shannon, Weaver et al.).

## **CHAPTER 3: SYMBOLIC, LOGICAL EMPIRICIST, CLASSICAL UNDERPINNINGS OF HOW CDA HIGHLIGHTS MYSTIFYING TEXT**

### **3.1 Orientation**

The seminal CL texts - Hodge and Kress (1993) / Kress and Hodge (1979) and Fowler et al. (1979) - draw heavily upon early Chomsky (1957, 1965) and adapt his competence model to one of performance. Much of this adaptation is still present in the language processing assumptions of recent CDA and how CDA locates mystifying text. Having outlined symbolism, logical empiricism and the classical approach to categories in chapter 2, and their influence on Chomsky, in this chapter I show how these influences make up many of the language processing assumptions of CDA, often via Chomsky. I indicate also how symbolism, logical empiricism and the classical approach to categories influence how CDA isolates mystifying text. To begin this chapter, let me outline the relationship between CDA and Chomsky.

### **3.2 CDA and Chomsky**

#### *3.2.1 Hodge, Kress and Transformations*

##### *Orientation*

Much 'borrowing' of Chomsky's theory of transformations occurs in Hodge and Kress (1993), Fowler et al. (1979); a more formal discussion of this 'borrowing' is located within Hodge and Kress (1974). In this section, I deal mainly with Kress and Hodge (1979) /

Hodge and Kress (1993)<sup>1</sup>. This is one of the major sources in CDA for the highlighting of textual mystification and has gone on to be favourably cited in more recent CDA such as Fairclough (1992a) and (1995a).

Here are Hodge and Kress (1993: 10):

In our account, transformations are a set of operations on basic forms, deleting, substituting, combining, or reordering a syntagm or its elements. So *The car was wrecked* is transformed from (*someone or something*) *wrecked the car*, with the actor (*someone or something*) deleted and the elements of the syntagm reordered in the passive.

Hodge and Kress's signalling of 'actor' rather than 'subject' point to a semantic notion of transformations in contrast to Chomsky's syntactic emphasis. Another difference between Hodge and Kress's conception of transformations and Chomsky's is that Hodge and Kress (1993: 10) do not regard transformations as innocuous operations:

In transformational theory it is assumed that transformations are always innocent (that is, they do not alter the meaning of the basic form) and can always be reversed. In actual discourse this is, sadly, not always the case. Transformations serve two functions, economy and distortion, often so inextricably mixed that even the speaker cannot separate them.

The typical function of transformations is distortion and mystification.

Hodge and Kress (1993: 35)

### *Transactivity*

Hodge and Kress found much of their examination of texts upon a set of basic models. These models are semantico-syntactic and comprise what Hodge and Kress (1993: 9) call *actionals* and *relationals*. The category of actionals, as the name suggests, subsumes



‘actions’. Actionals consist of two sub-models, *transactives* and *non-transactives*. Transactives are explicit about causal relations and non-transactives are not. Take ‘the batsman struck the ball’. This is transactive because, in the event, action flows from actor to affected and so, for Hodge and Kress, this sentence is explicit about causation. ‘The batsman runs’ is non-transactive since the absence of an affected means the sentence is inexplicit about the causal process. While transitivity, as traditionally defined, is a syntactic notion, transactivity is semantico-syntactic. For Hodge and Kress (1993: 8), while ‘the parcel weighs ten pounds’, ‘John plays tennis’ and ‘Bill resembles his father’ are all transitive, they are not transactive since action does not flow from an actor to an affected.

A corollary of all this for Hodge and Kress (1993) is that since transactives are explicit about causality, they are the preferred sub-model for describing causal processes. The second basic model, relationals, is concerned with the classification system of the language.

Hodge and Kress differentiate two sub-models: equatives and attributives. Equatives pertain to relations between nouns that are commutative, e.g. Philip Brown is the headmaster of Belpont Comprehensive. Attributives institute relations between nouns and qualities which are not commutative, e.g. Philip Brown is a kind man.

#### *Mystification and Transformations: Hodge and Kress’s Analysis of a News Text*

Consider now an example of how, for Hodge and Kress, transformations are meant to *mystify* or distort. Recall from 1.4.2 the examination by Hodge and Kress (1993: 21) of part of a newspaper editorial on the miner’s overtime ban during the winter of 1972-73. First, here again is the sentence from the editorial:

The Government knows that in early 1972 it was caught out by picketing of power stations which curtailed coal deliveries.

and here are Hodge and Kress (1993: 22):

*Picketing curtailed coal deliveries*

has, underlying it, a considerable complexity, a varied history of transformations. As readers of this editorial we should have to be alert and willing to engage in mental exercise to get beyond the seductive simplicity of the final form, with just three entities, and seemingly precise relations, where everything seems to be there on the surface...we can see **that few commuters on the 8.05 from Brighton would have the energy to perform the mental gymnastics required.**'

Hodge and Kress (1993: 22) [my bold]

The passage is rich in tacitly logical empiricist and symbolic notions. The notion that a 'surface form' conceals an 'underlying' form echoes a similar assumption in logical empiricism we saw in the last chapter. A symbolic perspective is also echoed since mental representation for Hodge and Kress is tacitly *sentence-like*, where the necessary mental exercise will transform the surface form of 'picketing' into 'strikers picket a power station'. However, this mental representation is not the same as *mentalese*. Rather than a 'language of thought' we have something like 'thought consisting of language', a notion popularly associated with Whorf. And indeed, Whorf (1956) is twined with Chomsky in Hodge and Kress (1993) and cited throughout as theoretical substantiation. The result is that the underlying view of mental representation in Hodge and Kress (1993) is an amalgam of 'Whorfianism' and symbolism. So, rather than advocating mentalese, for Hodge and Kress, thought seems to consist of what might be called *linguese*.

### 3.2.2 *Linguese*

Following Chomsky's 'surface' and 'deep' structures, Hodge and Kress seem to be advocating 'surface' *linguese* and 'deep' *linguese*. But Chomsky's distinction between

‘surface’ and ‘deep’ linguistic structure is itself a logical empiricist notion (redolent of Carnap, Russell and the early Wittgenstein) and so a logical empiricist notion has been absorbed into Hodge and Kress. However, Hodge and Kress’s ‘deep linguese’ consists of language and is not, then, the same as the Chomskyan deep and abstract underlying string. Here are Hodge and Kress again (1993: 28):

All the processes mentioned here work to obscure the originally chosen models; deletion, simplification, collapsing of forms into single units, all act to alter the way in which a reader meets the material and tend to structure his interpretations in specific ways. He is continually coerced into taking the surface form as the real form; and the surface form is a radically transformed version of the originally chosen linguistic form.

Hodge and Kress seem to imply that graphic symbols are cognitively reiterated as ‘surface’ linguese. Like mentalese, this time, in CDA the sentence is both the vehicle of computation **and** the vehicle of mental content. Returning to the quote in the last section, cognitive labour is involved in transforming ‘surface’ linguese into transactive ‘deep’ linguese. Transactives represent reality more ‘fully’ (they do not entail distortion of what really occurred etc), but for Hodge and Kress (1993), there is a paucity of transactives in the text. So, for a reader to have a ‘fuller’ appreciation of the events in question, considerable cognitive effort is necessary to transform the ‘surface’ linguese into ‘deep’ transactive linguese (‘about a dozen times on every full line of newsprint’ (Hodge and Kress, 1993: 22)). Since such effort will not normally be invested (see also 1.4.2), textual mystification occurs.

### *3.2.3 Hodge and Kress (1993) and The Derivational Theory of Complexity*

Hodge and Kress’s (1993) processing assumption above is very much akin to the ‘Derivational theory of Complexity’ (DTC) which we saw in the last chapter. In their

formulation of ‘transformations’, Hodge and Kress (1993: 35) assert the following:

We take a strongly realist position and regard all transformational analyses as hypothetical reconstructions of psychologically real processes.

But as we saw in chapter 2, the early promise for the psychological reality of transformations from the experiments of Miller (1962); Miller and McKean (1964); Savin and Perchonock (1965) was dashed comprehensively by the experimental evidence of Fodor and Garrett (1966; 1967). And all this is fairly well-known. So, in the light of the psycholinguistic evidence against DTC, it is surely misplaced of Hodge and Kress (1993: 35) to claim that their position is explicitly ‘strongly realist’ regarding ‘all transformational analyses as hypothetical reconstructions of psychologically real processes.’ [See Widdowson (1997) for a more detailed questioning of the ‘psychological reality’ of Hodge and Kress’s (1993) transformational model]. Let me now return to Hodge and Kress’s (1993) analysis of the miner’s overtime ban text. They fix upon the ‘nominalisation’ ‘picketing’ and argued that the reader would have to spend considerable labour to recover the deep form ‘strikers picket the power station’; in other words, to transform the ‘surface’ linguage into ‘deep’ linguage. But if transformations have no psychological reality, then Hodge and Kress’s processing assumption is dubious to say the least.

#### *3.2.4 Difficulties with Transactives as a ‘Basic Model’*

Hodge and Kress (1993: 60) highlight Chomsky’s use of kernel sentences (simple, active, affirmative, declarative structures) as some sort of legitimacy for assigning primitiveness to transactives:

Most linguistic theories assume that there is a limited number of basic sentence patterns or sentence types in

a language, and that the vast variety of actual utterances are constructed around this basic set. Chomsky's *Syntactic Structures* (1957) used the concept of 'kernel sentence'. Though it had a technical definition (sentences derived from the application of obligatory transformations only), in fact this was a set of forms, between five and seven, which could not be analysed in more basic terms.<sup>2</sup>

But this justification is *circular* - i.e, most linguistic theories are predicated on basic sentence patterns, so it is legitimate for Hodge and Kress to predicate their model on basic sentence patterns. One reason that there seems to be no need for 'outside' justification for the rudimentary status of transactives is because the concept of a simple, via the weighty heritage of *logical empiricism*, can appear so natural. Because of this circular reasoning, Hodge and Kress's simple - the transactive - is merely a variation on simples put forward as epistemological primitives over the years by logical empiricists.

In common with the logical empiricists, Hodge and Kress (1993: 35) see their version of a simple as having propositional structure:

*Transformations always involves suppression and / or distortion, but they are normally reversible. The standard that acts as the measure of what has been suppressed or distorted is given by the underlying structures uncovered by reversing transformations. The 'relevant truth' which acts as a standard then is given by full propositions in the form of basic models.*

The above is again circular since from the offset it is presumed that the standard is the basic model (transactive or relational) without supplying any evidence to this effect. Reliance on the notion of a transactive as an epistemological primitive, as having completeness of a 'thought' and thus propositional structure, is seen in many other CDA practitioners such as Martin (1989: 30-1; 42) and Kress (1989a: 77-78), Fairclough (1992a) and Wodak (1996).

In chapter 4, we shall see how connectionism represents a challenge to this mode of propositional knowledge representation and ultimately to the idea of a simple, be it a kernel

sentence or a transactive in the context of text processing.<sup>3</sup>

### 3.2.5 Difficulties with Transactives as the Basis of a Scientific Language

*Hodge and Kress (1993: 40-1)*

Hodge and Kress (1993: 40-1) figure that a *scientific* language should contain a high proportion of transactives. The corollary of this is that the description of scientific processes that proceeds through non-transactives is more primitive. As one example of this, they describe how the growth of the turnip, in the children's story, 'The Enormous Turnip', is described in non-transactive terms, thus *mystifying* the causal process for the child reader (1993: 48-50). Here is a fragment which they focus upon:

As time went by, the rain fell on the seeds and the sun shone down on them, and the turnips began to grow.

Hodge and Kress (1993: 50) then compare both transactive and non-transactive descriptions arguing that the transactives are now de-mystifying of causality:

non-transactive: The rain fell on the seeds.

transactive: The rain (water) moistened the seeds.

non-transactive: The sun shone down on them.

transactive: The sun warmed the soil.

Below I argue that Hodge and Kress's prescription that a scientific language should consist of transactives is redolent of logical empiricism.

*Hodge and Kress & Logical Empiricism*

Despite the titles of Carnap's books, *Philosophy and Logical Syntax* (1935) and *The Logical Syntax of Language* (1937), the nature of semantic categories was also important to this logical empiricist in assigning epistemic primitiveness. Primitive statements consisted of 'physicalist' categories rather than experiential ones. In what follows, I highlight how Hodge and Kress (1993) is redolent of Carnap's 'physicalism'. Firstly, here is Carnap on the thesis of physicalism (1935: 89-90):

...this physical language is the basic language of all science..., it is a universal language comprehending the contents of all other scientific languages. In other words, every sentence of any branch of scientific language is equipollent to some sentence of the physical language, and can therefore be translated into the physical language without changing its content...

For purposes of elucidation, let us take the following psychological statement: 'At ten o'clock Mr. A was angry'. The equipollent sentence of the physical language is: 'At ten o'clock Mr. A was in a certain bodily condition which is characterized by the acceleration of breathing and pulsation, by the tension of certain muscles, by the tendency to certain violent behaviour, and so on.

This emphasis on *physical* categories rather than *experiential* ones is very much a logical empiricist one since it tries to remove the contribution of the human understander as much as possible (2.2.4). Now, when we compare the sentences that Hodge and Kress recommended as primitive, 'the sun warmed the soil' and 'the water moistened the seeds', we see that they are of a less *directly experiential* nature than their non-transactive counterparts, but all the same are not outrightly physicalist. So while Carnap's physicalist / experiential distinction is absolute, we might regard Hodge and Kress's prescription as only moving towards physicalist descriptions. Indeed, it could be argued that the clause 'the water moistened the seeds' is *not* demystifying of the causal process and could provide the wrong impression about the nature of the causal processes. That is, it is not so much that the water moistens the seeds but rather it is the osmotic potential of the seeds' sugars which

sets up conditions for drawing water into the seed. So if we are to follow Hodge and Kress's quasi-Carnapian line, a more 'representational' sentence might be 'the seed draws water into itself'. But that sentence still of course does not tell us the precise chemical processes that occur within the seed, e.g., water triggers off hydrolytic enzymes which begin to break down food stores in the seed; soluble, mobile food molecules are then free to take part in growth etc [cf Cook (1994: 75-8) on infinite detail].

### *Quine vs Carnap*

The logical empiricist 'Carnapian' assumptions above, upon which Hodge and Kress's approach rests, have been criticised by the philosopher Quine (1953). Quine denied the feasibility of Carnap's project of translating experiential language into a physicalist language. For Quine, the sentences that report on our direct sensory experiences are also part of a web of other sentences, as part of general theory. So on a Quinian perspective, the sentence 'the water moistened the seeds' is not discrete in its correspondence to a part of reality. Rather it is 'webbed' with the sentences 'the water moistened the outer-covering of the seeds', 'the water hydrated the seeds' sugars' etc... It follows that a sentence like 'the water moistened the seeds' only makes sense holistically, within the reader's 'store of other sentences'. Since the young children for whom the story is intended will not have such a dense 'store' of sentences relating to germination, respiration and photosynthesis, Hodge and Kress's transactives will still be 'mystifying' for young children. Another corollary of this is that the 'physicalist' sentence provided by Carnap above is also not *discrete* in its correspondence to a part of reality. Rather, each seemingly discrete and *final* physicalist sentence is potentially webbed to a set of other sentences whose detail depends on the degree of encyclopaedic knowledge of the person describing the reality. A final corollary of Quine's position is that it leads one to see sentences not as discrete *representations* of reality



but as *cues* of encyclopaedic knowledge (i.e., one's 'store' of sentences).

### *Sentences as Representation vs Sentences as Cues*

Consider the following from Slobin (1982: 131-2):

A sentence is not a verbal snapshot or movie of an event. In framing an utterance, you have to abstract away from everything you know, or can picture, about a situation, and present a schematic version which conveys the essentials. In terms of grammatical marking, there is not enough time in the speech situation for any language to allow for the marking of everything which could possibly be significant to the message. Probably there is not enough interest, either. **Language evokes ideas; it does not represent them.** Linguistic expression is thus *not* a natural map of consciousness or thought. It is a highly selective and conventionally schematic map. At the heart of language use is the tacit assumption that most of the message can be left unsaid, because of mutual understanding (and probably also mutual impatience). The subset of semantic notions which is formally marked in a particular language serves more to guide the listener to the appropriate segments and categories of analysis than to fully represent the underlying notions.' [my bold]

Paraphrasing what I bolded above: sentences *evoke* encyclopaedic knowledge, they do not *represent* the world.<sup>4</sup> Since linguistic meaning transpires within an extant theory, it is patent that Hodge and Kress have neglected the *cueing* aspect of language, its function in evoking encyclopaedic knowledge. Or to put it another way, problematically, they are in tune with the logical empiricist notion that sentential structure can *reflect* or *represent* the situation in the world independent of the contribution of the human understander (2.2.4). So, if the encyclopaedic knowledge is rich, as would presumably be the case with a reading of the story by a botanist, then Carnapian equipollent translation even to the limited extent of Hodge and Kress would be otiose. But if encyclopaedic knowledge is impoverished, as is the case with children, then Carnapian equipollent translation and thus Hodge and Kress' translation into transactives will be otiose too, as we saw in the last section. This is all because mental representation is the *output* of what sentences *cue*. This mental

representation necessarily goes *beyond* the sentential structure and so is not, as implied in Hodge and Kress (1993), a facsimile of sentential structure. The notion of ‘language as cues’ will return when I highlight connectionist approaches to language processing in chapter 4. The importance of taking account of encyclopaedic knowledge (i.e., the contribution of the human understander) in the issue of textual mystification will be dealt with from a more psycholinguistic point of view in chapter 6.

I have dealt specifically with how *Chomskyan* ‘borrowings’ in Hodge and Kress (1993) have absorbed logical empirical / symbolic postulates. This is despite Hodge and Kress’s (1993) more *semantic*-syntactic notion of transformations. We saw how Hodge and Kress’s (1993) postulation of ‘psychological reality’ for their brand of transformations was problematised by the discrediting of the derivational theory of complexity. In 2.5.1, I outlined Fodor and Garrett’s (1966; 1967) rebuttal of the possibility of there being psychologically real transformations. Fodor and Garrett (1967) went on to argue parsing was performed by perceptual heuristics or surface structure cues, claiming nevertheless that the end-product of syntactic processing was a syntactic representation equivalent to the deep structure as posited by Chomsky (1965). I indicate in the following how *symbolic* ‘perceptual heuristics’ also inform processing assumptions in CDA.

### 3.3 Perceptual Heuristics and CDA

In 2.5.2, we saw that Fodor and Garrett (1967) argued instead that parsing was performed by perceptual heuristics or surface structure cues. One strategy became known as *canonical sentence structure* which in English is subject-verb-object, Fodor and Garret positing that comprehenders initially assume that many sentences would conform to this structure. On

such a heuristics, comprehenders supposedly try a simple strategy like CSS first, and failing this, move on to other more complex ones. We also saw that the strategies of late closure and minimal attachment and the phenomenon of garden-pathing are in accordance with an initial preference in comprehension for subject-verb-object structure.

As we have seen, CDA authors (e.g. Hodge and Kress (1993)) exploit something similar to DTC in their analyses of news texts. At other times, something akin to perceptual heuristics is drawn upon by CDA authors. Consider the following from Fowler and Kress (1979a: 41-2), an examination of a fragment from a set of university regulations:

Consider, for example, the sentence:

‘All students matriculating in the University shall, so long as they remain in attendance, be bound by the following Regulations and by such other Regulations as the University may from time to time determine.’

The deep structure is actually something like:

The University (AGENT) binds (PROCESS) all students (OBJECT) by Regulations (INSTRUMENT).

But in the passive surface structure, the nominal designating the object (‘all students’) has been placed in the position of theme, i.e. the left-most noun phrase in the sentence, and this is a position normally associated with the role of the agent. **The syntax strongly encourages one to read the first part of this sentence with the expectation that it is going to describe some action carried out by ‘all students’;** this illusion is heightened by the presence of an active verb of a subordinate clause (‘matriculating’) immediately following ‘all students’; and by the extreme distance between the subject ‘all students’ and the main verb ‘be bound’, a distance which forces the reader to cling on to a hypothesis about the way the sentence is going to turn out. The easiest hypothesis is that we are waiting for a main verb which will tell us what action ‘all students’ perform; but this hypothesis will prove incorrect, since it is actually the University which is doing something. **The reader has, however, made strong use of the hypothesis that ‘all students’ are the agent of the sentence, and is likely to retain some sense that the sentence does mean that...** [my bold]

Fowler and Kress (1979a: 41-2) conclude that, in making use of the above hypothesis, the reader processes ‘all students’, the people affected by the process, as occupying a syntactic

position 'which makes them appear responsible for their own fate'. Also, recall from 1.4.3, Kress's (1993: 181-2) analysis of the sentence 'his parents could not afford a uniform' and his argument that:

...many readers of the *Daily Express* may read across this clause in reading the text, and read it as fully semantically transitive, which I shall call, following Hodge and Kress (1993), a 'transactive'. In support of that reading, these readers might say: 'We scrimped and saved, and we afforded a uniform for our children, so why can't they?!

What characterises these and many other CDA analyses is a syntax-first approach, in line with the symbolic separation between syntax and semantics, which we saw in the last chapter. In other words, the reader initially follows syntactic cues as part of their perceptual heuristics, trying a simple strategy like CSS first. The extra step here for these CDA practitioners is that having begun from subject-verb-object as a perceptual heuristic, the reader will then 'read across' this syntactic form reading it in semantic transitive terms whether the sentence has this semantic structure or not. So for Kress (1993), (and implied by Fowler and Kress (1979a)), this perceptual heuristic is capable of misleading - the syntactic form itself is so powerful that it suggests semantic (transactive) meaning. In short, Kress details a type of *shallow* processing, since the processing of the *actual* semantics of the clauses outlined is shallower than the processing of the syntax. The theme of shallow processing is something I return to and develop in chapter 6.

The last two sections have been concerned with how logical empiricist / symbolic notions of *syntax* have imbued CDA. In the next section, I show how the symbolic / logical empiricist notion of *compositionality* - lexical categories as mental building blocks - informs CDA.

### 3.4 CDA and Compositionality

#### 3.4.1 *Orientation*

In 2.2.5, I highlighted the notion of compositionality: that the meaning of a sentence is regarded as the meaning of its parts and that these parts are discrete, enduring symbols such that ‘a word makes approximately the same semantic contribution to the meaning of every sentence in which it occurs’ (Fodor and Pylyshyn, 1988). In semantic analysis, compositionality seems a reasonable assumption. In order to make a discussion of semantic meaning manageable, it seems heuristically acceptable to ‘atomise’ meanings by removing environmental constraints. However, standard semantics differs from CDA in being largely unconcerned with the actualities of cognition. As Langacker (1987b: 1) submits:

...the whole point of truth-conditional semantics is to avoid any postulation of mental constructs in the characterization of semantic structure.

CDA is partly concerned with processing of a non-analytical reader (although, as we saw in chapter 1, their non-analytical reader is undeveloped, lacking a systematic psycholinguistic account). Because of this concern with the *processing* of a non-analytical reader, compositionality, rather than being a heuristic, has been tacitly treated as being a psychologically real facet of cognition in CDA. In order to see this, I consider an extract from Fairclough (1995a: 110-3) who presents an analysis of how ‘the poor’ in the ‘Third world’ are represented in a television documentary.

#### 3.4.2 *Fairclough (1995a: 110-3)*

Fairclough (1995a: 112) criticises the documentary for representing the poor as passive,

merely the outcome of a set of circumstances beyond their control, ‘as Patients - as people who are affected by the actions of others’. Firstly, here is an excerpt from the narration, taken from the ‘third extract’ that Fairclough offers for examination:

Everywhere in the Third World life in rural areas gets harder - and poor people flock to the city. The urban poor get poorer.

Here now is Fairclough (1995a: 113) who, having commented that the text *backgrounds* ‘the poor’ as *agents* of their circumstances (i.e., the agency and responsibility of the ‘poor’ is not made clear), mentions what seems to be an exception to this analysis at first glance:

...there are only two Actors in the third extract, the New People’s Army and, exceptionally, the poor, in *the poor people flock to the city*. Interestingly, the Action here is one more usually associated with sheep - notoriously passive - than people, so the exception does not really contradict what I have said so far.

Fairclough naturally assumes that the meaning here of ‘flock’ is that of a *collective of sheep*. What assumption about cognition and language might enable him to postulate this? One supposition would be that compositionality is an actual feature of cognition, symbols being *discrete* and *enduring*. Fairclough’s neglect of how the semantico-syntactic environment might close down on the possible meaning of ‘flock’ is consonant with such a supposition. Moreover, compositionality facilitates his choice of the ovine sense of ‘flock’ so as to suit his line. But the standard polyseme of ‘flock’ (going in large numbers, e.g.: ‘people flock to football matches on Saturdays’), where the idea of an agentless herd is absent, is surely a higher contender for the meaning of ‘flock’ here. This is because the lexical environment of ‘flock’ effects closure on the possibility that ‘flock’ refers to a collection of sheep. In chapter 4, I examine the connectionist approach to processing which in contrast takes into account the ‘closing down’ effects of lexical environment and does not treat meaning

compositionally, treating meaning as *subtractive* along a linguistic string rather than *additive* as in the above.

### 3.4.3 Fowler (1986: 20)

As another example of compositionality being drawn into notions of mental representation, consider the following from Fowler (1986: 20):

...consider phrases like 'my wife', 'my son', 'my assistant'. Being any of these people involves activity, relationship. But the syntax which is conventional in English - Possessive + Noun - has the unfortunate effect of encoding a human relationship as an *object*, a possession of another person, so that 'my wife' seems to be as totally owned by me as my hand or my books or my car. Obviously, this syntactic structure, apparently so 'natural', embodies a theory of personal relationship as ownership with dominance, with the dominated partner reduced to the status of an object. Once recognized, such processes are seen as ideological and objectionable. It is claimed that they encourage habits of mind and behaviour which are prejudicial to the dignity and the economic progress of the people presented as 'possessed'.

Notice how the compositional arrangement of the grammar has become conflated with how the 'human relationship' is understood. The notion of possession is equated to a discrete symbol. Indeed, the symbolic assumption of *compositionality* - that the symbol 'my' is discrete and enduring - is so strong that it goes unnoticed that 'my' is actually a deictic determiner whose meaning is reliant on context. Taking Fowler's compositional building block approach to meaning to its conclusion, the meaning of 'my' in 'my boss', for example, would also have to be regarded as possessive.

One of the origins of this assigning of meaning to syntactic categories derives from an amalgamation of compositionality with 'Whorfianism' [see 3.2.1 / 3.2.2 and also Hodge and Kress (1993)]. In 'Whorfianism', language is equated with thought - we think in language (i.e. *linguise*). Even in circumspect discussion of 'Whorfianism', Leech and Short (1981:

146-7) still assert what is bolded below:

The classical statement of the view that languages determine the way in which their speakers interpret and categorize experience is that of B.L. Whorf (1956). Recent thinking has suggested that there is a great deal more in common between languages in this respect than Whorf acknowledges; but whatever stand one takes on the issue, the fact remains that **we conceptualise in terms of the categories our language provides for us.** [my bold]

If we actually think in the categories of our language, that is, if the compositionality of a sentence on a page is replicated in thought, then it is not a huge step to believing that the syntactic nature of a category is intact in our compositional thought. In other words, our thinking will be guided by the syntactic nature of the category. This point has particular relevance for 3.5.1 below.

#### 3.4.4 CDA's Use of Lakoff and Johnson (1980)

Recall that in 1.5, I indicated that CDA draws on Lakoff and Johnson's (1980) perspective on metaphor that human conceptual systems, in terms of the ways in which we think and act, are at base metaphorical. Here are Lakoff and Johnson (1980: 4):

...let us start with the concept ARGUMENT and the conceptual metaphor ARGUMENT IS WAR. This metaphor is reflected in our everyday language by a wide variety of expressions:

#### ARGUMENT IS WAR

Your claims are *indefensible*.

He *attacked every weak point* in my argument.

His criticisms were *right on target*.

I *demolished* his argument.

I've never *won* an argument with him.

You disagree? Okay, *shoot!*

If you use that *strategy*, he'll *wipe you out*.

He *shot down* all of my arguments.



It is important to see that we don't just *talk* about arguments in terms of war. We can actually win or lose arguments. We see the person we are arguing with as an opponent. We attack his positions and we defend our own. We gain and lose ground. We plan and use strategies. If we find a position indefensible, we can abandon it and take a new line of attack. Many of the things we *do* in arguing are partially structured by the concept of war. Though there is no physical battle, there is a verbal battle, and the structure of an argument - attack, defense, counterattack, etc. - reflects this. It is in this sense that the ARGUMENT IS WAR metaphor is one that we live by in this culture; it structures the actions we perform in arguing.

Again what underlies the above is the *classical* assumption that concepts are enduringly atomic, i.e., compositional, that they are not influenced by the *accommodative* effects of lexical company (see section 7.2.2). For example, *win* in 'I'd love to *win* some money in the lottery' surely has little to do with *war*; cf Lakoff and Johnson (1980) above, 'I've never *won* an argument with him'.

In 2.4, I outlined the classical theory of categories, which is interrelated to symbolism and logical empiricism. In the next section, I highlight how some postulates of language processing in CDA are bound up with the classical theory.

### 3.5 CDA and the Classical Theory of Categories

#### 3.5.1 *Nominal Form, Objectifying Effects and Mystification*

As I highlighted in 1.4.2, there is a tendency in CDA to regard all nouns as things, to see 'thingness' as a necessary and sufficient condition for nounhood. As an example, consider the following from Hodge and Kress (1993: 23-4):

In discussing the next two examples, *the miners lift their overtime ban*, and *the ban cuts production*, we begin to deal with words whose status as stable nouns is unquestionable...Both ['ban' and 'production'] are descriptions of actions which involve participants, both in fact are descriptions of transactive actions:

someone bans something → ban  
 someone produces something → production

In the case of *production* the *-ion* ending is an outward sign of its derivation, but *ban* has no such marking. There seems therefore a choice for a hearer as to how he or she might interpret these two words. First, we might assume that the speaker had in fact started from the full sentence form..., or at least was aware of the expanded form at some earlier stage in the production of the utterance. In this case it would be quite proper to regard these as nominalizations, though of a kind which have become so conventional as to be clichés. Second, we might assume that speakers use these words, and hearers understand them, as though they were like *apple* or *bench*, but referring to things which happen to be abstract, not concrete physical things. For this kind of speaker or hearer, the linguistic form creates a world of thing-like abstract objects, which are capable of acting or being acted on. Here language determines perception in two ways, by creating an alternative world which can only be 'seen' in language and by imposing this alternative world, with its apparent solid reality, on the material world, so that we no longer see or believe in the world of physical events.

By nature of the fact that 'ban' and 'production' function as nouns in the editorial, it is assumed that they must share the same properties of nounhood as 'apple' and 'bench', i.e., 'thingness'. As we saw in 1.4.2, for CDA, an excess of nominalisations 'objectifies' events being described. This *mystifies* the nature of the events since, for these authors, events are better 'captured' with material action *processes*. Clearly, the classical theory of categories is in the background here, i.e., the all-or-nothing criterion that all nominals are 'thingy'. In chapter 5, I will outline the enterprise of cognitive linguistics and highlight problems with this all-or-nothing criterion for category membership and in turn for what aspects of language CDA regards as being mystifying.

Ascribing meaning to grammatical form, as Hodge and Kress (1993) do, assumes that syntactic information and semantic information are discretely processed. This is consonant with symbolism since 'rules of composition, as well as other rules that operate on symbols, are syntactic and can be applied without regard to the semantics of the symbols', Bechtel and Abrahamsen (1991: 211). An accompanying feature of isolating syntax from the semantics

of the sentence and then attributing meaning to the syntax is that the notion of *compositionality* becomes reinforced. This is because while we can imagine such a phenomenon as the global semantic meaning of a sentence, it is difficult to imagine the ‘global syntactic meaning’ of a sentence. Thus, in focusing on syntax separate from semantics, *individual* syntactic categories inevitably become associated with a *discrete* meaning.

### 3.5.2 *Nominals and ‘Scientific Language’*

Consider the following from Ogborn et al. (1996: 51) (one of whose authors is the CD analyst, Gunther Kress):

Scientific texts are well known for their high concentration of events and processes presented as if they were things. Simple examples include evaporation, crystallization, ionization, speciation, oscillation. Any scientific textbook or journal will yield a multitude of them, as transparent as ‘magnification’ or as opaque as ‘commensurability oscillations in the resistivity’ (culled from a relatively non-specialized journal). Their presence is not due to the barbarous linguistic habits of scientists. They exist in texts and talk as entities because they exist in the thinking of scientists as entities. They are...things with which to think.

This has parallels with section 3.5.1 where nominal expressions *necessarily* correspond to ‘nominal’ thoughts. It also has parallels with 3.2.5 since in the over-emphasis on the form of scientific expressions and their capacity to influence thinking, there is a neglect of the scientist’s own ‘scientific knowledge’.<sup>5</sup> There is no notion that ‘crystallization’, ‘ionization’ etc are linguistic cues rather than ‘thinking entities’.

So far we have only looked at how symbolicism / logical empiricism assumptions underpin CDA’s perspective on sentence processing. I now want to look more broadly at how these assumptions underpin their approach to text processing.

### 3.6 CDA, Text and Symbolicism / Logical Empiricism

#### 3.6.1 Lexical Relations in Text

The following consists of an excerpt from a booklet issued to expectant parents by hospitals (Fairclough, 1992a; 170-1), and then some of Fairclough's (1992a: 182-2) commentary upon it:

A complete physical examination will then be carried out which will include checking *your breasts, heart, lungs, blood pressure, abdomen and pelvis*. The purpose of this is to identify any *abnormalities* which might be present, but which so far have not caused you any problems. [my italics]

Nominalization turns processes and activities into states and objects, and concretes into abstracts. For example, it is one thing to refer to concrete processes in pregnancy which may not be developing normally; it is another to refer to identifying 'any abnormalities which may be present', which creates a new category of abstract entities. The creation of new entities is a feature of nominalization which is of considerable cultural and ideological importance. For instance, an advertisement for cosmetic surgery has the headline 'Good looks can last you a lifetime!'; 'good looks' is a nominalisation (from concrete relational processes such as 'you look good!') which entifies a local and temporary condition into an inherent state or property, which can then itself become the focus of cultural attention and manipulation (good looks, can, for example, be cultivated, enhanced, looked after; they can be said to bring people good fortune, make them happy, give them trouble). Accordingly, one finds nominalizations themselves taking on the roles of goals and even agents of processes. (For further discussion of the properties of nominalization, see Kress and Hodge 1979: chapter 2).

In the text, 'abnormalities' patently lexically is cohesive with 'checking *your* breasts, heart, lungs, blood pressure, abdomen and pelvis'. One readily makes the *inference* that 'abnormalities' refers to *possible concrete* problems with 'breasts', 'heart' etc. So from a *processing* point of view, while 'abnormalities' may be a distinct lexical item as *input*, this is not the same as saying that it is, in processing *output* terms, 'a new category of *abstract* entities'. Why does Fairclough suppose that 'abnormalities' *is* 'a new category of abstract entities'? One probable explanation is that Fairclough is operating on the symbolic assumption that symbols are discrete and enduring in cognition. Another assumption may

be the logical empiricist postulate that symbols derive their meaning from referring to the world. On this perspective, the use of a category presupposes a *referent*. For Fairclough, it seems, the use of a ‘new’ category presupposes a ‘new’ referent.

Echoes of logical empiricism, similar to that of Fairclough, are found in the following by Chilton (1988: 70-1):

To say that someone or something is *sovereign* or *paramount* (in relation to something else) may make some sort of verifiable sense; the two words would also be virtually synonymous. Transform them into nouns [*sovereignty* and *paramountcy*] and you appear to be staking out separate bits of semantic ground with corresponding bits of reality. Creating the word seems to create the thing; once created they can be manoeuvred to form pseudo-explanations of events and states of affairs.

In the above from Chilton, we find again something of the logical empiricist insistence that linguistic items must have denotations. We have seen something of how category relations are treated in CDA. Let us now see how sentential relations are treated.

### 3.6.2 *Sentential Relations in Text*

What follows is from Clark (1992) and constitutes an extract from the ‘Sun’ newspaper with her critical examination together with Simpson’s (1993) supporting critical commentary. The main thrust of her critical examination below is to highlight how the agency of the event of ‘rape’ can be mystified through sentential configuration. I highlight this particular work here, and refer to it in other chapters of this thesis, since it has had a certain resonance in CDA as well as feminist linguistics. Like Trew (1979) it is one of the most referred to CDA articles. Clark’s analysis is not only reproduced in part and commented upon in Simpson (1993) but also in Montgomery (1995). Simpson goes on to extend the analysis. Its significance as an article in feminist linguistics has more recently been bolstered in being

cited favourably in West, C. et al. (1997). This is a chapter synopsis on feminist discourse analysis in a volume edited by the Critical Discourse Analyst, Teun van Dijk. The volume consists of chapter synopses of differing strands within discourse analysis. Even more significantly, the article is reproduced in Cameron (1998), the second edition of a high-profile anthology of feminist writings on language. Here is the excerpt from the 'Sun' text in Clark (1992: 215):

- 1) Two of Steed's rape victims - aged 20 and 19 - had a screwdriver held at their throats as they were forced to submit.
- 2) His third victim, a 39 year old mother of three, was attacked at gunpoint after Steed forced her car off the M4.
- 3) Two days later, he gunned down call-girl, Jacqueline Murray, 23, after picking her up in London's Park Lane.

Clark (1992: 215) alleges that in 1) and 2) '...the perception of Steed as rapist is reduced by making the sentences passive and deleting him as Agent', that is, the semantic-syntactic encoding diminishes Steed's responsibility. Simpson (1993: 170-1) comments upon Clark's analysis:

GOAL	PROCESS	CIRCUMSTANCES
His third victim...mother of three	was attacked	at gunpoint

In fact, the agency involved in this process has to be inferred by implication from the process expressed by the second clause where Steed does now feature in the role of ACTOR / AGENT:

AGENT	PROCESS	GOAL	CIRCUMSTANCES
Steed	had forced	her car	off the M4

The message is constructed in such a way as to obscure the relationship between Steed and the attack. The only entity upon which Steed acts as AGENT is 'her car', whilst the victim of the attack, although prominent in the information structure of the report, is acted upon only by an implicit and unspecified agency. Indeed, so obscured is the relationship between attacker and victim that it allows a possible reading wherein someone else

attacks the woman at gunpoint while Steed only forces her car off the road...We see a wilful refusal to 'tell it like it is'. What, for instance, is so difficult about presenting the details of the story in the following way, where the relationship between attacker and victim is not obfuscated:

- (1) Steed held a screwdriver at the throats of two of his victims as he forced them to submit.
- (2) Steed attacked at gunpoint his third victim, a 39-year-old mother of three, after he had forced her car off the M4.

Again, the assumptions here are logical empiricist and symbolic ones. One sonorous echo of logical empiricism in Simpson is the assumption that the structure of meaningful sentences must mirror the structure of the event (independent of the contribution to processing of the human 'understander': see 2.2.4). This is because the 'absence' (i.e. Steed's agency), on Simpson's account, has to be *inferred* by implication, i.e. requiring the contribution of a human understander. As a result for Simpson, with regard to agency, the text is mystifying. [This is an echo of Trew (1979: 98-9) (see 1.4.1) where an inference by implication was regarded as being *weak*.]

Truth-functionality (a feature of logical empiricism) is a covert issue in the above, also.

Consider first the following from Ellis (1992: 74):

When considering entire texts and sequences of interaction, according to standard structural semantics, it is necessary for every sentence in a text to be true if the entire text is going to be true. Such a requirement leads to some logical inconsistencies that cannot be resolved because the structure of a text, and not only truth values, can determine whether or not a sentence is true.

For 'standard structural semantics', (which has similar foundations to logical empiricism), for a text to be true, each of the sentences should be true. The rather odd corollary of this is that sentences in a text should be logically independent of one another. Ellis is, of course, right to point out that the structure of a text has a bearing on 'whether or not a sentence is

true'. But as in logical empiricism and 'standard structural semantics' where simples are logically independent of one another, so Simpson sees sentences in terms of logically *independent* simples - his simple being the *agent-process-patient* structure. He makes no allowance for how the local, co-textual meaning, and as Ellis implies, global co-text might close down on the meaning of a particular sentence. This is also borne out by the sentential alternatives he offers. Indeed, the grip of logical empiricism and the need to find the underlying semantic-syntactic form is so strong that Simpson's arithmetic goes astray. The number 'two' precedes rape victims and now we have the ordinal 'third'. We are, then, expecting the third of **his** (Steed's) rapes; it cannot logically be anybody else's. So Simpson's alternative reading where '*someone else* attacks the woman at gunpoint while Steed only forces her car off the road' is difficult to accept.

Recall some of the premises of logical empiricism. The world is constituted by well-defined objects with well-defined inherent properties, the result being that objects can be ascribed discrete names. Since these objects have well-defined properties, then, each of these properties can be ascribed a one-place predicate corresponding to each of those properties.

And since the objects stand in fixed relations to one another, then, a series of many-place predicates can be ascribed so as to correspond to each relation. As we saw in 2.2.4, the natural consequence of such a perspective was to regard semantics in a 'building block' or *compositional* fashion. Also, since for meaningful discourse to transpire, sentential structure must necessarily mirror the structure of the world; the syntax needs to be composed of 'simples'. [As we have seen, for Clark / Simpson, these simples are *agent-process-patient* structures.] On this reasoning then, if one of the 'necessary' building blocks required to construct a simple is absent [i.e. the 'agent'], as Simpson and Clark highlight, then this would render the sentence less meaningful. But not only that. Because it is assumed in



CDA that the sentence is both the vehicle of computation and the vehicle of mental content, then, an absence of a 'necessary' semantic component is consonant with an absence of a necessary thought - hence mystification.

I shall refer again in this thesis to the symbolic assumptions of Clark (1992) and Simpson's (1993) analyses when I problematise them from the perspective of connectionism (chapter 4), cognitive linguistics (chapter 5) and recent psycholinguistic work on inference generation (chapter 6). In chapter 6 especially, we shall see that when the contribution of a human processor *is* taken into account, *pace* Clark / Simpson's logical empiricism, the text is not mystifying of Steed's agency for a non-analytical reader (6.6.2). In the final section of this chapter, I want to say something more generally about the issue of inference generation in CDA and its relationship to symbolism / logical empiricism.

### 3.7 CDA / CL, Inference Generation and Logical Empiricism / Symbolicism

#### 3.7.1 Orientation

In 1.4, I outlined how different CL / CDA authors regarded the issue of inference generation. Trew (1979) [endorsed by Toolan (1988), Lee (1992) and Montgomery (1995) as well as Simpson (1993) in the above] contends implicitly that sentential structure should reflect reality. Thus any inferences that have to be generated across sentences or clauses can be treated as weak representations. We also saw in 3.2.5 that Hodge and Kress (1993), in emphasising the importance of semantic-syntactic structure, implicitly downplay the importance of encyclopaedic knowledge in *processing* and thus inferences generated on the basis of encyclopaedic knowledge. CDA's alliance with something akin to the derivational

theory of complexity also accounts for the emphasis on propositional form and deeper structure. In turn this also means that non-propositional factors such as encyclopaedic knowledge and inferences generated on the basis of this knowledge are downplayed. We also saw in the last section how inferences were downgraded in Simpson's (1993) analysis. It should be apparent by now that this downgrading of inferences *more generally* in CDA is also in line with the *syntax-first* position of logical empiricism / symbolism, and the notion that sentential structure should reflect events independently of the (inferential) contribution of the human understander. Let us see in a little more detail how these assumptions have affected the CD analysts in 3.6 with regard to inference generation.

### 3.7.2 *Returning to Fairclough (1992a); Clark (1992)*

We noticed that Fairclough does not see 'abnormalities' as being 'instantiated'<sup>6</sup>, i.e. that a reader would construct an inference *across adjacent sentences* to the effect that 'abnormalities' is made more *concrete* via the information 'your breasts etc...' Consider also the following from Clark (1992: 215):

In both descriptions of the rapes (1) and (2), the perception of Steed as rapist is reduced by making the sentences passive and deleting him as Agent. This perception is further reduced by using the euphemism 'attacked' to mask the terrible details of abduction, repeated rape, and death threats (not mentioned at all in this newspaper).

Again, like Fairclough, the more general category 'attack' is not seen as being 'instantiated' by the more specific category 'rape' even though it is in an adjacent previous sentence (see 3.6.2). So, Clark, like Fairclough, neglects how inferences might be generated across sentences in *reading*. In other words, both Fairclough and Clark are in line with the *syntax first - inference last* perspective which we saw in 2.6 is in the spirit of logical empiricism

and symbolism. Or put another way, what underlies both Fairclough and Clark is the logical empiricist assumption that all the information should be ‘in’ the sentence, i.e., the structure of the sentences should mirror reality regardless of the processor’s (inferential) input.

### 3.7.3 Fairclough, Schank and Abelson (1977) and Inference Generation

Similar to the authors cited above, in Fairclough’s (1989) analysis of the ‘Quarry load-shedding problem’ text from 1.4.1, we saw that inference generation was implicitly downplayed with an emphasis on ‘surface’ sentential structure. That is, though the sentential structure did not actually *represent* the causal relationship *directly*, it was not acknowledged that the causal relationship could actually be inferred. By implication, an inference as to causal antecedence was seen as a weaker representation than if causal antecedence had been *directly* specified in the sentential structure. However, I also pointed out in 1.4.1 that this downplaying of inference generation was in *conflict* with Fairclough’s analysis of the ‘Jenny Keeble’ text where inference generation was emphasised. Indeed, in this analysis, the implicit assumption was that inferences were *strong* representations since they could lead to the reproduction of sexist!

If it is clear that logical empiricist / symbolic assumptions govern why inference generation is so downplayed in the examples I cited above, why should then inference generation be given prominence in Fairclough’s analysis of the Jenny Keeble text? This is because what informed Fairclough’s analysis of the ‘Jenny Keeble’ text was the *script* theory of Schank and Abelson (1977). Schank and Abelson (1977) regard the encyclopaedic knowledge component in a processor as supplying the bulk of processing requirements. This encyclopaedic knowledge component embodies a series of well-delineated knowledge

scripts. Scripts delineate information associated with particular events or situations such as eating in a restaurant. Scripts are something akin to a complex semantic network that forms the parameters of a stereotypical set of events. A restaurant script might house information about menu selection, the order of courses, paying the bill, leaving a tip etc. Scripts enable a reader to *infer* stereotypical information when confronted with a text that alludes to a restaurant. Schank and Abelson's (1977) assumption is that a cue such as 'restaurant' evokes the entire script, the reader making default assumptions that stereotypical events transpire. With such an emphasis, Schank and Abelson regard *higher-level processes* as influencing parsing, their model exhibiting *strong interaction* between semantic and syntactic processing.

To see Schank and Abelson's influence on Fairclough, here is Fairclough's (1989: 160) follow-up to the Jenny Keeble text analysis:

...the text implicitly conveys the meanings that Jenny Keeble is a 'good wife' and admirable person, through the expressive values of attributes attached to her.

...the meaning that Jenny Keeble is a 'good wife' is not explicitly expressed in the text, and it is only because interpreters have in their heads a mental representation of what a 'good wife' is stereotypically supposed to be that they are able to recognize attributes thereof which occur in the text and **so infer the meaning**. In terms of the preceding section, interpreters make use of a *script* for 'the good wife'. In fact, schemata and frames as well as **scripts** can be regarded as playing a role in the interpretation of point: they act as stereotypical patterns against which we can match endlessly diverse texts, and once we identify a text as an instance of a pattern, we happily dispense with the mass of its detail and reduce it to the skeletal shape of the familiar pattern for purposes of longer-term memory and recall. Fairclough (1989: 160) [my bold]

It should now be clear why there exists a conceptual tension towards inference generation in Fairclough (1989). That is, when Fairclough draws upon script theory he is invoking a top-down approach to processing where syntax and semantics are treated interactively. However, other assumptions in Fairclough's work are not top-down based, but in line with

logical empiricism and Chomsky where syntax is treated as a separate phenomenon from semantics. Indeed, this conceptual tension in Fairclough (1989) is all the more marked since Schank (1980: 36) has been publicly hostile to Chomskyan modularity, promoting the view that semantics and pragmatics are central in language and downplaying the role of syntax:

It is impossible to produce a model of language alone...apart from beliefs, goals, points of view and world knowledge (1980:36)

*Section C* of this thesis builds cumulatively towards an alternative framework for the highlighting of mystifying discourse. However, my framework derives from a different set of cognitive, philosophical and psycholinguistic assumptions from those of CDA. My framework instead draws from compatibilities between connectionism, cognitive linguistics and recent psycholinguistic evidence on inference generation. Consequently, my framework will not necessarily highlight text which leads to mystification in reading that would be highlighted as such by *symbolic* CDA.

I have mentioned that Schank and Abelson's (1977) model treats semantics and syntax interactively. In the next chapter, I outline the enterprise of connectionism which also favours interaction between syntactic and semantic information. I show how connectionism problematises many of the precepts of symbolicism / logical empiricism with regard to mental representation, in turn problematising the language processing assumptions of CDA and thus what CDA highlights as being mystifying text.

## Notes

1. Hodge and Kress (1993) is the 'second edition' of Kress and Hodge (1979). However, unusually it is not revised and only differs significantly from the first edition through the inclusion of a new final chapter which is more oriented to work on CDA in the eighties and early nineties.

2. In fact, Chomsky had jettisoned the idea of kernel sentences by 'Aspects' (1965).

3. Widdowson (1998: 140) highlights a contradiction in CDA in relation to the notion of the transactive:

'In the Hodge and Kress conception, they convert one kind of sentence into another. This would seem to imply the existence of neutral non-transformed sentences which are, by definition, innocent of any representational significance...even if we were able to identify the neutral sentences, their very existence means that it is in principle possible, by a judicious avoidance of transformation, to produce language which is entirely free of representational subjectivity. But this contradicts the critical linguistic tenet that there is no neutral language: *all* of it is loaded, 'ideologically saturated' as Kress puts it (Kress 1993: 174).'

4. Widening the perspective with Peircian categories, while *iconic* signs can be taken to 'represent' the world to varying degrees, *symbolic* (linguistic) signs do not usually unless they are configured iconically.

5. On the issue of 'scientific English', consider the following from Halliday (1993: 68) (referred to in Graddol (1996: 178) also):

'What is *lung cancer death rates*: how quickly lungs die from cancer, how many people die from cancer of the lung or how people die if they have it? What is *increased smoking*: more people smoke, or people smoke more? What is *are associated with*: caused by (you die because you smoke), or cause (you smoke because you are - perhaps afraid of - dying?). We may have rejected all but the right interpretation without thinking - but only because we know what it is on about already.'

Halliday (1993: 68)

Graddol (1996: 178) comments upon this excerpt as follows:

'...scientific English **often** requires a certain knowledge and understanding of the subject matter: it may be better at high-level, abstract argument than at low-level, explicit description.'

[my bold]

But Graddol betrays some attachment to *textual empiricism* and thus to logical empiricist / symbolic assumptions. Given what we have seen in 3.2.5, especially Quine's critique of Carnap, it should be axiomatic that scientific meaning would *always* require knowledge and understanding of the subject matter - linguistic meaning is *always* an interaction between the language and the cognitive resources. Similar to Graddol, Halliday seems to bemoan the lack of explicitness of nominal expressions. But, as we saw with Hodge and

Kress, even translation into transactives, i.e. *verbal expressions*, does not adequately increase the explicitness.

6. Instantiation is the type of inference generated when a general category is made more concrete in context e.g. 'the fish attacked the swimmer' where fish is instantiated readily by the reader as 'piranha' or 'shark' etc. Instantiation, along with other types of inference, is dealt with in chapter 6.

## **Section B: Connectionism and Cognitive Linguistics - Problematizing Symbolicism etc and Implications for CDA**

This section comprises two chapters. In chapter 4 and chapter 5, I outline connectionism and cognitive linguistics respectively, how they problematise symbolism, logical empiricism and the classical approach to categories, and thus what CDA highlights as being mystifying text.



## **CHAPTER 4: CONNECTIONISM - PROBLEMATISING HOW CDA HIGHLIGHTS MYSTIFYING TEXT**

### **4.1 Introduction**

In this chapter, I provide a general portrait of an approach to cognitive modelling known as connectionism. I also detail two well-known connectionist models of sentence processing, namely, McClelland and Kawamoto (1986) and Elman (1990). In the last chapter, I highlighted the symbolic assumptions at work in many CDA analyses of mystifying text. Symbolicism and connectionism in cognitive science are the two dominant approaches to modelling. But on many, if not most, of the central issues in cognitive science they are diametrically opposed to one another. The function of this chapter is to problematise the symbolic cognitive assumptions of CDA from the most established oppositional paradigm and thus problematise what CDA regards as mystifying text.

I should stress from the outset that in this chapter I do not offer connectionism as some kind of universal explanation of all issues of language comprehension. As MacDonald and MacDonald (1995: xvi) acknowledge, connectionist architectures show ‘promise’; the theory of such architectures is ‘still in its infancy’. Similarly, as Schopman and Shawky (1996: 70) remark: ‘the connectionist point of view does not mean that no problems are left, perhaps even new ones have been introduced.’ Rather my aim in section B, at least, is merely to produce a timely situating of the paucity of appreciation of the issue of mental representation in CDA within as wide a contemporary cognitive framework as possible. A brief history will make this clearer. Many of the assumptions of cognition in current CDA derive from the work of Fowler, Hodge, Kress and Trew (1979) and Hodge and Kress (1979), work

which was developed in the seventies. Connectionism only became a serious challenger of the symbolic paradigm in the late eighties, continuing to flourish in the nineties. Because of this, and because the linguistic models of Fowler et al. (1979) and Hodge and Kress (1979) have a sociological orientation, looking at CDA in term of ‘neglect’ of the issue of mental representation would be an unfair *ex post facto*. The symbolic assumptions operative in the above authors would have seemed obvious in the absence of another paradigm. It is timely then to consider issues of mental representation in CDA from the perspective of connectionism, given its contemporary prominence and the challenge it poses to symbolicists. Because of the constraints of this thesis my use of connectionist modelling is selective. Consequently my use of the term ‘connectionist’ is short-hand for a selected group of connectionist authors. Significantly, though, this group includes one of the ‘fathers’ of connectionism, James McClelland.

## 4.2. Connectionism: The General Picture

### 4.2.1 Orientation: *The Interaction of Syntactic and Semantic Information*

In the last chapter, we saw that CDA’s approach to locating mystifying text places a great deal of emphasis upon syntax and that this was in line with the syntax-first approach to processing based on modularity. One of the significant features of connectionism is that it regards syntactic and semantic processing as being simultaneous and *interactive* and so does not operate on a syntax-first basis. Non-connectionist research into this kind of interactive processing, and thus the non-autonomy of syntax, can be found in Just and Carpenter (1980) and Marslen-Wilson and Tyler (1980) Crain and Steedman (1985), Altmann and Steedman (1988) and Taraban and McClelland (1988). Moreover, I highlighted in 3.6.2 how Schank

and Abelson (1977) regard semantics and syntax as interactive in processing. Taraban and McClelland are two authors associated with connectionism but the following outline of experimental evidence from their paper of 1988, for the interaction of semantics and syntax in comprehension, does not have an explicit connectionist emphasis.

#### 4.2.2 Taraban and McClelland (1988)

The evidence from Taraban and McClelland's (1988) experiments is offered as refutation of the syntax first / minimal attachment perspective (see 2.5.4). The following consists of what subjects had to read in one of the experiments in Taraban and McClelland (1988):

The janitor cleaned the storage area with the

- a) broom
- b) solvent
- c) manager
- d) odour

because of many complaints.

For Taraban and McClelland (1988), the above sentence contains a verb which leads to the expectation that an instrument will be mentioned in the PP; that is, the PP will attach to the VP [and inadvertently in accordance with the minimal attachment principle]. This is shown in a). A less expected instrument is shown in b), a less expected role is shown in c), and the less expected syntactic attachment (attaching to the NP 'storage area') is shown in d). Now, the results indicated a significant increment in reading times for the word ('because') following the focal noun when the focal noun was an unexpected role c) or an unexpected attachment d). McClelland and Taraban induced that it is *semantic content* and not syntax which determines interpretation. That is, the slow reading times in c) and d) were due to the presence of unexpected roles irrespective of whether interpretation is consonant with the

minimal attachment principle as in c) or not as in d). Because of this involvement of semantic knowledge, in place of the syntax-first model, Taraban and McClelland argue that a system of parallel constraint satisfaction (i.e, a basis of connectionism), where syntactic and semantic information are processed coterminously, is a more valid explanation of the process of sentence processing.

The papers of McClelland and Kawamoto (1986), McClelland, St. John and Taraban (1989), St. John and McClelland (1992), St. John (1992), which are drawn upon in this thesis, also deal with interactive processing. Where they differ from Taraban and McClelland (1988) is that these authors ultimately attempt some simulation of neural activity. This modelling of neural processing is known as *connectionism* or *parallel distributed processing*. In the next few sections, I outline the basis of connectionist models and then go on to discuss connectionist approaches to multiple constraint satisfaction in sentence processing. Throughout, I explore the repercussions for the symbolic cognitive postulates of CDA and its approach to highlighting mystifying text.

Firstly, let me outline some common arguments that processing models should attempt to simulate neural activity.

#### *4.2.3 Arguments for Neural Modelling*

We saw in 2.2.3, briefly, reasons why the mind has been treated separately from the brain in symbolicism. In symbolicism, change in mental states is achievable through the rule-governed manipulation of sequences of mentalese-sentence strings. Language cognition on the symbolicist perspective is, then, essentially *mentalese-sentence-crunching*. Since mentalese-sentence crunching is algorithmic, it need not be implemented in brains, thereby

reducing the importance of locating mental states with brain states. A large number of thinkers regard this perspective, however, as suffering from a series of difficulties; for example, Anderson and Hinton (1981), Churchland, P.S. (1986), McClelland, Rumelhart and the PDP Research Group(1986) and Edelman (1992). Below is a synopsis of these difficulties:

i) Tasks that pose no problems for von Neumann machines (mathematical calculations, logical reasoning [cf my outline of Wason (1966) in 6.4.2] are usually either poorly performed by humans or require substantial effort. By the same token, tasks which humans take for granted (e.g. facial recognition) are poorly simulated on von Neumann architectures.

ii) The brain is an extremely dense interconnected neural network. Purkinje cell neurons can have more than 80, 000 input connections and neurons in the cerebral cortex can have in excess of 10, 000 output connections. Its *parallel* nature conflicts, then, with the serial architecture of von Neumann machines. Further evidence for the parallel status of the brain comes in the following. Many tasks such as facial recognition and simple responses to questions take about 0.5 seconds. On the basis of what is known about neural firing rates, conduction velocities and synaptic delays in neurons, this allows approximately 5 milliseconds per stage in the cognitive process. This would allow only about 100 stages, which for a serial program supported on a von Neumann architecture would be far too slow. Feldman and Ballard (1982) have coined this the *hundred-step rule*.

iii) Neural networks degrade *gracefully* and are relatively tolerant of localised deterioration similar to brains. Von Neumann architectures, however, cannot tolerate localised deterioration since this will lead to either total breakdown or significant handicap.

iv) The existence of extensive individual variation in cognitive systems negates the fundamental postulate of symbolicism that representations have meaning independent of their physical instantiation. Human experience is not based on so simple an abstraction as a Turing machine; to get our meanings we have to grow and communicate in a society. In contrast to computers, the patterns of nervous system response depend on the individual history of each system, because it is only through interactions with the world that appropriate response patterns are selected.

v) Basing cognition on mentalese-sentences proves problematic for non-verbal animals. Either non-verbal animals do employ mentalese-sentence / logic as a basis for cognition or their cognitions are distinct from human cognitions. Neither alternative seems to be plausible. The first one is lacking in evidence and the second entails a disruption of evolutionary processes. The additional entailment is that evolutionary biology and developmental neurobiology are somehow erroneous. Indeed as Churchland, P.S. (1986: 388) argues:

Sentence-crunching is certain to have been a latecomer in the evolutionary scheme of things, and it must have knit itself into the pre-existing nonsentential cognitive organization or...evolved out of preadaptive nonsentential structures. To be sentence-crunching 'all the way down' implies that cognition must have been sentence-crunching 'all the way back', which is implausible, or that sentence-crunchers have no cognitive heritage from earlier species, which is also implausible given the evolution of the brain.

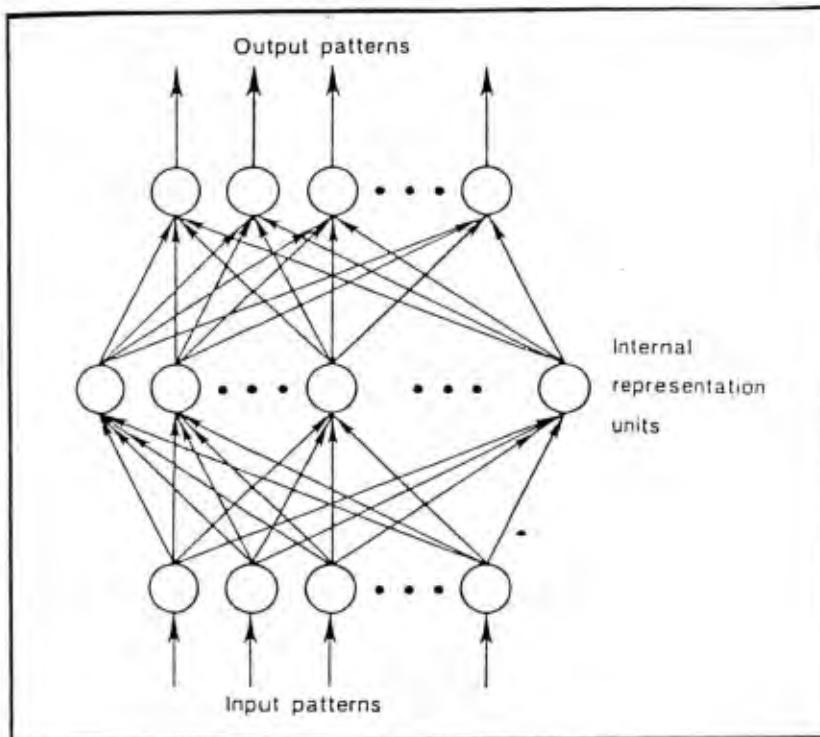
The symbolic emphasis on 'sentence-crunching' or the syntax-first approach, as inadvertently employed by CDA, does not then sit easily with an evolutionary perspective.

#### 4.2.4 Modelling Neural Activity

Firstly, let me outline briefly the characteristics of neurochemical processes. These processes are predicated upon a handful of elementary principles. Neurons are simple processing elements which gather electrochemical pulses on their input side. If the combined total of incoming pulses reaches a particular threshold of activation, a neuron will generate an *action potential*. An action potential is a pulse that is conducted along the axon, a long, thin fibre on the neuron's output side. The rate of the action potential is consonant with the 'strength' of the signal. Since neurons are massively interconnected in a parallel network, the firing or non-firing of particular maps of neurons can either *excite* or *inhibit* the activity of other neuronal maps.

Artificial neurons in connectionist networks are much simpler compared to the heterogeneity of neuronal structure. A connectionist model consists of a multitude of simple processing elements. Each element is known as a *node* or a *unit* and each node has many connections with other nodes. Nodes affect other nodes through either exciting or inhibiting them. Nodes have *activation levels*, a number usually between the (arbitrary) limits 0 and 1, or -1 and + 1, which fluctuate along with the activity around a node. The strengths of connection between nodes are known as *weights* and these carry a numerical description according to the strength of firing. The weight may be either positive or negative. With two active nodes A and B, A will tend to excite B if the weight of the A / B connection is positive. When the weight of the A / B connection is negative, A will tend to inhibit B. Because connectionist networks are usually densely interconnected, usually any node will be simultaneously excited and inhibited by a host of other nodes, the activation of each node calculated as weighted sum of its inputs [see figures a) and b)]. Connectionist processing ends when the system reaches a *stable state*; in other words, where activation through the network does not

Figure a)



A multi-layered connectionist network with a layer of input units, a layer of internal representation units or hidden units, and a layer of output units. Input patterns can be encoded, if there are enough hidden units, in a form that allows the appropriate output pattern to be generated from a given input pattern

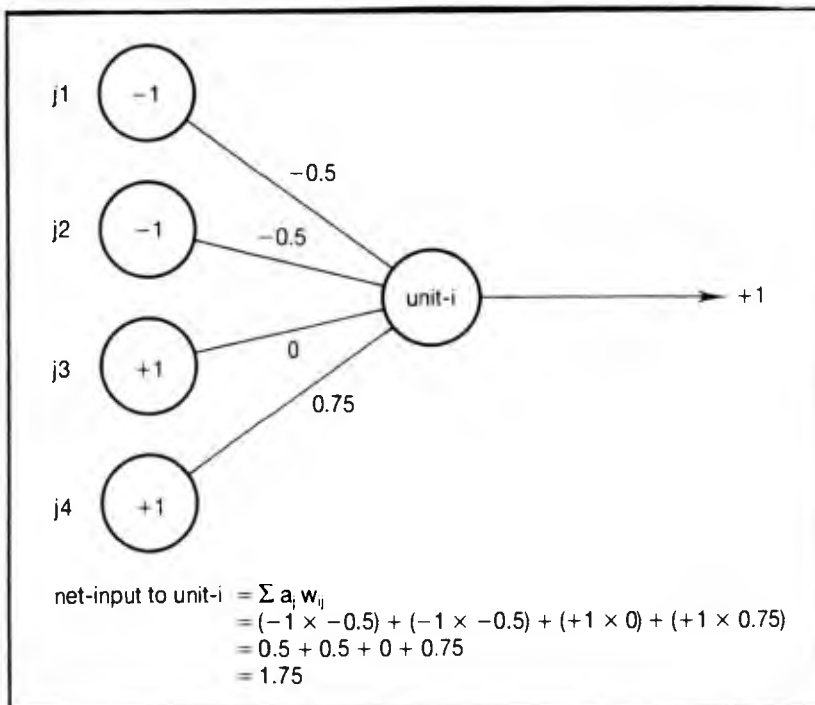


Diagram showing how the inputs from a number of units are combined to determine the overall input to unit-i. Unit-i has a threshold of 1; so if its net input exceeds 1 then it will respond with +1, but if the net input is less than 1 then it will respond with -1.

Figure b)



lead to activation strengths in the units being changed.

### *Training the Network*

Networks are trained to ensure a particular response in the output layer to activation in the input units. The network 'learns' the association between different inputs and outputs by adjusting weights between units. In order to achieve such 'learning associations', *learning rules* are used to systematically calibrate the connection strengths between particular nodes.

When learning rules are implemented, the weights are adjusted until the network effects the desired output patterns given particular input patterns. A widely known learning rule is the *backward propagation of errors rule (BACKPROP)*. BACKPROP permits a network to associate inputs with outputs. Initially, the weight values within the network are randomised. In the early learning phase, following introduction of the input pattern, the output units often spawn responses that are not the desired outputs. To remedy this, BACKPROP compares the undesired output with the desired output, registering the errors.

What it then does is to backpropagate activation though the network modifying the connection strengths until the desired output pattern is produced. What is interesting about 'learning associations' between input and output in connectionist networks is that cognitive processes are portrayed *without* appealing to the kind of *explicit rules* (despite the unfortunate phrase 'learning *rules*') characteristic of symbolic models. All the same, when the network has 'learned' to create a particular response at the output layer, it may appear that a rule of the form 'IF X THEN Y' is being followed.

In general, two kinds of network have been employed - *localised* and *distributed* networks.

In *localised* networks each unit stands for an object or property. So in a localised network model for word recognition, each unit would represent an aspect about a feature, letter or

word (see for example, Feldman and Ballard, 1982; Rumelhart and Norman, 1982; Cottrell and Small, 1983). In *distributed* networks, on the other hand, representation of a particular phenomenon is over a coalition of units. In other words, representation is consonant with distributed patterns of activation (Rumelhart, D.E, McClelland, J.L. and the PDP Research Group, 1986). Connectionist networks have been employed in a variety of ways. One example is Sejnowski and Rosenberg's (1987) network, NETtalk. NETtalk is a connectionist network which has learnt how to transform graphemes into phonemes. From an input of English graphemes, NETtalk is able to produce English phonemes with about 90% accuracy. NETtalk appears then to have 'learned' 'English pronunciation rules' although of course it has *not* been programmed with explicit rules. Following a BACKPROP learning rule, the network is given graphemic inputs and informed of the correctness of the phonemic outputs. On each training, the automatic learning procedure modifies the connection strengths, bringing the system closer to the desired output. It is *following* this training, then, that further English graphemic input is able to produce a 90% accurate phonemic output.

#### 4.2.5 Sub-symbols / Microfeatures

One important characteristic of *distributed* representations is their *sub-symbolic* nature. In contrast to those operating within symbolism, connectionists do not regard cognition as the computation of structured symbolic strings such as the proposition THROW (JULIE, BALL). In contrast, connectionists hold that a predicate such as 'ball', depending on the particular situation, semantico-syntactic environment etc, might be construed differently in a distributed representation. This is because in a distributed representation concepts are profiled in terms of *sub-features*. Ball for example could be profiled with the following sub-features for 'sphericalness' / 'non-sphericalness', 'hollowness' / 'non-hollowness',

‘hardness’ / ‘softness’ etc. These sub-features are termed *sub-symbols* or *microfeatures*.

Consider the following:

- i) The toddler kicked the ball
- iii) The batsman struck the ball and it smashed the window

Looking at ‘ball’ in terms of microfeatures, the *likely* meaning of ‘ball’ in the above depends on the environmental constraints. That is, from a connectionist perspective, the environment of i) might be said to *inhibit* the microfeature ‘non-hollow’ and ‘hardness’ and *excite* instead ‘hollowness’ and ‘softness’ where the opposite would be the case for ii). We shall see in the next section that the probabilities of sub-symbolic inhibition and excitation are one of the fundamentals of a connectionist network. The ‘concepts’ of ball in i) and ii) in connectionist network processing can be activated over a distributed representation. I flag ‘concepts’ with inverted commas since for Smolensky (1988) our conscious, verbal notion of a stable, well-defined ‘ball’ is merely an approximation. Instead ‘concept’ is a convenient fiction that stands for the convergence of situationally specific aspects of a ball which are *emergent* in cognition. For Smolensky (1988: 68) the differences between the symbolic and connectionist paradigms can be seen in their approaches to context:

in the symbolic paradigm, the context of a symbol is manifest *around* it and consists of *other* symbols; in the sub-symbolic paradigm the context of a symbol is manifest *inside* it, and consists of sub-symbols. [my italics]

Smolensky’s emphasis on the microfeatural level is also endorsed by other cognitive scientists:

In our view, the most interesting relation between subsymbolic emergence and symbolic computation is one of *inclusion*, in which we see symbols as a higher-level description of properties that are ultimately embedded

in an underlying distributed system...symbols are not taken at face value; they are seen as approximate macrolevel descriptions of operations whose governing principles reside at a subsymbolic level.'

Varela, F.J., Thompson, E. and Rosch, E. (1991: 101-2)

However, while Smolensky argues that symbols are approximators of sub-symbols, the actual relationship between symbols and sub-symbols in cognition still needs to be accounted for in much greater detail.

Let me now consider a connectionist approach to the modelling of *sentence* processing. The connectionist approach to modelling that I want to consider is that of McClelland and Kawamoto (1986), one of the most cited articles in connectionism. Indeed, McClelland and Kawamoto (1986) derives from the connectionist 'bible' - McClelland, Rumelhart and the PDP Research Group (1986). This model has exerted considerable influence, continuing to be cited in connectionist literature. In other parts of this chapter, I refer to models of sentence processing and text processing which build on McClelland and Kawamoto (e.g. McClelland, St. John and Taraban, 1989; St. John and McClelland, 1992). In chapter 7, I provide in some detail St. John's (1992) connectionist simulation of a short text to highlight how a connectionist model handles inference generation. Here, the reason I outline McClelland and Kawamoto (1986) is principally to problematise symbolic assumptions of mental representation in CDA and thus what is highlighted in CDA as mystifying text.

### **4.3 An Outline of McClelland and Kawamoto's (1986) Connectionist Model of Sentence Processing and its Implications for How CDA Highlights Mystifying Text**

#### *4.3.1 Outline of McClelland and Kawamoto (1986)*

This model aims to display in a simplified way how the capacity of PDP models for

simultaneous interactive processing of syntax and semantics might be utilised in sentence comprehension and in particular in case (semantic) role assignment. McClelland and Kawamoto (1986) were interested to see whether a network, following training, could generalise its learning to new sentences comprised of novel word combinations and thus provide a case-role assignment for these novel sentences. The model is a distributed one and comprises two sets of units: one for representing the syntactic structure of the sentence and one for representing the case role structure. The model is *trained* so as to ‘learn’ the association between sentential input and the desired output of correct case-role assignment. In *testing* (i.e. subsequent to training), the model is presented with surface-structure sentential input and the output the model produces is examined to see if the model has successfully matched case-structure to surface structure. The sentences processed in the model comprise a verb and from one to three NPs of which one is always a subject NP. If an object NP is present, there may also be another NP as a sub-constituent of a PP. An example sentence is ‘the boy broke the window with the hammer’

Words from the input sentences in training are represented as groups of microfeatures. For both nouns and verbs, the features are assembled into dimensions, and for each dimension there are a set of mutually exclusive microfeatures. I detail the dimensions and microfeatures below:

## NOUNS

### *DIMENSION*

### *MICROFEATURES*

HUMAN	human	nonhuman		
SOFTNESS	soft	hard		
GENDER	male	female	neuter	
VOLUME	small	medium	large	
FORM	compact	1-D	2-D	3-D
POINTINESS	pointed	rounded		
BREAKABILITY	fragile	unbreakable		
OBJ-TYPE	food	toy	tool	utensil

furniture

animate

nat-inan

[nat-inan = natural inanimate]

**VERBS***DIMENSION**MICROFEATURES*

DOER	yes	no			
CAUSE	yes	no-cause	no-change		
TOUCH	agent	instrument	both	none	AisP
NAT_CHING	pieces	shreds	chemical	none	unused
AGT_MVMT	trans	part	none	NA	
PT_MVMT	trans	part	none	NA	
INTENSITY	low	high			

[AisP = Agent is Patient; NA = not applicable]

The noun dimensions are fairly self-explanatory. The verb dimensions are regarded as capturing aspects of a scenario designated by the verb:

‘DOER’ - whether agent initiates an event

‘CAUSE’ - whether verb is causal. If not, this dimension indicates whether this is due to an absence of a specified cause ‘the window broke’ or because there is no change ‘the boy touched the girl’.

TOUCH - whether the Agent, Instrument, both, or neither touches the Patient;

AisP - coincidence of Agent and Patient as in ergatives, e.g. ‘The cat moved’

NAT\_CHNG - nature of change in the Patient

AGT\_MVMT - movement of the Agent

PT\_MVMT - movement of the Patient

INTENSITY - forcefulness of action

The dimensions and values are chosen on the basis that they are often salient ones, particularly in the case of verbs, in semantic role ascription. However, the authors are explicit about that fact that these dimensions and microfeatures are not meant to be comprehensive (McClelland and Kawamoto, 1986: 278).

### 4.3.2 How the Model Shades Meaning

In the model, a word is represented by a *vector* (an ordered pattern of distributed representation) in which one microfeature of a dimension is ON and the other is OFF. Microfeatures that are ON are represented in the vectors as 1s. OFF values are represented as dots ('.'). In the training *inputs*, the noun 'ball' was assigned the microfeature SOFT (SO). Going back to the noun dimension table above, the microfeatures for SOFTNESS are SOFT and HARD in that order. So the training *inputted* microfeature value of SOFTNESS for 'ball' was (1 .) (see below and refer back to the noun dimension table above for the significance of HU, GND etc.):

	HU	SO	GND	VOL	FORM	PO	BR	OBJ_TYP
ball	. 1	1 .	.. 1	1 ..	1 ...	. 1	. 1	. 1 .....

However, following training, when the model was *tested* with the sentence *The ball broke the vase*, the *output* for the SOFTNESS dimension of 'ball' was (. 1); the microfeature HARD was activated instead of SOFT. In a sense, this could be treated as an 'error' since the model had been trained with the information that 'balls' were SOFT. However, the adjustment that the model made was perfectly reasonable since all of the other instruments of BREAKING (e.g. 'rock', baseball-bat, 'hammer' etc) were HARD. The model responded to this and shaded its interpretation of the meaning of 'ball' in *The ball broke the vase* accordingly. For the model then, while 'balls' may be SOFT, 'balls-used-for-breaking' are HARD. This property of the model to shade meaning according to lexical environment is impressive to connectionist commentators such as Clark (1989: 109):

...this property, which comes for free with parallel distributed storage and retrieval (at least with all genuinely *distributed* approaches), allows PDP models to provide a mechanism well suited to supporting a variety of

important semantic phenomena. Of all the interesting properties of such models, this one, I believe, most firmly fixes any conceptual or qualitative advantages that PDP might have over other approaches. And indeed, McClelland and Kawamoto (1986: 314) themselves describe the capacity to represent ‘a huge palette of shades of meaning’ as being ‘perhaps...the paramount reason why the distributed approach appeals to us’.

This shading of meaning in the model concurs with the flexibility and holistic grasp that humans enjoy and so in a sense the connectionist network’s ability to shade meaning can be treated as some kind of simulation of what takes place in human brain neural networks:

...if it is the action in us of something operating according to the principles of such networks that enables us to be as flexible and holistic in our grasp of meaning as we are, then the study of such mechanisms surely illuminates how we succeed in grasping the meanings we do. Clark (1989: 108)

A further point to make is that microfeatural shading of meaning is *automatic* in connectionist networks. ‘Shading of meaning’ in humans intuitively appears to be an automatic process. Again, since connectionist networks attempt brain network simulation, then, this lends support to the idea that ‘microfeatural shading of meaning’ in the brain really *is* an automatic process. Or to put this another way, *non-compositional* processing of words in a sentence *is* automatic in a connectionist network, lending support to the notion that *non-compositional* processing of words is automatic in humans too. All this conflicts with the symbolic notion of compositionality in mental representations we saw in the last chapter.

#### 4.3.3 Implications for CDA of McClelland and Kawamoto (1986)

##### *Shading of meaning*

Recall Fairclough’s (1995a: 113) analysis of the following piece of text:

Everywhere in the Third World life in rural areas gets harder - and poor people flock to the city. The urban poor get poorer.



Fairclough (1995a: 113), having commented upon the absence of ‘the poor’ as agents of their circumstances, mentions what seem to be exceptions to this analysis at first glance:

...there are only two Actors in the third extract, the New People’s Army and, exceptionally, the poor, in *the poor people flock to the city*. Interestingly, the Action here is one more usually associated with sheep - notoriously passive - than people, so the exception does not really contradict what I have said so far.

I commented in 3.4.2 how Fairclough isolates ‘flock’ from its semantic-syntactic environment as though it functioned *compositionally* in mental representation as a discrete symbol. However, the capacity for semantic closure and multiple constraint satisfaction of McClelland and Kawamoto’s connectionist model, i.e. *non-compositional / sub-symbolic* processing, is incongruous with Fairclough’s analysis which is essentially *symbolic*. In 3.4.4, I highlighted how Lakoff and Johnson’s (1980) view of metaphor, incorporated into CDA, rested on the assumption of compositional processing. We shall see in chapter 7 that a connectionist approach to metaphor conflicts with Lakoff and Johnson’s (1980) approach.

#### *Attaching Meanings to Syntactic Categories in CDA*

As I have said, McClelland and Kawamoto (1986) were interested to see if their model could correctly assign case roles to sentential input. The *training* input representation of the microfeatures of verbs were the same, regardless of context. For ‘broke’ these were:

	DO	CAU	TOUCH	N_CHG	A_MV	P_MV	IN
broke	1.	1..	.1...	1....	.1..	..1.	.1

The following consists of the microfeature patterns that the model had to choose from in *testing* to see if it could provide a contextually appropriate reading of the verb ‘break’ in a variety of sentences:

	DO	CAU	TOUCH	N_CHG	A_MV	P_MV	IN
broke AVPI	1.	1..	.1...	1....	.1..	..1.	.1
broke AVP	1.	1..	1....	1....	.1..	..1.	.1
broke IVP	1.	.1.	.1...	1....	...1	..1.	.1
broke PV	1.	.1.	...1.	1....	.1..	..1.	.1

[A = Agent, V = Verb, P = Patient, I = Instrument]

Thus, ‘brokeAVPI’ specifies the case frame in which the surface subject is the Agent, the surface object is the Patient, and the with-NP the Instrument. In testing, the model successfully selected the appropriate output case-frame microfeature representations (*slots*) for the following sentences:

the boy broke the window with the hammer	<i>(brokeAVPI)</i>
the dog broke the plate	<i>(brokeAVP)</i>
the hammer broke the vase	<i>(brokeIVP)</i>
the plate broke	<i>(brokePV)</i>

Let me proceed now to consider what McClelland and Kawamoto (1986: 288) say about the success of the model to provide case-frame (‘slot’) representations for the above sentences:

Several things should be said about case-frame representations. **The first thing is that the slots should not be seen as containing lexical items.** Rather, they should be seen as containing patterns that specify some of the semantic properties assigned by the model to the *entities* designated by the words in the sentences. Thus, the pattern of feature values for the verb *break* specifies that in this instance (*the boy broke the window with the hammer*) there is contact between the Instrument and the Patient. This would also be the case in a sentence like *The hammer broke the window*. However, in a sentence like *The boy broke the window*, with no Instrument specified, the pattern of feature values specifies contact between the Agent and the Patient. Thus, the verb features provide a partial description of the scenario described by the sentence. The noun features, likewise, provide a partial description of the players in the scenario, and these descriptions... may actually be modulated by the model to take on attributes appropriate for the scenario in question. [my bold]

The case-frame representations do not, then, contain discrete lexical items. So there are **no syntactically intact** nouns and verbs in the case-frame *output* representations. Indeed, earlier McClelland and Kawamoto (1986: 283) state that ‘the model does not have any prior commitment to the idea that the features in the *input* representation should be preserved in the *output* representation’. Later in the paper, McClelland and Kawamoto (1986: 315-6) flesh this notion out by indicating that the results of their connectionist model support the view that *words (input) are clues to scenarios (output)*, a notion promulgated by Rumelhart (1979):

...words are clues to scenarios...A sentence assembles some words in a particular order, and each provides a set of clues that constrains the characteristics of the scenario, each in its own way. The verb, in and of itself, may specify a range of related scenarios and certain constraints on the players. The nouns further restrict the scenario and further constrain the players. **But the words themselves are no longer present in the scenario, nor is there necessarily anything in the scenario that corresponds to the literal meaning of any of the words...**

**...all the words work together to provide clues to the case frame representation of the sentence, and none of the words uniquely or completely determine the representation that is assigned to any of the constituents of the underlying scenario.** McClelland and Kawamoto (1986: 316) [my bold]

So, from a connectionist perspective if the discrete input words are no longer present in the output, then, nor are the words’ compositional syntactic nature. This all conflicts with the symbolic *compositional* assumption that there is a mental reflex for each constituent of a sentence.

Recall, in the last chapter, how CDA authors associated semantic meaning with the nature of a syntactic category. This is particularly the case with nominals. In CDA, if an event is described with nominalisations, the event is regarded as being objectified, thus mystifying the dynamics of the event, e.g. the causal relations etc; see Kress and Hodge (1979) and Fowler et al. (1979) and more specifically Kress (1989a: 58); Martin (1989: 43); Fowler

(1991: 80); Lee (1992: 95); Fairclough (1995a: 112). Now, for McClelland and Kawamoto's (1986), the units of an expression *conspire together* to produce the scenario and in doing so 'the words themselves are no longer present in the scenario'. Both verb and noun features provide only *partial* excitations and inhibitions of activation patterns, and the individual 'verbs' and 'nouns' of the input are not even present in the output pattern of activation. It follows that, from McClelland and Kawamoto's (1986) connectionist perspective on language processing output, it is inappropriate for (symbolic) CDA to focus too heavily upon the syntactic nature of the units in an expression. Moreover, it is in turn inappropriate to ascribe meaning to the syntactic nature of constituents in a sentence as CDA does. Indeed, from a connectionist perspective, syntactic and semantic information are always integrated (see also: Churchland and Churchland 1996: 238), further ruling out a compositional isolation of a syntactic category and ascribing meaning to it. To sum up: McClelland and Kawamoto's (1986) connectionist model problematises the CDA notion that when nominals are used to describe events, their syntactic nature *mystifies* the dynamics of the event. The integrative nature of semantic and syntactic information also conflicts with the processing assumption of Fowler and Kress (1979a) and Kress (1993) (see 1.4.3 and 3.3) that people begin with a perceptual heuristic of subject-verb-object but may read across this structure and confuse semantic transitivity (transactivity) with syntactic transitivity.

### *Shading of Case Roles*

McClelland and Kawamoto (1986: 312-3) highlight how there are problems with treating case roles as being *unitary* since:

...some but not all of the Patient properties generally hold for the role nominally identified as Patient. Similarly, some but not all of the Agent properties generally hold for the role nominally identified as Agent.

In certain cases, as with sentences like *The boy moved*, enough of these properties hold that we were led to assign *the boy* to both roles at once.

Trying to solve this problem by creating more roles leads to a proliferation:

...that is ungainly, unwieldy, and inelegant, and that detracts considerably from the utility of the idea of roles as useful descriptive constructs.

However, on a distributed representation made up of microfeatures:

If each role is represented by a conjunction of role properties, then far more distinct roles can be represented on the same set of role units. Furthermore, what the Agent roles of two verbs have in common is captured by the overlap of the role features in the representation of their roles, and how they differ is captured by their differences. The notion of a role that represents a combined Agent / Patient as in *The boy moved* is no longer a special case, and we get out of assigning the same argument to two different slots.

McClelland and Kawamoto's (1986) idea of a distributed representation of case-roles (i.e. without unitary status) has also been applied by Touretzky and Geva (1987) and is highlighted in McClelland, St. John and Taraban (1989).

McClelland and Kawamoto's (1986) fine-grained approach to case roles thus necessarily highlights the reductionism of coarse-grained case roles in a Hallidayan meta-language. Of course, the sharpness required of a tool depends on its purpose and if the purpose of a Hallidayan metalanguage, as used in CDA, is to highlight the broad semantic structure of a sentence, then, this is acceptable. However, what often happens in CDA is that meaning is ascribed to the metalinguistic description of a sentence and used to bolster a particular interpretation. Often this meaning is in excess of the actual sentential meaning. To indicate what I mean, consider the following news text and Trew's (1979: 102-3) analysis:

The Riots in Salisbury

The rioting and sad loss of life in Salisbury are warning that tension in that country is rising as decisive moves about its future seem to be in the offing. The leaders of the African National Council have ritually blamed the police, but deplore the factionalism that is really responsible.

No mention is made of 'police' except as those 'ritually blamed' - and note how even in this one reference the syntax has 'Africans' as agents and 'police' as affected participants (the victims of blaming!).

If we were to assign a fine-grained micro-featural profile to 'blame' in the above, we would designate *no-contact* between AGENT and PATIENT. There is also no *movement of the agent* (AGT\_MVMT) nor *movement of the patient* (PT\_MVT). It is not, then, an instance of prototypical material transitivity. Since there is no PT\_MVT, the police are hardly candidates for a reading of AFFECTED where the usual fillers are excited. And because of this, it is entirely tenuous to equate the semantic meaning of the case-role with 'victims'.

The coarse-grained analysis can be seen as projecting an extra meaning which suits Trew's line. We can see, then, that while *grammatical* metalanguage *describes* a clause, a certain type of use of the coarse-grained nature of Halliday's *semantic* metalanguage not only *describes* the clause but can distort the nature of the scenario.

#### 4.4 Connectionism and Inference Generation

##### 4.4.1 Orientation

In this section, I refer to the connectionist models of St. John and McClelland (1992) and of McClelland, St. John and Taraban (1989) which both build upon McClelland and Kawamoto (1986). Both sets of authors deal with connectionist processing of sentences but St. John and McClelland (1992) also deal with short text ('story') processing. Like McClelland and Kawamoto (1986), both models are successful at 'extracting' information

from a training input consisting of syntactically structured representations without resorting to rule-governed internal syntactically structured representations.

The issue I want to highlight in the connectionist networks of these authors is how these deal with *inference* generation. Given the constraints of the thesis and the fact that I have already outlined in some detail a connectionist sentential processing model, I confine myself to the broad principles and results of St. John and McClelland (1992) and McClelland, St. John and Taraban (1989). [The broad principles of St. John and McClelland (1992) and McClelland, St. John and Taraban (1989) derive from McClelland and Kawamoto (1986).]

This is also because in chapter 7, I provide a more detailed description of St. John's (1992) connectionist simulation of inference generation in *short text* processing whose principles are similar to those of St. John and McClelland (1992) as well as McClelland, St. John and Taraban (1989).

#### 4.4.2 *Inference Generation in Sentential Processing*

Both St. John and McClelland (1992: 100) and McClelland, St. John and Taraban (1989: 293) highlight human ability to readily infer other constituents such as instruments when the context is sufficiently constraining. So for example in:

- a) The boy spread the jelly on the bread
- b) The knife was covered with poison

coherence between b) and a) is smooth for the above authors since 'knife' will be readily inferred in a). Now, in the connectionist models of McClelland, St. John and Taraban (1989) and St. John and McClelland (1992), handling implied constituents is not a problem since inferences are an *inherent* part of sentence processing rather than an *extra* process.

McClelland, St. John and Taraban (1989: 316) explain why it was ‘natural’ for their connectionist network to ‘learn’ that *eating steak* always involved a *knife* as an instrument:

There is no special ‘inference step’ required to fill in the knife. This is in part a direct result of the fact that there is no prior stipulation that a particular part of the representation of the sentence corresponds to the internal reflex of each particular constituent of the sentence. It’s just that events described by sentences with ‘ate’ as the verb and ‘steak’ as the object always involve knives as instruments.

The prior stipulation McClelland et al. (1989) refer to is Fodor and Pylyshyn’s classical principle of compositionality (see: 2.2.5). Crucially, even though the network does build up an internal representation (the current sentence gestalt), the representation is not a classical, compositional representation with combinatorial syntax and semantics.

As already outlined, connectionism, in contrast to classical psychology, seeks to understand with reference to underlying *neural* mechanisms. The central idea of parallel distributed processing that information processing arises from the parallel interactions of large numbers of simple information processing elements, comes from the observation that the brain is, in some ways, just such a system. From such a perspective, then, the above models raise the following prospects:

- i) that *humans* can also understand sentences without representing them in a syntactically structured internal ‘language’ (mentalese).
- ii) in *human* processing, *inference generation* is not a special extra step but simultaneously integrative with the integrative processing of semantics and syntax.

But as Bechtel (1996a: 72) points out, some circumspection is necessary:



St. John and McClelland's network can only process a small fragment of English, and it remains a question whether networks of this kind could eventually handle the full range of complexity found in natural human languages. The answer to the question will only come from further empirical investigation.

Crucially, though, Bechtel does add that human ability should not be exaggerated, mistakes being common in the comprehension of complex sentences. Given this, networks should not be expected to perform better than humans when inputted with, for example, sentences with many clausal embeddings (see note 2 of this chapter).

### *Implications for CDA*

The connectionist 'principle' that inferences are inherent to sentence processing and not an extra stage *conflicts with* the downgrading of importance of inference generation in symbolism, the modularity hypothesis etc (see 2.5.3) where syntactic processing is emphasised, e.g. Fodor, Fodor and Garrett (1975). It also conflicts, then, with the neglect of inference generation in (symbolic) CDA in their focus on surface structure, e.g., Fairclough's (1989) analysis of the 'Quarry load-shedding problem' text.

#### *4.4.3 Strength of Interpretation in Short Text Processing*

McClelland and St. John (1992: 97) point out how 'traditional story processing algorithms have viewed comprehension as the sequential building and connecting of text-based propositions' (e.g. Charniak, 1983; Schank and Abelson, 1977; van Dijk and Kintsch, 1983; Wilensky, 1983) as though each sentence were a *context-free proposition*. And indeed in 3.6.2, we saw for instance that for Clark (1992) and Simpson (1993), CDA examination was along these lines. However, St. John and McClelland's (1992) model does not process

sentences in this way. This is because the network exhibits ‘parallel constraint satisfaction’ where interpretation of a sentence is predicated on the strength with which other sentences constrain support for that interpretation. Here are St. John and McClelland (1992: 123):

The strength of support can be observed in how strongly elements of the interpretation are activated. Early in a text, there may be little support for any interpretation. This condition will be manifest in the weak activation of possible interpretations. As text-based support for an interpretation grows, that interpretation will become more active. Further evidence may also revise the interpretation to one that is better supported by the combination of new and old evidence.

Again as in sentence processing, inference generation is *inherent* to the processing of text:

By allowing all of the evidence from the text to bear on the whole interpretation as each proposition is processed, a globally consistent interpretation is more likely to be found.

St. John and McClelland (1992: 124)

So, since the results of computation are not just based on context-free propositions, *extra* computational results (i.e., inferences) are present as a ‘natural’ consequence. On the same basis as before, if connectionist networks can be treated as simulations of brain networks, then, this raises the prospect that *human* text processing works also on the principle of parallel-constraint satisfaction.

### *Implications for CDA*

In chapter 7, I outline an example of a parallel-constraint satisfaction model of short-text processing from St. John (1992) and highlight how the model is able to handle inference generation. But for now, I want to dwell on the basic *principle* of a parallel constraint satisfaction model, i.e. the principle of *global constraints*. In light of what has been dealt with in this section, consider again the text from Clark (1992: 215):

- 1) 'Two of Steed's rape victims aged 20 and 19 had a screwdriver held at their throats as they were forced to submit.
- 2) His third victim, a 39 year old mother of three was attacked at gunpoint after Steed forced her car off the M4.
- 3) Two days later, he gunned down call-girl, Jacqueline Murray, 23, after picking her up in London's Park Lane.'

and what we have already seen from Simpson (1993: 170):

Indeed, so obscured is the relationship between attacker and victim that it allows a possible reading wherein someone else attacks the woman at gunpoint while Steed only forces her car off the road.

From the basic principle of parallel constraint satisfaction, global constraints, the 'weights' for Simpson's alternative scenario that 'someone else might have attacked' would be very low. Simpson's view that it is a possible interpretation is emblematic of the *local* fixation CDA have on individual sentences, as though they were context-free propositions. As we saw in 3.6.2, this fixation is influenced by symbolism, the antithesis of connectionism.

#### 4.4.4 *Summing-Up*

In this section, I have highlighted how connectionist models conflict, on the issue of inference generation, with symbolic accounts where inference generation is downplayed and syntactic processing made prominent. Inference generation is not a special, extra process to syntactic processing but, by nature of connectionist models, *intrinsic* to linguistic processing. A corollary then for CDA, from the point of view of connectionism, and especially if connectionist networks are seen to simulate human processing, is that concentrating on surface structure to the detriment of inferential processes is misleading. Having detailed a model of sentence processing in 4.3, and having highlighted the centrality

of inference generation to connectionist networks in 4.4, let me now widen out the discussion to the issue of mental representation in connectionism and how this differs from the symbolic attitudes to mental representation in CDA.

## 4.5 Connectionism and Mental Representation and Implications for CDA

### 4.5.1 *Gross Descriptivism vs Gross Internalism*

In Clark (1996: 1-2), the following positions are listed in a discussion of whether there are mental analogues to folk discourse:

i) *Gross Descriptivism*: the common-sense constructs (concepts, beliefs, propositionally identified contents etc) are *nothing but* descriptions of large-scale behavioural dispositions of whole agents. According to this view, no neat inner analogies to the folk constructs are to be found.

ii) *Modest Internalism*: the common-sense constructs serve to pick out transient and / or large-scale features of internal (e.g. neural or computational organization). Examples might include the identification of concepts with distributed, context-dependent patterns of neural activity (see Clark, A, 1993) or the identification of mental images with temporarily time-locked activity in multiple neural regions (see Damasio, 1994). In such cases the folk items (images, concepts) do not have neat, highly manipulable and / or spatially localizable inner analogues. But there remain fairly robust patterns of widespread neural / computational activity which the folk discourse at times succeeds in tracking.

iii) *Gross Internalism*: The common-sense constructs (or, in this case, some favoured subset such as concepts or (most) lexical items) have matching, highly manipulable, object-like inner analogues, e.g. a complex thought, folk-psychologically described might thus appear as a complex inner state with independently manipulable parts which match the independently recombinable concepts we deploy in its common-sense characterization.

We can see that CDA, in its highlighting of ‘transactives’ as a privileged representation, chimes with ‘gross internalism’. Consider the following from Simpson (1993: 88):

In this study, the term *transitivity* is used in a much wider sense than that employed in traditional grammars.

Here it refers generally to how meaning is represented in the clause. It shows how speakers encode in language their mental picture of reality and how they account for their experience of the world around them.

Simpson's semantic version of transitivity is more or less the same as Hodge and Kress's (1993) transactivity. It should be apparent that Simpson's notion that speakers encode their 'mental picture of reality' into 'language' is an example of *gross internalism*. Now, for McClelland and Kawamoto (1986: 315-6), 'words themselves are no longer present in the scenario'; for McClelland, St. John and Taraban (1989: 316), 'there is no prior stipulation that a particular part of the representation of the sentence corresponds to the internal reflex of each particular constituent of the sentences.' That is, in their connectionist networks, linguistic input is not the same as output. In contrast to Simpson (1993) above, this position chimes with *gross descriptivism*. So, if connectionist networks in some way simulate brain networks, then, linguistic symbols are less likely to be *encodings* of a mental reality, *pace* Simpson, but rather *input cues* for it.

Now, consider the following:

The propositional attitude statement provides a gloss on the system's state, but not a description of its internal structure. Bechtel and Abrahamsen (1991: 290)

...the currency of our systems is not symbols, but excitation and inhibition.

Rumelhart and McClelland (1986: 132)

Connectionist networks dispense with internalist propositionally based representation consisting of symbols. So, from the vantage point of a connectionist network, CDA's simple - the transactive - would, then, lose its privileged status. This is not only because propositions in connectionism are merely glosses of the system, but because while *symbolic* representations are discrete, highly grammatical and concatenatively compositional (all

features of *transactives*), *connectionist* distributed representations, on the contrary, are continuous, non-grammatical and non-concatenative (Miikkulainen (1993: 21)).

#### 4.5.2 Representation vs Enactment

In chapter 2, I referred to the notion of a ‘simple’ in logical empiricism. The ‘simple’ in logical empiricism is connected with the notion of epistemological foundation. But the desire for epistemological foundations is actually prevalent throughout the history of philosophy, common to both empiricist and rationalist traditions:

[philosophy] understands the foundations of knowledge, and it finds these foundations in a study of man-as-knower, of the ‘mental processes’ or the ‘activity of representation’ which make knowledge possible. To know is to represent accurately what is outside the mind; so to understand the possibility and nature of knowledge is to understand the way in which the mind is able to construct such representations. Philosophy’s central concern is to be a general theory of representation,...

Rorty (1980: 3-4)

But in connectionist philosophy, inspired by a desire to ape in some way the behaviour of the brain, there has been a shift away from concerns of representation. The principal activity of networks is their *self-modification* rather than *representing* the external world. As an outline of such a perspective, consider the following from Schopman and Shawky (1996: 69-70):

...during the course of learning the neural structure reorganizes itself, so that the result is a changed neural structure. This structural change cannot be called a representation of the original input, because the restructuring is the outcome of an interactive process between the state of the organism and the input. **That means that the changed structure does not represent the external world, but it represents - if one wants to stick to the term - the interactive process: input-organism’s or environment-organism’s interaction. Thus one can say that it has a relation with the input, the external world...one can say that the learned structure-change has its own intrinsic semantics; nothing has to be ascribed to it and it requires no external interpretation. The important consequence for our story is that the problem of how to find the semantic relation has evaporated.**

[my bold]

Let me just dwell a moment on ‘the problem of how to find the semantic relation has evaporated’. In 2.2.4, I outlined how logical empiricism strove to mirror a pre-given world that is constituted by well-defined objects with well-defined properties, with the objects standing in fixed relations. This desire, from the connectionist perspective outlined by Schopman and Shawky, is no longer a concern. There is a shift away from the idea of a world being independent to the idea that the world is inseparable from the patterns of self-modification. Like Shopman and Shawky, for Varella et al. (1991: 140) such systems do not work on the basis of representation. Crucially:

...instead of *representing* an independent world, they *enact* a world as a domain of distinctions that is inseparable from the structure embodied by the cognitive system.

Viewing cognition in terms of *enactment*, the realism - idealism dichotomy is circumvented. That is, we no longer have to see linguistic description as ‘mirroring’ the world [realism]; nor do we see our linguistic representations of the world as mediating between the world and what we can know of it [idealism]. Instead, sentences are *enactors* of understanding of the world, not representations [whether realist or idealist] of the world or of cognitive processes. The circumvention of the realist-idealist dichotomy is attractive for another reason. It avoids the realist emphasis on ‘exact fit’. Since ‘connections between actions and the facts of the world can be represented as statistical correlations’ and ‘connectionist reasoning is evidential rather than logical’ (Waltz, D. 1989: 58), instead of *exact match* we have something like *best fit*. A connectionist perspective on mental representation, then, does not require the assumption that the world was contrived to be recognised by humans nor that humans are bestowed with a ready-made key to it all. Sentences are not direct reflections of structures in the external world, *pace* a Carnapian logical empiricist perspective and *pace*

the inadvertent use of this perspective in CDA (see chapter 3).

#### 4.5.3 Connectionism and the 'Representational' Language of Science

For the connectionist philosopher, Bechtel (1996b), rather than being something *innate*, 'our knowledge of grammar, for example, may consist of knowledge of procedures for comprehending and producing sentences in spoken and written speech' (Bechtel, 1996b: 127). On the basis of such a perspective, then:

...written documents are not transcriptions of our mental representations, but specifically constructed representations with which we have learned to interact. (1996b: 128)

That is, natural language *text* provides an *externalist* rather than *internalist* system of 'representations'. Bechtel develops this point by asserting that although scientific theories may take a *sentential* form:

...these representations are not translations of what is in the heads of the scientists; rather, they are devices used by scientists. Scientific theories may take a sentential form even if, in using these theories, scientists rely on weights on connections within their heads. Consequently, we should not seek to localize the story of scientific development in representations and processes occurring in the head. Instead, we need to take seriously the fact that scientists are situated cognizers whose cognitive processes involve interactions with external representations as well as physical devices.' Bechtel (1996b: 141)

All this problematises the analysis of scientific language in Hodge and Kress (1993) (see 3.2.5). This is because their analysis of scientific language is predicated upon the notion that discrete 'scientific' transactive sentences in texts can *reflect* reality which in turn become a mental representation which is *directly reflective* of that reality. Bechtel (1996b) would also conflict with the notion in Ogborn et al. (1996: 51), outlined in 3.5.2, that nominal descriptions of scientific phenomena mean they are treated mentally as 'things':



Scientific texts are well known for their high concentration of events and processes presented as if they were things. Simple examples include evaporation, crystallization, ionization, speciation, oscillation. Any scientific textbook or journal will yield a multitude of them, as transparent as ‘magnification’ or as opaque as ‘commensurability oscillations in the resistivity’ (culled from a relatively non-specialized journal). Their presence is not due to the barbarous linguistic habits of scientists. They exist in texts and talk as entities because they exist in the thinking of scientists as entities.

As a final point, Bechtel’s view of text as not being a transcription of mental representation but rather an external *device* for interaction with a cogniser is on a par with the approach to discourse analysis I outlined in 1.7. It is also in line with the notion of language as a set of cues rather than as a representational medium (see 3.2.5 and 4.5.2).

In 4.5, I have broadened out the issue of mental representation by indicating the poles of gross descriptivism and gross internalism. Gross descriptivism chimes with the positions of connectionist scientists and philosophers I have cited and gross internalism chimes with symbolicists. Two symbolicists I mentioned in chapter 2, Fodor and Pyslyshyn (1988), have written a well-known critique of connectionism. In the final section of this chapter, I highlight this critique and, in the main, how Elman (1990) has tried to answer these criticisms by way of demonstration of connectionist networks. I do this also because Elman’s (1990) response makes very clear the differing perspectives between connectionists and symbolicists with regard to the notions of *compositionality*, *systematicity* and *productivity*, which I outlined in 2.2.5.

## 4.6 Connectionist Responses to Fodor and Pylyshyn's Arguments against the Connectionist Approach

### 4.6.1 Orientation

Fodor and Pylyshyn (1988) argue that only the classical model, one that incorporates a combinatorial syntax and semantics, can account for the *systematicity*, *productivity* and *compositionality* of language. These cannot be accounted for on a connectionist model since connectionist models do not incorporate a combinatorial syntax and semantics. Consequently, connectionism is inadequate on these central characteristics of language. In the following, I outline how the results of some connectionist networks offer a response to Fodor and Pylyshyn (1988).

### 4.6.2 Compositionality

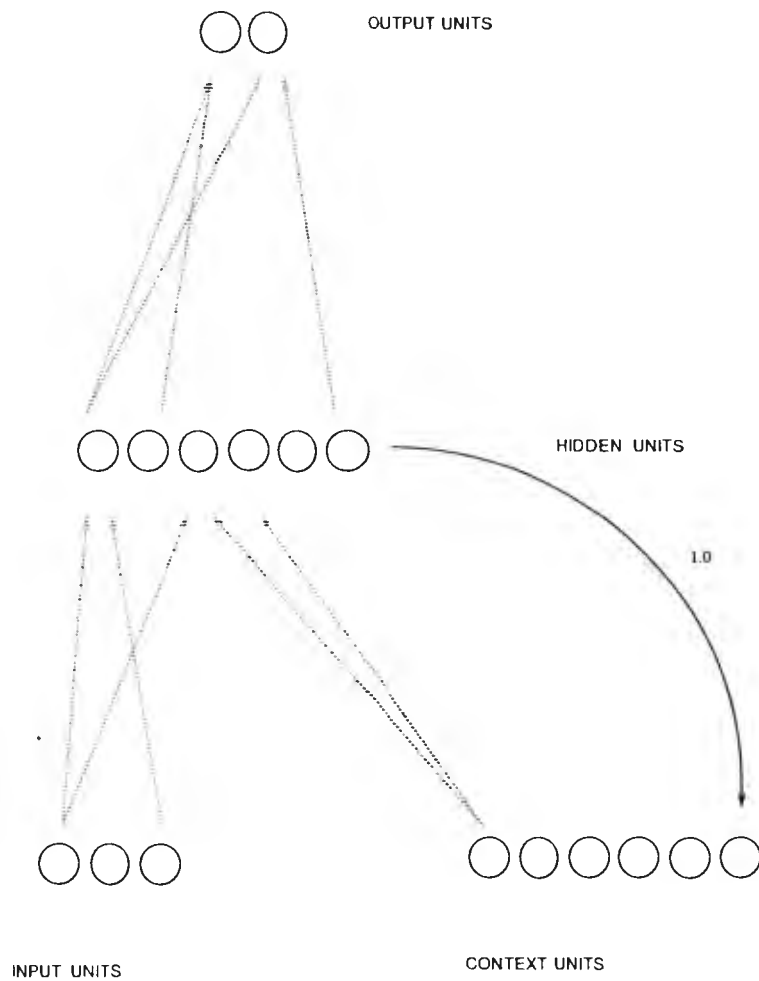
Compositionality refers to the idea that a word contributes the same thing to the meaning of all sentences in which it occurs. Fodor and Pylyshyn (1988) argue that a model of language should be able to account for the phenomenon of compositionality. However, connectionist authors would argue that compositionality is an impoverished view of what happens in processing. We saw earlier that connectionist networks generate implied constituents [St. John and McClelland (1992); McClelland, St. John and Taraban (1989) etc]. This, then, conflicts with the notion of compositionality since there is no prior stipulation that a particular part of the representation of the sentence corresponds to the internal reflex of each particular constituent of the sentence. Furthermore, in McClelland, St. John and Taraban (1989), there was a tendency for the network to be too sensitive to context. [See also 7.4.1 for how the propensity for *top-down* processes of schema expectation to override bottom-up processes is incompatible with the notion of compositional mental representation.]

### 4.6.3 Systematicity

Systematicity refers to the fact that systematic exchange can occur in mental representations to yield new representations. That is, if someone can comprehend ‘John loves the girl’, they can also understand ‘The girl loves John’. For this to take place, ‘the two mental representations, like the two sentences, must be made of the same parts’ (Fodor and Pylyshyn, 1988: 39). Elman (1990) is in part a response to Fodor and Pylyshyn’s criticisms that a connectionist network cannot account for systematicity in language because it does not operate on the basis of a combinatorial syntax and semantics. Elman addresses the question of whether a connectionist model could *induce* lexical-category structure (e.g. the categories of noun and verb) from a series of exposures to sequences of linguistic input. If the network were able to induce structure from different linguistic inputs, i.e. that a verb typically follows a noun, then it could be said to exhibit *systematicity*. That is, the network could ‘understand’ sentences such as *John loves the girl* and *The girl loves John* because it ‘knows’ typically from the corpus sentences it was trained on that one thing (a verb) follows something else (a noun) without actually ‘knowing’ any rules about what constitutes nouns or verbs.

The type of network used is as in the figure c). The network is basically a three-layered network with feed-forward connections from input units to hidden units to output units. There is an additional set of units, context units, which provide for limited recurrence. The context units were set up to copy the activation at the hidden-unit layer, which thus simultaneously received the input and the ‘temporal context’ information (i.e. the copy of its own previous state). Because of this, the network is able to handle syntactic sequence, i.e., recognises that ‘the man likes the woman’ and ‘the woman likes the man’ are not equivalent. A lexicon of 29 nouns and verbs is used, and these were composed into a

**Figure c)**



The network used in the first simulation. Hidden-unit activations are copied with fixed weights (of 1.0) onto linear context units on a one-to-one basis. In the next time step the context units feed into hidden units on a distributed basis.

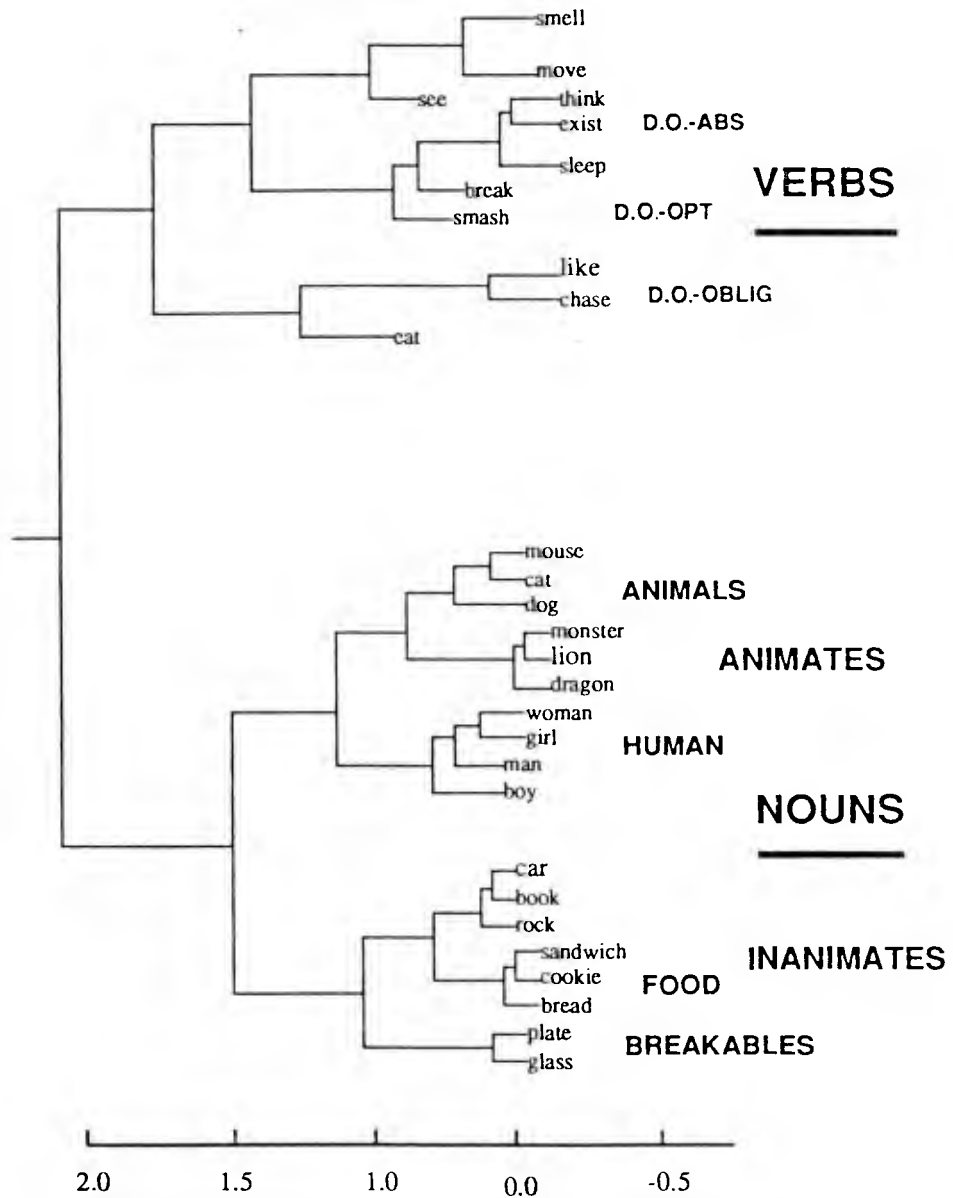
training corpus of 10, 000 two- and three-word sentences. The sentences reflected certain properties of the words. For example, only animate nouns occurred as the subject of the verb 'eat', and this verb was only followed by edible substances. Following training, the task for the network was to take successive words (from an input of words taken from the corpus itself) and to predict the subsequent word by producing it in the output layer. The prediction task was chosen partly on the basis that it does seem that much of what listeners do involves anticipation of future lexical input at least at an immediate local level (Grosjean 1980; Marslen-Wilson and Tyler 1980, Salasoo and Pisoni 1985).

Elman (1990) was not interested so much in what the actual predicted word was in each case but with what 'syntactic nature' the network ascribed to the predicted word. On analysing the internal representations at the hidden-unit layer, Elman found that it was partitioning the space into the recognisable 'syntactic distinctions' 'nouns' and 'verbs'. Figure d) shows the similarity structure of the internal representations of the 29 lexical items. For example, it displays very similar hidden-unit activation for 'mouse', 'cat', 'dog' and thus warranting a cluster label of 'animal'. Elman (1990: 351-3) asserts that this arises:

...from the fact that there is a class of items that always precedes *chase, break, and smash*, it [*the network*] infers a category of large animals' [my italics]

At a coarser perspective, it displayed a similarity in its treatment of *all* nouns in the lexicon. The same was true for *all* the verbs in the lexicon. The verb category is broken down into those that require a direct object, those that are intransitive, and those for which a direct object is optional. As the diagram shows, then, the category structure reflects facts about the possible sequential ordering within sentences in the corpus. That is, the network

Figure d)



Hierarchical clustering of mean hidden-unit vectors after presentation of each of the lexical items in context. The similarity structure of the space reflects distributional properties of the lexical items.

‘recognises’ in this corpus that there are two fairly distinct classes of phenomena, and that one class (verbs) typically follows another class (nouns). Crucially, what Elman’s network demonstrates is its ability to ‘realise’ syntactic / lexical structure, i.e., that internal structure is an *emergent* feature of his simulation without it having to be built into the simulation in the form of *explicit rules*. In a sense, then, the connectionist network did ‘discover’ the categories of *noun* and *verb*. Elman’s network, and thus by extension perhaps also brain neural networks, exhibit *systematicity* in ‘understanding’ sentences such as ‘*John loves the girl*’ and ‘*The girl loves John*’; however, this is not because the two mental representations must be made of the same parts, but because the connectionist network ‘knows’ a verb *follows* a noun from exposure to a large corpus where this is the case. So, while Chomsky invoked the ‘poverty of stimulus’ in linguistic input as support for his argument that language acquisition was only possible with innate linguistic competence, Elman’s network is actually able to induce grammatical distinctions from a limited input (Bechtel, 1996a: 68).

The network’s systematicity, then, has a different basis to Fodor and Pylyshyn’s systematicity and so Fodor and Pyslyshyn’s (1988) stipulation that a combinatorial syntax and semantics must inform systematicity does not follow. However the fact that the network ‘has no information available that would ground the structural information of language in the real world’ is not wholly positive:

This is both a plus and a minus. Obviously, a full account of language processing needs to provide such grounding. On the other hand, it is interesting that the evidence for category structure can be inferred so readily solely on the basis of evidence internal to the language. Elman (1990: 353)

Elman also acknowledges that the model barely scratches the surface ‘in terms of the richness of linguistic phenomena that characterize natural languages’ (Elman, 1990: 376).

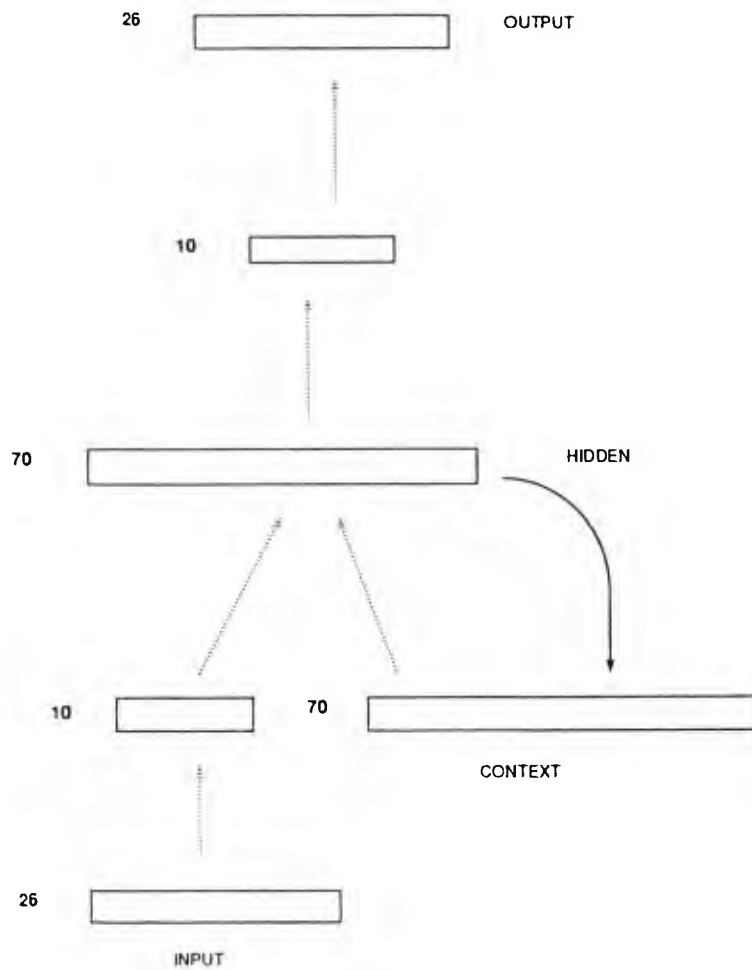
#### 4.6.4 Productivity

Productivity refers to the fact that, despite a finite number of simple concepts, humans can *produce* an infinite set of propositions. Fodor and Pylyshyn (1988) reason that a set of infinite propositions can only be explained on the basis that they are constructed according to structure-sensitive rules via repeatable units. In a further simulation, Elman (1990) showed that a network could acquire a simple grammar by repeated exposure to grammatical sentences, but again without generating or applying structure-sensitive rules, in contrast to the classical approach. Elman (1990) trained a network, similar to the first simulation (see figure e), that is, once more incorporating context units, to discriminate all and only the grammatical sentences generated by a simple but productive grammar over a lexicon of 21 words. The grammar permits considerable complexity in the form of embedded relative clauses as in the sentence 'Boys who kiss girl who feeds dog chase cats'. The grammar demands, as does the trained network, that the plural verb 'chase' agree with the plural subject 'Boys' despite being separated from it by six words, two defining relative clauses, and two distracting singular nouns.

The trained network is presented with a candidate sentence one word at a time, and it yields as output (for that word) a list of grammatical categories permissible, at that point, for the next word in the sentence. In order to do that, it regularly needs to know not just what the current input word is but also what the preceding word was, and sometimes even the two or three words before that. It is the recurrent loop, from the middle layer to the 'auxiliary' input layer, that gives the network this continuous access to its own recent past. In the diagram, what the network returns to along with each new input word is not just the previous word, but rather its subsequent 'digest' at the middle layer, a digest that places the previous word into the context of its preceding word, and so on. Information about earlier words



**Figure e)**



The network used in the second simulation. The architecture is similar to that in figure 1. Additional hidden layers provide for compression of localist input and output vectors.

survives repeated cyclings through that recurrent loop, and thus exercises an influence on the network's response to an input word.

Since the network was able to generate a list of possible grammatical categories at each stage of the input, for Elman (1990), the network could be said to exhibit *productivity*. Again, like the first simulation, the network uses a representational strategy quite different from any classical or rule-based approach which Fodor and Pylyshyn (1988) argue for. That is, Elman's connectionist network exhibits a different kind of productivity to that of Fodor and Pylyshyn (1988). It does not represent any structure-sensitive rules, nor retrieves and applies any such devices in rendering its ongoing verdicts about grammatically permissible next words. For Elman, those rules correctly *predict* the network's behaviour, but they do not *explain* it.<sup>2</sup>

#### 4.7 Endpoints

In this chapter, I have shown how connectionist principles problematise those of symbolism and thus, by extension, how these principles also problematise language processing assumptions in symbolic CDA and how CDA highlights mystifying text. I was not attempting to promote connectionism as a linguistic panacea since connectionism is still in its 'infancy' but rather to show how connectionism problematises the processing assumptions of CDA. However, the ability of connectionist networks to shade meaning non-compositionally and to include inference generation as an integrative part of language processing are features which capture the automatic flexibility and holistic grasp of meaning in human language processing more readily than approaches we saw in chapter 2. This is especially pertinent given that one of the aims of connectionist networks is to provide some

simulation of human brain network activity. In chapter 6, we shall see that the capacity of connectionist networks to tie in with the above aspects of human language processing is in fact supported by psycholinguistic evidence. In the next chapter, I outline the enterprise of cognitive linguistics, an enterprise with many similarities to connectionism, and tease out how cognitive linguistics problematises the symbolic assumptions of CDA also.

## Notes

1. McClelland, St John and Taraban (1989: 322) argue that there is an aspect of the productivity of language which appears to be better explained by a connectionist approach. This is the use of context to shade meanings (see 4.3.2). They refer back to McClelland and Kawamoto (1986) and the example of the 'ball'. In the initial inputs, 'ball' was assigned the microfeature 'soft'. However, with the sentence *The ball broke the vase*, 'hard' was activated instead of soft. That is, the network was able to make necessary adjustments so as to *produce* the novel sentence.

2. Elman's network shows a discrepancy in the type of productivity it manifests and the kind of open-ended generative productivity in Chomskyan (competence) models. I have mentioned that 'rules' predict the network's behaviour. But in reality, this prediction is by no means perfect. This is because the network makes systematic errors. In a long sentence with more than three relative defining clauses, the network loses the ability to specify the correct number of the final verb. This is due to the fact that the subject-term information that is cycling through the recurrent loop is progressively diluted with each cycle: after three-word cycles, the information is effectively gone. Classical algorithms do not show such a failure pattern, but intriguingly humans do and so does Elman's network. This lack of open-ended productivity is not only in marked contrast with Chomsky's 'classical' model, but also strengthens the case for connectionism over the symbolic paradigm, as a better simulation of human language processing, given that limitation on open-ended productivity is a characteristic of human cognition as well.

## CHAPTER 5: COGNITIVE LINGUISTICS - PROBLEMATISING HOW CDA HIGHLIGHTS MYSTIFYING TEXT

### 5.1 Introduction

In 2.4, I mapped out the classical theory of categories. The salient assumption in the classical theory is that categories are defined in terms of a set of necessary and sufficient conditions. In the post-war period, a number of philosophers and linguists, gathering momentum under the banner of 'cognitive linguistics', have begun to challenge this classical postulate. Notable among these challenges is the experimental work of Rosch, and Lakoff's (1987a), 'Women, Fire and Dangerous Things'. Lakoff sets out an alternative philosophico-linguistic position to objectivism which he and Johnson (1987) term 'experientialism'. He terms the classical theory of categories - *objectivism*. I will come to an outline of this position in due course. But to begin this chapter, I draw attention to two challenges cognitive linguistics poses to objectivism. In objectivism:

- i) no members should be better examples of the categories than any other members.
  
- ii) categories should be independent of the peculiarities of beings doing the categorising i.e. should not involve neurophysiology, human body movement, and human capacities to perceive, to form mental images etc.

Cognitive linguistics is founded on the rebuttal of the absoluteness of i) and ii). In 5.2 and 5.3, I discuss the cognitive linguistic dissension from i) and ii) respectively.

## 5.2 Categories and Prototypicality

### 5.2.1 Prototype Effects

#### *Early Rosch*

In the early to mid-seventies, the results of a series of experiments conducted by Eleanor Rosch cast doubt on the notion that category membership was along the lines of Aristotelian absoluteness (see 2.4.1). In Rosch (1975b), for example, students were asked to assess the ‘typicality’ of members of certain categories. Subjects tended to define categories (e.g. ‘bird’) by identifying certain prototypical members of the category (e.g. ‘robin’), where there exists the greatest density of attributes for the category, and they recognise other non-prototypical members (e.g. ostrich) that differ in various ways from the prototypical ones.

Rosch’s experimental results, then, conflict with the assumption inherent within the classical theory that no members should be better examples of the categories than any other members. This is not to say that all of the objectivist or classical theory is challenged by Rosch’s results as Lakoff (1987a: 586) points out. Certain geometrical shapes such as squares and spheres etc can be described by a complete set of necessary and sufficient conditions, although what constitutes the expected size of such shapes, for example, may be grounded in prototypicality (Armstrong et al., 1983).

Referring to Rosch’s results, Lakoff (1987b: 63) highlights how prototype effects were interpreted in two ways as indicating something *direct* about the nature of categorisation:

- i) *The Effects = Structure Interpretation*: Goodness-of-example ratings are a direct reflection of degree of category membership.
- ii) *The Prototype = Representation Interpretation*: Categories are represented in the mind in terms of prototypes (that is, best examples). Degrees of category membership for other entities are determined by their

degree of similarity to the prototype.

### *Later Rosch's More Tentative Position*

In the early part of her career, Rosch accepted the above as valid interpretations of her experimental results. Later, Rosch (1978: 40-1) was to form a much more tentative position:

The pervasiveness of prototypes in real-world categories and of prototypicality as a variable indicates that prototypes must have some place in psychological theories of representation, processing and learning. However, prototypes themselves do not constitute any particular model of processes, representations, or learning. This point is so often misunderstood that it requires discussion:

1. To speak of a *prototype* is simply a convenient grammatical fiction; what is really referred to are judgments of degree of prototypicality...
2. Prototypes do not constitute any particular processing model for categories...
3. Prototypes do not constitute a theory of representation for categories...As with processing models, the facts about prototypes can only constrain, but do not determine, models of representation...
4. Although prototypes must be learned, they do not constitute any particular theory of category learning.

Lakoff (1987a: 44-5) elaborates upon this position. In the early to mid 1970s, empirical goodness-of-example ratings were taken as supporting the notion that a penguin is less a member of the category 'bird' than a robin. But this was later regarded as a mistaken interpretation of the data because the responses by subjects as to the goodness-of-example are just *ratings*. And indeed as Lakoff (1987a: 45) states, the ratings:

...are consistent with the interpretation that the category *bird* has strict boundaries and that robins, owls, and penguins are all 100 percent members of that category. However, that category must have additional internal structure of some sort that produces these goodness-of-example ratings.

Lakoff (1987a: 45), then, tallies with Rosch's circumspection. He goes on to state that prototype effects are really 'superficial' in the sense that they are *by-products* of 'internal structure'.

### *Idealised Cognitive Models*

One of the aims of Lakoff (1987a) is to outline the sources for prototype effects. For Lakoff (1987a: 45), 'prototype effects result from the nature of cognitive models, which can be viewed as 'theories' of some subject matter.' The central thesis of Lakoff (1987a) is that human knowledge is organised via *idealised cognitive models* (ICMs) and that prototype effects (as well as category structure) are *by-products* of such cognitive organisation. What does Lakoff mean by *idealised*? He means that the cognitive model is *created*. For example, if the Western seven-day calendric cycle is compared with that of the Balinese, it becomes easier to see that the Western model is an idealised model since it does not exist independently in nature (Lakoff, 1987a: 68-9). How, for Lakoff, do prototype effects arise? Lakoff (1987a: 70) gives the example of the category 'bachelor', arguing that it is defined with respect to an ICM of human marriage in a typically monogamous human society, and a typical marriageable age. It is as a result of this ICM that fuzzy cases can arise: homosexuals, the pope<sup>1</sup> etc. For Lakoff (1987a: 70) an ICM:

...may fit one's understanding of the world either perfectly, very well, pretty well, somewhat well, pretty badly, badly, or not at all. If the ICM in which *bachelor* is defined fits a situation perfectly and the person referred to by the term is unequivocally an unmarried adult male, then he qualifies as a member of the category *bachelor*. The person referred to deviates from prototypical bachelorhood if either the ICM fails to fit the world perfectly or the person referred to deviates from being an unmarried adult male.

Prototype effects arise because of the degree to which the ICM fits our knowledge (or assumptions) about the world. They arise in the comparison between two cognitive models in the above - one for *bachelor* and one characterising one's knowledge about an individual, say the pope. To allow such a comparison, Lakoff (1987a: 71) thus argues that 'one needs the concept of 'fitting' one's ICMs to one's understanding of a given situation and keeping track of the respects in which the fit is imperfect.'<sup>2</sup>

*Barsalou (1983)*

For Lakoff (1987a: 45-6), the work of Barsalou (1983) confirms that prototype effects arise from the nature of cognitive models. Barsalou focuses on ‘ad-hoc categories’ - non-conventional categories which are created ‘on the fly’ for some particular *goal* in a particular *context*, e.g. ‘things to take from one’s home during a fire’. Certain features (e.g. one’s children) can be regarded as prototypical and others (e.g. a box of paper clips) as non-prototypical. However, despite the fact that the category is non-conventional, it can be said to have a prototype structure. For Barsalou, such a non-conventional category is determined by goals and goals are part of a speaker’s cognitive models. From a Lakoffian perspective then, prototype effects arise when the ‘on the fly’ created ICM fails to adequately fit the understanding and goals of the situation. The importance also of goals in the type of prototype effects created becomes clear when we see that in Rosch’s early experiments the respondents’ *goal* was to deliver ratings of typicality of particular categories.

*Context (co-text) -Dependency Of Prototype Effects*

What Barsalou’s non-conventional examples also indicate is that prototype effects are context-dependent. Context can also include co-text. Co-textual influence on prototype generation is appreciable in the following taken from Ungerer and Schmid (1996: 43-4) (see also Roth and Shoben, 1983):

- i) The hunter took his gun, left the lodge and called his dog
- ii) She took her dog to the salon to have its curls reset

The first dog, prototypically would be a retriever, while the prototype for the second sentence might be a pekinese. Outside of the above co-texts, a retriever could be construed



as a prototypical dog but a pekinese is unlikely to be. What these examples indicate is that Rosch was *right* to adopt a more tentative position since the influence of context (including co-text), engagement of cognitive models and specific goals indicate that prototypes are created ‘on the fly’ rather than being stored representations; see also Rosch (1978: 42-3) for a discussion of the role of context in prototype effects. All this is not to deny the *reality* or salience in human reasoning of prototype effects, only that these effects depend on other phenomena for their *emergence*. The importance of prototypes and their relation to human reason is indicated in the following from Lakoff (1987a: 45):

In many cases, prototypes act as *cognitive reference points* of various sorts and form the basis for inferences (Rosch 1975a, 1981). The study of human inference is part of the study of human reasoning and conceptual structure; hence, those prototypes used in making inference must be part of conceptual structure.

The importance of prototypes to inference generation will become apparent in section C of this thesis. We shall see also in section 5.4 that the internal structure upon which superficial prototype effects depend relates to our neurophysiological propensities for motor-interaction and gestalt perception.

Having introduced the phenomenon of prototypicality, in the next section I outline the relationship between prototypicality and syntactic categories and highlight implications for CDA.

### *5.2.2 Prototypicality and Syntactic Categories: Implications for CDA*

The work of Hopper and Thompson (1985) has indicated that the status of a word within its respective grammatical category is by no means a fixed property of the word in question.

The semantically relevant properties - in the case of nouns, the extent to which the noun

refers to an identifiable, enduring thing, with verbs whether the verb refers to a specific dynamic event - can vary. On a similar note, here also is Lakoff (1987a: 57)

...it is important to see that prototype effects occur not only in nonlinguistic conceptual structure, but in linguistic structure as well. The reason is that linguistic structure makes use of general cognitive apparatus, such as category structure.

Consider the following sentences:

- i) The lorry shed its load on the building site
- ii) Load-shedding is frequent on the building site

In i), 'shed' and 'load' are fairly prototypical. Consistent with this is the fact that both verb and noun can take on the whole range of typical verb and noun properties. The verb is marked for tense, polarity, mood and voice. 'Load' could also appear as singular or plural, and be preceded by a determiner, adjective(s) etc. In sentence ii), though, 'load' does not refer to a discrete identifiable object, and neither does 'shed' refer to a single identifiable event. Symptomatic of this loss of semantic categoriality is the fact that neither word can be inflected or modified. It follows that 'load-shedding' is a non-prototypical noun. Recall from 1.4.2, Fairclough's (1989: 51) analysis of the headline 'Quarry load-shedding problem':

...the grammatical form in which the headline is cast is that of a *nominalization*: a process is expressed as a *noun*, as if it were an entity.

The assumption here is that mystification can occur because the reader is prevented from seeing reality in *active* terms. Understanding of the *dynamism* of the reality (i.e., the nature of causal relationship and agent responsibility) is prevented by the 'objectifying' effects of

the nominalisation. But we have seen that, on the basis of Hopper and Thompson's analysis, 'load-shedding' suffers decategorisation. In also functioning as a modifier, it loses some of the morphological and distributional attributes of the noun class. In chapter 3 and 4, we saw that much of CDA's approach to mental representation was predicated upon the *internalist* view of mind. Adopting this CDA-ratified perspective, since 'load-shedding' is a non-prototypical nominal, the reader's mental representation of 'load-shedding' would not for a CD analyst, at least, be that much of an 'entity'.

The oddness of seeing all nouns as entities can be seen by considering a noun such as 'departure'. On Hodge and Kress's argument, readers would have to override its semantic meaning by treating the nominal 'syntactic meaning' as completely separate and somehow make the idea of, say, the 'train departure' difficult to understand as an event. The absurdity of such a suggestion is even more evident when we consider the lexemes 'verb' and 'nominalisation', the latter being a central facet of Hodge and Kress's conceptual apparatus. Does the nounhood status of 'verb' impede understanding that a great number of verbs are processes? Does the nounhood status of 'nominalisation' prevent understanding that this term refers to a process - i.e. the process of making verbs into nouns? On Hodge and Kress's reasoning the answer would have to be in the affirmative, but this paradoxically would problematise their use of the concept 'nominalisation' since its grammatical form would mystify understanding of how such processes are supposed to mystify! The problematising I applied to Fairclough can also be applied to Fowler's (1986: 20) treatment of 'my wife' (see 3.4.3). In other words, Fowler ignores the fact that there are prototypical uses of 'my' to indicate possession (e.g. 'my wallet') as well as non-prototypical uses ('my round of drinks') where possession diminishes as a feature of the situation. Similar points can be made against how Hodge and Kress (1993: 23-4) regard nouns such as 'ban' and

‘production’ (see 3.5.2):

...as though they were like *apple* or *bench*, but referring to things which happen to be abstract, not concrete physical things.

In the light of all this, and as I argued in chapter 4, it is unwise to simply ‘read off’ metalinguistic descriptions of syntax as is common in CDA.

Recall from 5.1 how in objectivism:

Categories should be independent of the peculiarities of beings doing the categorising i.e. should not involve neurophysiology, human body movement, and human capacities to perceive, to form mental images etc.

I now move to a discussion of the basic level of categorisation, a type of category salient in cognitive linguistic explanations, which is linked to human neurophysiology and thus in tension with the above precept of objectivism.

### **5.3 Basic-Level Categories**

#### *5.3.1 Orientation*

To begin this section, let me outline the work of Berlin et al. (1974). The aim of this project was to scrutinise the ‘folk taxonomy’ used by Tzeltal people, a community in southern Mexico, for classifying and naming the plants in their environment and compare this taxonomy with Western scientific classifications. It was found that Tzeltal plant categories were most numerous on the *generic* level with 471 genera, e.g., corn, bean, pine, willow.

In contrast, the number of *superordinate* categories was exceedingly low - no more than

four plant names - tree, vine grass, broad-leafed plant. 'Species' was well represented but membership was more restricted than on the generic level, there being 273 species, e.g. genuine pine, red pine, white bean, common bean etc. The *subordinate* level to the species level was minimal, consisting of five plant names, e.g. red common bean, black common bean. Generic categories were not only more numerous but were the ones that were most commonly chosen by Tzeltal speakers. Moreover, in Tzeltal, *culturally* salient categories are much more likely to be generic categories than superordinate ones. For example, 'corn' and 'beans' form two basic ingredients of the Tzeltal diet.

What is interesting about the folk taxonomy in Tzeltal compared to folk taxonomies used by English speakers is that the latter's cognitively basic level is also the generic level. Generic category names such as 'dog', 'car' are first learned by children, are used frequently and consist of simples (i.e. undecomposable morphemes) (Brown, 1958, 1965). This also applies for Tzeltal generic categories. Speakers also tend to prefer category names such as 'dog' and 'car' in neutral contexts or when introducing new categories into conversation (Cruse, 1977). Brown's observations about the use and acquisition of cognitive categories together with Berlin et al's investigation of the Tzeltal seems to indicate that the generic or intermediate level of categorisation is more important. This is partly due to the fact that categories at this level not only have cultural salience but *biological* salience also. Such categories are used to refer to objects and actions which are bound up with motor interaction (Lakoff, 1987a: 37). Following Lakoff (1987a), I shall refer to the salient generic-level of categorisation as the *basic level*.

### 5.3.2 Direct Understanding of the Basic Level

For Lakoff (1987a: 302), meaningfulness involves the structuring of experience. Some types

of experience are preconceptually structured because of our neurophysiological capacities for motor interaction, gestalt perception and mental imagery. For Lakoff (1987a: 292), ‘a sentence is directly understood if the concepts associated with it are directly meaningful’, i.e., if it involves *basic-level* categories. As an example, consider the difference in direct understanding between *the pig ate the carrot* (which uses basic-level categories) and *the organism ingested the food* (which uses superordinate categories).

It is because of the capacity of basic-level categories to give rise to direct meaningful sentences that basic-level *metaphors* are comprehensible. Here is Lakoff (1987a: 303):

In domains where there is no clearly discernible preconceptual structure to our experience, we import such structure via metaphor. Metaphor provides us with a means for comprehending domains of experience that do not have a preconceptual structure of their own.

But basic-level metaphors may not only communicate conceptual content but *interactional* properties also. As an example of a basic-level metaphor, consider the ‘internal mouse’ on a laptop computer. It has become common to refer to the small mound in the keyboard as the ‘mouse’ as an extension of ‘external mouse’, itself a metaphor. Obviously this piece of rubber bears no relationship to the ‘external mouse’, i.e. it does not have the inherent properties - made of hard plastic, has a ball mechanism underneath etc. But referring to this laptop item with the *basic-level* category ‘mouse’ enables the hearer to realise the particular motor-interactionality and purposive properties associated with a laptop. That is, in context, when instructed as to the name of the small mound, perceptual properties (it looks like a mouse) are inhibited and purposive properties (it serves the purpose of an external mouse) and motor-activity properties (with some adjustment, you handle it like an external mouse) are activated.<sup>3</sup>

### 5.3.3 Cognitive Economy of the Basic Level

The basic level is where perception of obvious differences amongst organisms and objects is salient, consonant with ‘the most obvious discontinuities in nature’ (Kay 1971: 878). Experimental evidence suggests that there are more attributes associated with the basic level than with the *superordinate* level (Rosch et al. 1976). In other words, the basic level is *informationally rich*. Compare, for example the basic-level category ‘dog’ to the superordinate category ‘mammal’. In contrast to the basic level, it should be apparent that the superordinate covers such a disparate array of items that any similarities are only apparent from a general perspective, i.e., there is no common shape for the category which could be applied to dogs, elephants, giraffes etc; see section 5.6 for discussion of the superordinate level. Conversely, we find the basic level easy to understand because we understand objects on this level of categorisation (e.g. dog) in terms of an overall shape, that is holistically via a gestalt (Lakoff, 1987a: 33). Like basic-level categories, *subordinate* categories (e.g. ‘retrievers’) do possess a common defining shape. However, the differentiating power of subordinates does not match that of basic-level categories. Differentiation between ‘retrievers’ and ‘poodles’ is much less than between ‘dogs’ and ‘giraffes’ for example (Ungerer and Schmid, 1996: 69).

As I have said earlier, it is on the basic level that we *motor-interact* with objects and organisms, e.g. where cats are stroked, spoons are held etc (Rosch et al. 1976). Moreover, it is on the basic level that objects and organisms can be *distinguished* by the differences in how humans interact with them. For the *subordinate* level, though, ‘it is difficult to imagine that different kinds of cats are stroked in different ways’ (Ungerer and Schmid, 1996: 69).

Conversely, it is difficult to distinguish between superordinate categories (e.g. ‘furniture’ and ‘mammals’) in terms of one’s motor-interaction.

To sum up, compared to the superordinate and subordinate levels, i) we *readily* understand basic-level objects in terms of their overall shape, holistically via a gestalt; ii) we *readily* make distinctions between basic-level objects in terms of the differences in our motor-interactions with these objects. So, the basic level is not only *informationally rich* but also more *readily* yields information compared to the superordinate and subordinate levels. In other words, the basic level is the category level where the *largest* amount of information about an item is understood with the *least* cognitive labour (cf ‘Relevance Theory’: Sperber and Wilson, 1995). This characteristic of the basic level has been termed *cognitive economy* (Rosch, 1978) and, as Ungerer and Schmid (1996: 68) assert, this phenomenon ‘probably explains best why the basic level is particularly well suited to our cognitive needs.’

#### 5.3.4 Challenging Objectivism

Let me now indicate how what I have outlined challenges the objectivist approach to categories that was outlined in chapter 2. For Lakoff (1987a: 32), Berlin’s research into the Tzeltal taxonomy can be considered as a response to the classical doctrine of *natural kinds* - that to a large degree the world comprises natural kinds of things and that languages consist of names (‘natural kind terms’) that *fit* those natural kinds. Typical natural kinds would include cows, dogs, tigers, gold, silver, water, etc. Cognitive linguistics takes issue with the doctrine of natural kinds since it highlights how we *interact* with the world rather than how categories naturally fit the world independent of the user. Lakoff and Johnson (1980: 119-121) provide a good example of how objectivist views of categories omit any treatment of *interactional* properties by referring to the concept of a ‘fake’ gun. On an objectivist entailment, ‘This is a black gun ENTAILS this is a gun’ and ‘This is a fake gun ENTAILS this is not a gun’. Fakeness is taken to be an inherent part of the gun, i.e. it will not fire real bullets. But the gun’s fakeness is not so simple. This is because while ‘fake’ negates a



gun's inherent properties, it still *preserves* interactive properties. I elaborate upon Lakoff and Johnson (1980: 121) in the following:

FAKE preserves:	perceptual properties (a fake gun looks like a gun)
<i>INTERACTIVE</i>	motor-activity properties (you handle it like a gun)
<i>PROPERTIES</i>	purposive properties (it serves some purpose of a gun - to threaten or scare).
FAKE negates:	mechanical properties (a fake gun doesn't shoot)
<i>INHERENT</i>	
<i>PROPERTIES</i>	

Lakoff and Johnson suggest that from this we conceptualise a gun in terms of a multidimensional gestalt of properties where the dimensions are perceptual, motor-activity, purposive, functional, mechanical etc. What is interesting is that most of these properties are not an inherent part of the gun but *interactional properties*. This causes a problem for the objectivist Aristotelian view of 'essence' (see discussion in 2.4.1) and thus for deciding on set membership of a category according to attributes of an object. Objectivism discounts the mind of the observer since, from an ontological viewpoint, it only deals with what Searle (1995) calls the *ontologically objective*. The interactional properties of an object are what Searle terms the *ontologically subjective* [although *ontologically interactional* would be more felicitous]. One very well-known view of meaning associated with the Wittgenstein (1953) of the 'Philosophical Investigations' is that the meaning of a word is bound up with its *use*. Certainly, this is a valid observation. But it can be made more specific from the perspective of cognitive linguistics. It needs to be supplemented with an account of how *basic-level* categories are associated with motor-interactional *use*.<sup>4</sup>

From the perspective of objectivism, it would be expected that the categories that were

easiest to process would correspond to conceptual primitives which by definition would have no internal structure. However, basic-level categories are centrally placed in taxonomic hierarchies and possess a fair amount of internal structure. Even though they are not, then, primitives, they actually have the structure which humans find easy to process, easy to learn, remember etc. As Lakoff (1987a: 199) says: 'In short, what should be cognitively complex from an objectivist point of view is actually cognitively simple.' The ease of processing stems from their connection with motor-interactivity and thus ontologically subjective properties. But the (objectivist) predicate calculus description of the hyponymy sense relation,  $\forall x[C(x) \rightarrow D(x)]$ , where C could be a basic-level category like 'chair' and D a superordinate like 'furniture' reflects only ontologically objective properties.

The relatively new perspective on categorisation, which takes into account the ontological subjectivity associated with basic-level categories, Lakoff (1987a) terms *experientialism*. In experientialism, *meaningful* thought is *embodied*, conceptual systems growing out of bodily experience and the interactions of the body with the environment. Since in experientialism, language is bound up with other aspects of cognition, it conflicts with the modular hypothesis promoted by Chomsky that syntax is a *separate* system independent of the rest of cognition (Langacker, 1987a; Lakoff, 1987a: 225-6). On these grounds, then, cognitive linguistics problematises the attributing of meaning in CDA to the nature of a syntactic category *separate* from its *actual* semantic meaning. Thus, what was problematised on these grounds by connectionism in chapter 4 is also problematised by cognitive linguistics, e.g., the assumption of Kress (1993) (see 1.4.3) that people may read across the syntactic structure 'subject-verb-object', and in doing so, inadvertently bestow the *semantic* structure of 'agent-process-patient' even if this is not actually the semantic structure of the clause. In chapter 7, I draw out more explicitly similarities between

connectionism and cognitive linguistics. However, I also explore tensions between connectionism and cognitive linguistics which problematise the approach to *metaphor* in Lakoff and Johnson (1980) and Lakoff (1987a) and its use in CDA.

### *5.3.5 The Role of Context and the Basic Level*

As a coda to this section on basic-level categorisation, it should be stressed that degree of expertise of a particular domain will affect what is regarded as the basic level, and accordingly the basic level is fluid in accordance with expertise and cultural models. As Harley (1995: 195) points out:

Birdwatchers, for example, know nearly as much about the subordinate members such as blackbirds, jays, and Dartford warblers, as they do about the basic level.

Similarly, Rosch (1978: 42) indicates that a man in a furniture store surrounded by an array of chairs will obviously be speaking and thinking at a level subordinate to the basic level.

In the next section, I demonstrate the link between the basic level and prototypes.

## **5.4 The Symbiosis of the Basic level and Prototypes**

### *5.4.1 Orientation*

I mentioned in 5.2 that prototype effects are superficial, based on internal structure. I now want to look at one provenance for these effects - our neurophysiological capacity for motor-interaction, gestalt perception, image generation. It is because these give rise to basic-level

effects that prototype effects are associated with the basic level. Here are Ungerer and Schmid (1996: 72):

The basic level provides the largest amount of relevant and digestible information about the objects and organisms of the world (e.g. information about bird-like animals) or, to put it more technically, it offers the largest bundles of correlated attributes. These attributes are accumulated in their most complete form in the prototype (ROBIN in the case of BIRD) and expressed by the category name (e.g. *bird*).

Another reason is because:

The basic level is where the overlap of shapes is so great that it permits reliable gestalt perception, which is particularly easy for prototypical examples (like the ROBIN).

The symbiosis between prototypicality and basic-level categorisation can be seen clearly if a basic-level expression is used to refer to a non-prototypical instance. Here is Lakoff (1987a: 452):

...if a sparrow lands on the front porch, it is not misleading to report this by *There's a bird on the porch*. But it would be quite misleading to use such a sentence to report that an eagle had landed on the porch or that a penguin had waddled up the front steps. Similarly, if John hit a baseball with a bat in the usual way by swinging the bat at the ball, we could straightforwardly report that *John hit a ball*. But if he hit a beachball with a pizza platter, or if he hit a ball by throwing a rock at it, it would be misleading to describe such an event to someone who didn't see it as *John hit a ball*, even though such a description, strictly speaking, would be true. *Hit a ball* has an associated conventional image that characterizes the normal case, and with no further modification we assume that the normal case holds. Thus, conventional images are used to understand even the simplest, most straightforward sentences with no idioms in them.

A corollary of all this is that prototypes are more likely to be generated from basic-level categories than non-basic-level categories. To understand the readiness of generation of prototypes from the basic level compared to the superordinate level, again compare 'the pig ate the carrot' with 'the organism ingested the food'.

### 5.4.2 Causation as Interaction

I now come to making explicit the relationship between prototypes and the basic level of categorisation with regard to *causation*. For Lakoff (1987a: 54) prototypical causation is understood in terms of a cluster of ‘interactional properties’ since prototypical causation ‘appears to be direct manipulation’. The cluster acts as a gestalt which is psychologically simpler than its parts. I list the properties of prototypical causation below:

1. There is an agent that does something.
2. There is a patient that undergoes a change to a new state.
3. Properties 1 and 2 constitute a single event; they overlap in time and space; the agent comes in contact with the patient.
4. Part of what the agent does (either the motion or the exercise of will) precedes the change in the patient.
5. The agent is the energy source; the patient is the energy goal; there is a transfer of energy from agent to patient.
6. There is a single definite agent and a single definite patient.
7. The agent is human.
8.
  - a. The agent wills his action.
  - b. The agent is in control of his action
  - c. The agent bears primary responsibility for both his action and the change.
- 9. The agent uses his hands, body or some instrument.**
10. The agent is looking at the patient, the change in the patient is perceptible, and the agent perceives the change. [my bold]

Lakoff (1987a: 55) continues by highlighting how the most representative examples of ‘humanly relevant causation’ possess all of the above properties, citing the examples ‘Max broke the window’, ‘Brutus killed Caesar’. I have bolded condition 9 - that the agent uses his hands, body or some instrument - to highlight the link between the *basic level* of categorisation, prototype theory and causation. Prototype effects, with regard to causation, are *by-products* of our internal neurophysiological capacity for *motor-interaction*, which in turn is associated with the basic level. For Lakoff (1987a: 51):

... interactional properties form *clusters* in our experience, and prototype and basic-level structure can *reflect* such clusterings. [my italics]

Lakoff (1987a: 55) continues:

Many languages of the world meet the following generalization: The more direct the causation, the closer the morphemes expressing the cause and the result. This accounts for the distinction between *kill* and *cause to die*. *Kill* expresses direct causation,<sup>5</sup> with cause and result expressed in a single morpheme - the closest possible connection. When would anyone ever say 'cause to die'? In general, when there is no direct causation, when there is causation at a distance or accidental causation...

From this Lakoff concludes that:

...the best example of the *conceptual category* of causation is typically marked by a grammatical construction or a morpheme and that the word *cause* is reserved for noncentral members of the conceptual category.

### 5.4.3 Implications for CDA

Recall from 1.4.2 the following from Martin (1989: 43):

The CWF article uses three types of nominal structure in place of verbs to realise actions. One puts the action into the modifier of an abstract noun: e.g. *sealing operation*, *killing techniques*. Another makes use of a nominalised form of a verb: *statements*, *definition*, *death*, *coverage*, *constraints*. A third simply realises the actions as a noun: *the whitecoat harvest*, *the East Coast seal hunt*, *the seal hunt*.

...The CWF text...tends to refer to the killing indirectly, using incongruent forms: *killing techniques*, *the whitecoat harvest*, *the slaughter of animals*, *the East Coast seal hunt*, *a slaughtering operation*, *killing methods*, *an almost instantaneous death*, *a humane death*, *the seal hunt*, and so on. In this way the ACF text focuses on the process of killing, while the CWF text treats the killing as a kind of thing. This has the effect of immobilising the most unsavoury part of the seal hunt and helps draw attention away to other 'factual' considerations.

I reproduce this extract again as it neatly encapsulates symbolic assumptions which are scattered elsewhere in CDA. I deal with one of these immediately below and refer to the other two later in this chapter. From a cognitive linguistic perspective, 'killing techniques'

can hardly be said to be mystifying of the ‘unsavoury part of the seal hunt’. Since ‘kill’ encapsulates both cause and effect (Lakoff, 1987a: 55) and is associated with a high degree of interactionality, the *syntactic position* of ‘killing’ in ‘killing techniques’, and its capacity to mystify, diminishes in significance.

Recall Hodge and Kress’s (1993: 8) stipulation (3.2.1) that transactives are a privileged mode of representation where there is ‘action going from an actor to an affected’. Hodge and Kress’s (1993) viewpoint is merely *linguistic*. For them, the agent and patient of the actual circumstances and the process that links them should be indicated by the presence of the *semantic roles* of agent and patient linked adjacently in a clause. We saw in chapter 3 how their focus assumes that graphic linguistic structure is cognitively reiterated intact into *linguistic* mental representation. However, from a cognitive linguistic perspective, the rationale for highlighting transactives in Hodge and Kress (1993) does not include the notions that prototypical causation involves: an experientially simple gestalt, basic-level categories, and ‘direct manipulation’ via hands, body, an instrument etc. Their focus ignores the broader *cognitive* (interactional) perspective of processing in which syntactic structure is only *one* aspect.

As a more concrete example, recall the following from Trew (1979: 98-9), which I highlighted in 1.4.1 and referred to in 3.6.2 in order to indicate symbolic assumptions of processing in CDA:

Eleven Africans were shot dead and 15 wounded when Rhodesian Police opened fire on a rioting crowd of about 2,000.

We saw that Trew assumes that because of the passive-agent deletion in the first clause, the

'police' as 'agents' of the shooting needs to be inferred from the second clause; however, this would be a weak mental representation and so agency is mystified. On the basis of CDA's internalist postulates of mental representation, a better version for Trew would presumably be:

Rhodesian Police shot dead 11 Africans and wounded 15 when they opened fire on a rioting crowd of about 2,000.

But from a cognitive linguistic point of view, mental representation of both sentences would include the *same* 'neurophysiological understandings' related to gestalt formation, motor-interaction etc regardless of the different semantic-syntactic structures. In other words, the CDA perspective does not take into account ontologically subjective properties and so in this respect acts inadvertently in an objectivist manner. This extra neurophysiologically related phenomena in mental representation, similar for both sentences, thus diminishes the significance that CDA attach to sentential structure.

Having shown the symbiosis of the basic level of categorisation with prototypes, I now go on to consider another type of symbiosis: the cognitive interdependence of basic-level nouns and basic-level action categories.

## **5.5 Cognitive Interdependence of Basic-Level Nouns and Basic-Level Action Categories**

### *5.5.1 Basic-Level Action Categories*

Aside from basic-level categories which are associated with physical objects, Lakoff (1987a:



270-1) also argues that actions and properties can be basic-level:

We have basic-level concepts not only for objects but for actions and properties as well. Actions like *running*, *walking*, *eating*, *drinking*, etc are basic-level, whereas *moving* and *ingesting* are superordinate, while kinds of *walking* and *drinking*, say, *ambling* and *slurping*, are subordinate. Similarly, *tall*, *short*, *hard*, *soft*, *heavy*, *light*, *hot*, *cold*, etc are basic-level properties, as are the basic neurophysiologically determined colors: black, white, red, green, blue, and yellow.

A few words of caution on the prospect of basic-level *actions*. Clearly, to be *directly meaningful*, to produce a sharply defined gestalt, ‘drinking’ is more likely to be treated as being basic-level in the presence of basic-level objects such as ‘cat’, for example, in ‘the cat drinks’. Conversely, it is easier to decide on the status of objects as being basic-level, since we motor-interact with objects much more so than with ‘actions’. Now, when attribute lists of respondents for basic-level action categories like ‘eat’ or ‘drink’ are compared with related object categories such as ‘bread’ and ‘soup’, there is considerable overlap between the two. As Ungerer and Schmid (1996: 104) state:

The names of some basic level food categories will be found in the attribute list of EAT, and conversely, the names of basic-level action categories like EAT will certainly rank among the more important of the attributes of the basic level food categories.

The conclusion that Ungerer and Schmid (1996: 104) draw from this is that there is a strong *cognitive interdependence* between action and object basic level categories.<sup>6</sup>

### 5.5.2 Basic-Level Event Categories

The cognitive interdependence of basic-level action and basic-level object categories can also be seen in basic-level *event* categories, e.g. the category of breakfast where there is a

fusing of objects and activities (Ungerer and Schmid, 1996: 104). Breakfast is a complex basic-level category since it comprises basic-level object categories (spoon, cup, table etc) and basic level action categories (e.g. eat, drink, cut). Ungerer and Schmid (1996: 105) cite the experiments of Rifkin (1985) as supporting the view that event categories such as 'breakfast' are basic-level. Using an attribute listing test, Rifkin (1985) demonstrated that subjects could supply large numbers of attributes for basic level event categories such as 'breakfast', 'lunch', 'dinner', 'seeing a movie', 'taking a shower', 'murder', 'rape' and that such attributes did not just include object categories but action categories. For instance, subjects included for the (basic-level event) category of 'murder' both 'gun' and 'kill'. When subjects were asked to supply attributes for related superordinates ('meal', 'entertainment', 'hygienic activity', and 'crime'), their number was much smaller. For subordinates such as 'quick breakfast', subjects did not provide significantly more attributes than basic-level categories. Here are Ungerer and Schmid (1996: 105) on basic-level events:

Checking our examples of basic-level events against the other two criteria of basic level categories, prototype structure and gestalt perception, we have no difficulty in imagining more or less typical instances of breakfast or murders, and in categorizing these events as holistic gestalts.

### *5.5.3 Implications for CDA*

The cognitive interdependence of basic-level noun and verb categories have again consequences for CDA's emphasis that actions should be described by verbs only; for instance, Martin's (1989: 43) claim (5.4.3) that 'seal hunt' is mystifying because its nominal form is incongruent with action needs rethinking. 'Seal hunt' would constitute a basic-level event category where *action* and object categories are fused and so, from a cognitive linguistic perspective, its nominal compound status is not mystifying of the action that transpires. In other words, while 'seal hunt' is a nominal, cognition of 'seal hunt' goes

*beyond* its nominal status. Seeing ‘seal hunt’ through the perspective of cognitive linguistics again demonstrates CDA’s tendency to read off meaning from a metalinguistic description.

In chapter 4, I problematised the symbolic postulates operating in Simpson (1993: 170-1) [which develops Clark (1992: 215); see 3.6.2] from a *connectionist* point of view. I will now do the same from a *cognitive linguistic* point of view. Recall from 3.6.2 two sentences from the Sun text that Clark (1992) analyses:

- 1) ‘Two of Steed’s rape victims - aged 20 and 19 - had a screwdriver held at their throats as they were forced to submit.
- 2) His third victim, a 39 year old mother of three, was attacked at gunpoint after Steed forced her car off the M4.

Now also recall from the same section Simpson’s (1993: 169-171) extension of Clark’s (1992) analysis:

We see a wilful refusal to ‘tell it like it is’. What, for instance, is so difficult about presenting the details of the story in the following way, where the relationship between attacker and victim is not mystified:

- 1) Steed held a screwdriver at the throats of two of his victims as he forced them to submit.
- 2) Steed attacked at gunpoint his third victim, a 39-year-old mother of three, after he had forced her car off the M4.

The basis of Simpson’s ‘non-mystifying’ sentences is that the ‘agent-process-patient’ structure more readily reflects reality. But what is interesting about both of the ‘improved sentences’ is the absence of the category ‘rape’ which was present in the original Sun text (i.e., ‘two of Steed’s rape victims’). ‘Rape’ can be considered a *basic-level event* category (Rifkin, 1985) and so consists of a fusion of basic-level *action* and object categories, i.e.,

cognition of the noun ‘rape’ goes *beyond* its nominal status. So, in the original version, its ‘nominalisation’ status would not be mystifying of the activity taking place. Now, in sentence 2, without the basic-level event category ‘rape’, the action category ‘attack’ is not sufficiently constrained to indicate that the attack was sexual. By stressing the capacity for a sentence’s semantic-syntactic structure to ‘reflect’ reality (in logical empiricist vein) but not dealing with what the *cognitive output* of the sentence is likely to be, Simpson’s sentence 2 is ironically mystifying of what took place.

Having examined the basic level of categorisation in some detail, let me now give some attention to the *superordinate* level of categorisation.

## 5.6 Superordinates

### 5.6.1 Superordinate Noun Categories

How do superordinate categories compare to basic-level ones? Well, as I said earlier (5.3.3), there is no common shape and correspondingly no shared gestalt underpinning hyponyms of the superordinate. Rather, as Ungerer and Schmid (1996: 74) state, the gestalt properties of the superordinate are ‘borrowed’ from basic-level hyponyms. This ‘borrowing’, they refer to as *parasitic categorisation*. What this ‘borrowing’ also does is to highlight salient attributes of the basic-level hyponyms. So for example, in calling a car a vehicle, what is automatically stressed is the function of ‘moving persons or things around’ (Ungerer and Schmid, 1996: 78).<sup>7</sup> This ‘borrowing’ of attributes from the basic level means, again, that the basic level is the conceptually prominent level. Because of the phenomenon of parasitic categorisation, a crucial difference emerges between basic-level categories and

superordinates. Parasitic categorisation and the *lack* of a common gestalt means that hyponyms of the superordinate are only related via what Wittgenstein (1953: 32) referred to as *family resemblances* (Ungerer and Schmid, 1996: 98), Wittgenstein's example being 'game'. Wittgenstein meant by this that 'games' form a family' but that there was no necessary feature for all games. Rather different games are related through criss-crossing similarities between their different *aspects*.

Now, because basic-level categories, by definition, have a *common gestalt*, hyponyms of the basic-level category 'chair', for example, can be linked together via *prototype* structuring, i.e., there is a common gestalt for 'wooden-hard-backed chair', 'beanbag', 'bench', 'arm-chair', 'dentist's chair', 'electric chair' etc, where the prototype would be equivalent to 'wooden-hard-backed chair'. For the superordinate 'furniture', [or Wittgenstein's 'game'] with no common gestalt, it resists the formation of a prototype, there being one or a few category-wide attributes which are salient. That basic-level categories manifest *prototypicality* and non-basic-level categories may not is often unnoticed and misunderstood. For example, Hampton (1981) argues [cited in Harley (1995: 195) under the headline 'Problems with the prototype model'] that abstract concepts resist the formation of prototypes, e.g. it being difficult to talk meaningfully of a prototypical 'truth'!

### 5.6.2 Superordinate Verb Categories

For Lakoff (1987a: 270-1) and for Ungerer and Schmid (1996: 102), nouns are not the only type of superordinate category. Verbs such as 'cause' and 'become' can be seen as superordinate verbs because:

...their main function is to highlight one very general attribute which is part of a whole range of basic-level

action categories...Other candidates for superordinate action categories with a salient general attribute are HAPPEN, BECOME, BEGIN and STOP.

However, it is often the case that the superordinate status of a verb is less stable than that of a superordinate noun. For regarding a verb as superordinate, the following admonitions of Ungerer and Schmid's (1996: 104) are worth bearing in mind:

Of course, it is possible to assemble a hierarchy of actions, by arranging categories like STRIDE, WALK, MOVE, or MUNCH, EAT, CONSUME on the subordinate, basic, and superordinate levels respectively, but these hierarchies will more likely than not be scientific constructs and will not necessarily reflect the cognitive framework of the ordinary language user. In addition, these action hierarchies seem to be even more patchy than their counterparts in the domain of objects and organisms.

So we leave action categories with a feeling that they include a number of basic activities which are probably perceived in terms of prototype categories, but that the analysis becomes less conclusive as we turn to superordinates and subordinates and, more generally, to lexical hierarchies of action categories.

### *5.6.3 Translating Superordinates into the Basic Level*

Consider now the following well-known (inconsiderate) text from Bransford, J. and Johnson, M. (1973: 400):

The procedure is actually quite simple. First you arrange things into two different groups. Of course, one pile may be sufficient depending on how much there is to do. If you have to go somewhere else due to lack of facilities, that is the next step; otherwise you are pretty well set. It is important not to overdo things. That is, it is better to do fewer things at once than too many. In the short run this might not seem important, but complications can easily arise. A mistake can be expensive as well. At first the whole procedure will seem complicated. Soon, however, it will just become just another facet of life. It is difficult to foresee any end to the necessity for this task in the immediate future, but then one never can tell. After the procedure is completed, one arranges the material into different groups again. Then they can be put into their appropriate places. Eventually they will be used once more, and the whole cycle will then have to be repeated. However, that is part of life.

The text is initially difficult for many to understand. One way of understanding the comprehension difficulty is because of the excess of superordinate / abstract categories, e.g.

‘things’, ‘groups’, ‘complications’. In referring to the above text, Rosch (1978: 45) makes the point that:

...what Bransford and Johnson call context cues are actually names of basic-level events (e.g. washing clothes) and that one function of hearing the event name is to enable the reader to translate the superordinate terms into basic-level objects and actions. Such a translation appears to be a necessary aspect of our ability to match linguistic descriptions to world knowledge in a way that produces the ‘click of comprehension’.

When the basic-level category ‘washing clothes’ is introduced, the superordinates can be ‘translated’, making the text easier to understand; (see 7.3.1 for a more detailed commentary on the above text and how cognitive linguistic explanation relates to inference generation and shallow processing).

#### *5.6.4 Implications for CDA*

Repercussions for CDA? In the light of the above, consider the following from Fairclough (1995a: 112):

A lot of nominalizations in a text...make it very abstract and distant from concrete events and situations (Kress and Hodge 1979).

Let us look again at the Bransford and Johnson text from Fairclough’s perspective. So, for example, ‘the procedure is actually quite simple’ contains the ‘nominalisation’ ‘procedure’ and indeed this ‘nominalisation’ appears on two other occasions. ‘Denominalising’ ‘procedure’ to its verb form, we might arrive at the *second person* ‘how you proceed is quite simple...’, and something similar can be done for the other two instances. But is this any easier to understand just because a participant has been included and the verb ‘proceed’ is employed? Even though now there is less interpersonal distance, ideational distance still

exists not because of the *nominalisation* ‘procedure’ but because of the *type* of nominals, i.e., superordinate categories, which remain superordinate in *verb* form. The *basic-level* event category ‘washing clothes’ consists of a fusion of basic-level object categories and basic-level action categories which are cognitively interdependent. Given this cognitive interdependence, ‘washing clothes’ is able to lead to a rich gestalt. Conversely, since superordinates are *not* cognitively interdependent with one another, a rich gestalt is not forthcoming which explains why the text is difficult to process. Again, a corollary of all this is that it is easy to see how narrowly oriented CDA is in its *syntactic* focus away from larger *cognitive* concerns.

Having discussed the phenomenon of basic-level categorisation, let me now consider how cognitive linguists treat compound nouns together with implications for CDA.

### 5.7 Cognitive Linguistics and Compounds: Implications for CDA

Consider the following from Lakoff (1987a: 147):

It is often the case that meanings of compounds are not compositional; that is, the meaning of the whole cannot be predicted from the meanings of the parts and the way they are put together. The parts do play a role in the meaning of the whole expression - they *motivate* that meaning...

Lakoff (1987a: 144) gives a series of examples of which ‘red hair’ exemplifies the point perfectly. Because ‘red hair’ is not focal red, we cannot treat ‘red hair’ compositionally in terms of the intersection of a set of red things and the set of hairs. If then we cannot treat compounds compositionally, i.e. treating components of the compound atomically, then we are unlikely to see compounds strictly speaking in terms of modifier and head. Ungerer and



Schmid (1996: 93) support this with an analysis of the category ‘wheelchair’ arguing that:

...the cognitive categories underlying compound terms like WHEELCHAIR do not only rely on the two categories suggested by the linguistic form, but draw on a large number of other cognitive categories.’

They highlight the other ‘cognitive categories’ as ‘invalid’, ‘hospital’, ‘engine’. Non-objective interactional and experiential properties such as a patient’s motor-interaction with a wheelchair or a nurse’s pushing of a wheelchair also go beyond the surface compound form. Similar to Lakoff (1987a), the cognitive linguist, Langacker (1987a: ch. 12) indicates that a ‘building block’ or compositional approach to compounds is inadequate. Instead, he prefers to think of the components of a compound in what he terms the *scaffolding* metaphor, i.e., disposable when no longer needed. Seeing compounds in terms of the scaffolding metaphor rather than building block metaphor places a greater emphasis on how cognition output does not necessarily mirror linguistic input.<sup>8</sup> To sum up at this stage, here are Ungerer and Schmid (1996: 95):

(1) The standard view which posits a basic head item and a strictly specifying modifier element is far too rigid. With many compounds, even with model cases like *apple juice*, the modifier category supplies more than just the specifying attribute; these additional attributes may not all be ‘objective’ properties, but are often associative and ‘experiential’.

(2) The basic item, i.e. the dominant source category, is not necessarily expressed by the second element of a compound. Depending on the salience of the categories involved, the cognitive category corresponding to the first element may be equally important (as illustrated by *raincoat*) or even dominant (our examples were *coat collar*, *shoelace* and *washing machine*).

So we must be careful in analysing compounds in terms of a modifier and a head of a NP.

Now, recall from 5.4.3 Martin’s (1989: 43) point about how it is *incongruous*, in descriptions of action, to put ‘the action into the modifier of an abstract noun’: e.g. ‘sealing

operation, killing techniques, slaughtering operation, killing methods'. The assumption here is that because 'killing' and 'slaughtering' are modifiers, i.e. subservient to the head in an NP, that somehow this mystifies the 'unsavoury part of the seal hunt'. Patently, from what we have seen of the cognitive linguistic perspective on compounds, this is not the case. Again, my analysis here supports my analysis of Martin (1989) in 5.4.3 and 5.5.3 and thus problematises the imparting of semantic meaning to a grammatical metalanguage and making tacit internalist claims of a relationship between syntactic structure and mental representation.

## 5.8 Endpoints

From the above, it is apparent that the basic level is a privileged level of categorisation since i) we derive meaning from basic-level categories directly (Lakoff, 1987a: 279) and ii) basic-level categories are characterised by *cognitive economy* - a high level of information yield with the least amount of cognitive effort. For Lakoff (1987a: 271):

...it is basic-level physical experience that I believe will ultimately provide much of the basis for an experientialist view of epistemology that supersedes objectivism without giving up on realism.

The basic level is epistemologically privileged for Lakoff not because it provides a *mirror of nature* in the sense that *logical empiricist simples* (see chapter 3) were thought to be epistemologically privileged but because of the basic level's association with motor-interaction, image generation, gestalt perception etc, which assists the reader's or listener's understanding. In chapter 3, we saw CDA's logical empiricist fixation with having sentential structure 'reflect' the 'structure of the event'. In this chapter, we have seen how

this does not take into account the larger cognitive issue of how basic-level cues assist understanding through being associated with motor-interaction, common gestalt etc. Moreover, this diminishes the significance in CDA of supposing that certain sentential structure can lead to mystification. Similar to what I said about connectionism in chapter 4, I have not offered cognitive linguistics as a ‘linguistic panacea’. I do believe that the linking in cognitive linguistics of language with our neurophysiological capacities is a positive development from modular theories of language. However, in chapter 7, we shall see that not everything about cognitive linguistics is a positive development when I show how the Lakoff and Johnson (1980) view of metaphor conflicts with connectionism.

In 4.5.1, I cited Clark’s (1996) outline of three positions on mental representation. The two positions which are polarised are ‘gross descriptivism’ (connectionism), and ‘gross internalism’ (symbolicism). We saw that CDA in its highlighting of ‘transactives’ as a privileged representation chimes with ‘gross internalism’. The midway position, Clark (1996) phrased as ‘modest internalism’ and I quote this again:

*Modest Internalism:* the common-sense constructs serve to pick out transient and / or large-scale features of internal (e.g. neural or computational organization). Examples might include the identification of concepts with distributed, context-dependent patterns of neural activity (see Clark 1993) or the identification of mental images with temporarily time-locked activity in multiple neural regions (see Damasio, 1994). In such cases the folk items (images, concepts) do not have neat, highly manipulable and / or spatially localizable inner analogues. But there remain fairly robust patterns of widespread neural / computational activity which the folk discourse at times succeeds in tracking.

The cognitive linguistics emphasis on the basic level of categorisation would seem to place it in the modest internalism camp in terms of its place within Clark’s mental representation scheme. That is, in cognitive linguistics there is a rejection of gross internalism. But at the

same time, representation in cognitive linguistics cannot be regarded as simply gross descriptivism since there is a strong connection between one type of category - the basic level - and neural activity, e.g. capacity for image generation, motor-interaction etc.

## Notes

1. But from the perspective of *discourse*, it would be strange to regard 'the pope' as a non-prototypical bachelor in the same way that a penguin is a non-prototypical bird. Indeed, Lakoff's semantic analysis neglects the potential discourse meaning of 'bachelor' which often indicates much about the speaker's or writer's attitude.

2. Let me provide another situation where prototype effects are generated (this time a 'mind-reading trick') with an explanation in terms of ICM theory. A fuller appreciation of this may be derived from actually following the 'mind-reading instructions' below. Choose a number between 1 and 10. Multiply this number by 9. If you have a 1-digit number, leave it alone. If you have a 2-digit number, then add together the individual digits. From this sum, subtract 5. On the series, if A=1, B=2, etc, the number you have corresponds to a letter of the alphabet. Now, quickly, think of a country beginning with that letter. Locate the next letter along in the alphabet. Quickly, think of an animal beginning with that letter. Now think of a colour you would associate with that animal.

[see 'note 2 continued' below]

3. In the same way that a mouse can be a small piece of rubber on a laptop, a 'tricycle' can also be a 'bicycle'. Of course neither of these 'equivalences' can transpire on the basis of a necessary and sufficient set of conditions describing ontologically objective properties. It is, for example, a necessarily ontologically objective or inherent property that a 'bicycle' has two wheels. However, since ontologically subjective properties are habitually taken into account in everyday classification, a 'tricycle' may qualify as a 'non-prototypical bicycle' due to similar motor-interactions etc. Consider also the following lateral thinking puzzle:

A man has some wood. On Monday, he shapes it into a cube. On Tuesday he shapes it into a sphere and on Wednesday into a pyramid. He does not touch it nor uses an instrument of cutting etc. How does he do it?  
[See 'note 3 continued' below].

4. One salient method of classification in the west is found in *Roget's Thesaurus*. In Roget, knowledge is divided into six classes, [Abstract Relations; Space; Matter; Intellect; the exercise of the mind; Volition; the exercise of the will; Emotion, Religion and Morality], which are then systematically sub-divided into 'sections', further sub-divided into 'heads' and so on. But what is interesting about the cultural and biological salience

at the basic level is that it is *not* captured by Roget's classification. Roget's taxonomy can thus be construed as *objectivist* since it seeks independence from both cultural and cognitive salience, from how we interact with the world. In other words, independence from the category employer.

5. In relation to the verb 'kill' and Lakoff's comments, consider Orwell (1969: 146) on what he terms 'operators or verbal false limbs':

'These save the trouble of picking out appropriate verbs and nouns, and at the same time pad each sentence with extra syllables which give it an appearance of symmetry. Characteristic phrases are: *render inoperative, militate against, make contact with, be subjected to, give rise to, give grounds for, have the effect of, play a leading part (role) in, make itself felt, take effect, exhibit a tendency to, serve the purpose of, etc., etc.* The keynote is the elimination of simple verbs. Instead of being a single word, such as *break, stop, spoil, mend, kill*, a verb becomes a *phrase*, made up of a noun or adjective tacked on to some **general-purposes** verb such as *prove, serve, form, play, render.*' [my bold]

What Orwell is hinting at here with 'kill', 'break' etc is that they are more cognitively rich categories because they are not like the 'general purpose' or superordinate 'render inoperative'. Orwell, however, does not distinguish between morphological simplicity and the cognitive richness of a category. 'Stop', for example, may be morphologically simple but in cognitive linguistic terms it is a superordinate since it only highlights one very general attribute of a range of basic-level action categories (see 5.6.2 and Ungerer and Schmid, 1996: 102). So, 'stopping the car' does not produce as rich a gestalt as, say, 'parking the car'. 'Stopping the car', in this instance, is an ad hoc superordinate of 'parking the car'.

6. Gumenik (1979), cited in Garnham (1985: 164), investigated the cueing process for sentence recall. He was able to show, for example, how 'architect' is as good a cue for the predicate 'planned the house' as the sentence 'the man planned the house'. This demonstrates that there is cognitive interdependence between the *noun* 'architect' and the *verb* phrase 'planned the house'. Gumenik also showed that 'arctic' as well as 'Eskimos' is a better cue than 'group' for 'the group built their houses out of ice', which indicates that adjectives can be cognitively interdependent with nouns and verbs also.

7. The capacity of superordinates to highlight a function and thus downplay others may also serve a pragmatic deictic role in the same way that use of distal rather proximal demonstrative pronouns expresses emotional displacement e.g. 'Get *that* dog out of here' (see: Wales (1989: 112)). Consider the following. On being stopped by the police while driving, I was directed as follows: 'Could you step out of the vehicle, sir'. The use of 'vehicle' is marked since under normal circumstances the basic-level category of 'car' is the more likely. In this situation, the superordinate 'vehicle' highlights the function of transport but also necessarily *downplays* the motor-interactional properties associated with the basic-level 'car'. The effect, immediately, is one of 'dislocation' from the car. Since the police are implicitly not recognising how a human being interacts with the car, more formal relations are being instituted.

8. That ultimately the syntactic structure of a compound *depends on* its discourse meaning rather the other way around is often exemplified in quips, e.g., a mother taking her problem boy to the ‘child psychologist’ only to find out the psychologist is the same age as her son. It is only as a result of the inversion of expectation that ‘child’ is treated as the modifier of the NP head ‘psychologist’.

*note 2 continued:*

And the answer is: *Denmark, Elephant and Grey.*

On Lakoff’s scheme, all the countries beginning with ‘D’ can be defined in a classical way via a condition which is both necessary and sufficient. However, many if not most European dwellers will possess an ICM where certain countries are more prominent. For this reason, Denmark will be more salient than Diego Garcia, for instance. It is because of this particular ICM that prototype effects can emerge, even though the category satisfies classical criteria, i.e. Denmark belongs to the set of all countries beginning with the letter ‘D’. The corollary of all this is that the prototype model is not as simplistic as it often treated. Some categories such as colour categories *are* scalar, where degrees of membership vary. Other categories such as countries beginning with the letter ‘D’ have clear parameters, but within these parameters, graded prototype effects can arise.

*note 3 continued:*

The man is able to do the above because the wood is actually *sawdust* which he pours in to moulds in the shapes of cube, sphere and pyramid. The problem is difficult to solve because sawdust is non-prototypical wood from the point of view of motor-interaction and to many will be seen as having no function at all.

## Section C: Creating an Alternative Framework for the Analysis of Mystifying Discourse

In *section A*, I highlighted how CDA operated upon symbolic assumptions of mental representation. In *section B*, I outlined how cognitive linguistics and connectionism both problematise these assumptions, and in turn what CDA treat as being mystifying text. The derivational theory of complexity, on which some CDA strategies seem to be based, was discredited in psycholinguistic experiments regardless of the problematising effects of connectionism and cognitive linguistics. In *section C*, the focus is primarily on *inference generation* and how this relates to mystifying discourse. As yet, the understanding of inference generation in CDA is either neglected or superficially treated. However, an understanding of inference generation is crucial to any account of reading as Sanford (1990: 515) asserts:

The ubiquity of inferences in text comprehension makes the study of text comprehension look like a subset of the study of inference making.

In chapter 8, I will relate the issue of inference generation directly to a discourse analytical framework for highlighting how certain news text can lead to mystification in reading of the events being reported. The framework will enable the analysis of the mystifying discourse produced by a non-analytical reader who has little vested interest in the chosen news texts and is largely unfamiliar with their subject matter. The framework is derived from chapter 7 which highlights compatible elements from connectionism, cognitive linguistics and psycholinguistic evidence for shallow processing with particular regard to inference generation. The framework has, then, a *non-symbolic* basis. So the text the framework highlights as leading to mystification in reading of the events reported would not necessarily

be highlighted by CDA. Chapter 6 surveys recent psycholinguistic evidence for shallow processing, particularly with regard to inference generation, so that it can be drawn upon in chapter 7 and form an important part of my framework in chapter 8.



## CHAPTER 6: PSYCHOLINGUISTIC EVIDENCE FOR SHALLOW PROCESSING WITH PARTICULAR REGARD TO INFERENCE GENERATION

### 6.1 Introduction

This chapter is concerned with recent psycholinguistic evidence for shallow processing with particular regard to inference generation in reading. The notion of inference generation is present in the analysis of text by certain practitioners of CDA as we saw in 1.4. However, in CDA, there is little appreciation of the complexities and typology of text inference theory, or of recent psycholinguistic experimental data which has informed text processing theory. There is also little appreciation in CDA of how certain inferences are more likely to be generated for particular readers than others. Because of the constraints of the thesis, and what I try to achieve in chapter 8, I confine myself principally to four types of inference: causal antecedence, causal consequence, instrument and instantiations, i.e., inferences that are referred to in my framework in chapter 8. In the final section of this chapter, I shall show how the psycholinguistic research on inference generation in this chapter conflicts with assumptions of inference generation in CDA. Moreover, I show how this psycholinguistic evidence for inference generation, as well other evidence of text processing, conflicts with the *symbolically* underpinned strategies in CDA for highlighting mystifying text.

The experimental evidence I outline here is of course *experimental* and suffers from the standard problems of all psycholinguistic data generated under laboratory conditions. For example, *actual* skimming and scanning strategies in reading are difficult to replicate; artificially constructed texts may be lacking an interpersonal function found in most everyday texts (e.g. persuasion); the texts are of limited size and may even consist of single

sentences. Reading of course is a vastly complicated phenomenon and psycholinguistic experiments naturally are limited to dealing with a selection of variables. However, the evidence I marshal below possesses validity in the sense that it is largely *consensus* evidence in current psycholinguistics and is unobscure, readily available in standard psycholinguistic and cognitive psychology reference books (Gernsbacher, 1994; Eysenck and Keane, 1995; Harley, 1995). Because it is, as I say, largely consensus evidence, it has a consistency which compares favourably with the rather inconsistent psycholinguistic assumptions of processing in CDA which were outlined in chapter 1. Moreover, because CDA draws upon sociological theory to the detriment of psychological theory, my use of psycholinguistic evidence provides some sort of balance, facilitating articulation of problems with CD analyses, which are felt intuitively. This has already been demonstrated in one respect when I indicated in 3.2.3 how the derivational theory of complexity, a model of processing analogous to that used implicitly in CDA, is countered by experimental evidence. By extension, then, this experimental evidence was *also* in conflict with CDA's 'DTC model'.

## 6.2 Inferences in Text Comprehension and Likelihood of Their Generation

### 6.2.1 *Coherence vs Elaborative*

A common distinction drawn in psycholinguistics is one between *coherence* inferences and *elaborative* inferences. Consider the following from Potts et al. (1988: 405) [see also Sanford (1990: 516-7)]:

- i) No longer able to control his anger, the husband threw the delicate porcelain vase against the wall. It cost him well over one hundred dollars to replace.

The inference that the vase broke is needed to make coherence. Such an inference is

sometimes known as a *necessary* inference. It is also known as a *backward* inference since the reader works anaphorically to make coherence. Now consider a variant on the above:

- ii) No longer able to control his anger, the husband threw the delicate porcelain vase against the wall. He had been feeling angry for weeks, but had refused to seek help.

Here the inference that the vase broke elaborates upon the text and is thus also known as an *elaborative* inference since it is dependent upon information not made explicit in the text and draws therefore much more on encyclopaedic knowledge than the inference in i). Conversely, the inference in i) is not elaborative since it relies less on encyclopaedic knowledge and more on the textual material, making a *connection* between two pieces of text. The inference that the vase broke from ii) is essentially predictive and cannot be tied backwards to any supporting textual material. It is, then, also known as a *forward* inference. As we shall see when we consider *causal* inferences in more detail (6.2.3), elaborative inferences can be backward ones also!

## 6.2.2 Psycholinguistic Evidence Supporting the Likelihood of Generation of Inferences

### *Orientation*

Inferences habitually generated in reading, that psycholinguists have reached a firm consensus about, include referential inferences (e.g. anaphora<sup>1</sup>), case structure role assignment and causal antecedence (see Eysenck and Keane, 1995: 311). These are all coherence inferences, assisting the reader's local and global construction of coherence. As we saw in the previous section, elaborative inferences do not appear necessary to the coherence of a text because they are essentially forward rather than backward. Below, I outline positions which aim to predict when such *non-necessary* or elaborative inferences

might be produced.

### *The Early Constructionist Position*

The *constructionist* position is derived from the work of Bransford, (e.g. Bransford, Barclay and Franks (1972)) and later developed by others (e.g. Johnson-Laird, 1980). Bransford et al. argued that the comprehension process requires that the reader be actively involved so that information not explicitly textually rendered can be 'filled in'. A crucial aspect of the constructionist position is that there will be a rich set of *elaborative* inferences accompanying the generation of coherence inferences. Much of the early research that was marshalled in support of the constructionist position involved the use of memory testing for inference generation. For example, Bransford et al. (1972) gave their subjects sentences such as:

Three turtles rested on a floating log, and a fish swam beneath them.

contending that the inference generated would be that the fish swam under the log. To test this they used a recognition memory test and provided subjects with the sentence:

Three turtles rested on a floating log, and a fish swam beneath it.

most of whom replied they were confident this inference was the original sentence. Bransford et al. (1972) judged that text inferences were stored in memory in the same way as text information. But it soon became clear that on occasion, inferences might not actually be made at the time of comprehension but as a result of prompting during recall. The issue has been examined in detail (e.g. Singer, 1980) with findings suggesting that many inferences associated with recognition memory tests indicate reconstructive work of the

subject during recall. Memory measures can then be only indirect measures of comprehension and may give a distorting picture of the comprehension process. This may lead to over-estimation of the role of inference construction in comprehension. As a consequence, recognition memory tests are used much less in inference research (Eysenck and Keane, 1995: 308).

### *The Minimalist Hypothesis*

Because of the problem with cue-recall experiments, certain theorists have argued that the constructionist position lacks compelling evidence. Principal amongst its detractors are McKoon and Ratcliff (1992: 442) who offer an alternative viewpoint - the *minimalist hypothesis*:

In the absence of specific, goal-directed strategic processes, inferences of only two kinds are constructed: those that establish locally coherent representations of the parts of a text that are processed concurrently and those that rely on information that is quickly and easily available.

The minimalist hypothesis and the constructionist position both agree that coherence inferences are readily generated. Where the minimalist hypothesis and the constructionist position differ most starkly is in the area of which elaborative inferences can be considered to be *automatic*. For constructionists, automatic elaborative inferencing in reading is rife whereas under the minimalist hypothesis, automatic elaborative inferencing is much more restricted. That is, when the reader has no specific goal, automatic elaborative inferences are those which are dependent on information that is quickly and easily available. Indeed, in a series of experiments, McKoon and Ratcliff (1986, 1989a) demonstrated that elaborative inferences were only partially realised because information was *not* quickly and easily available for subjects. They had subjects read short texts containing sentences similar to:

The director and the cameraman were ready to start shooting when suddenly the actress fell from the 14th storey.

This was followed by a *lexical decision task* where target words had to be ascertained very quickly as being present or not. Each test word was succeeded by a signal, and subjects were told to provide a response immediately after registering the signal. Target words such as 'dead', were consistent with the predicted elaborative inference for causal consequence, i.e. *that the actress had died*. McKoon and Ratcliff (1986, 1989a) found that target words which were consistent with a predicted elaborative inference were only weakly identified as being present in the text. This was in line with the predictions of the minimalist hypothesis:

...the inference about death is not necessary for local coherence if the text ends with the sentence about the fall. The event of falling from a 14-storey building is not familiar enough to make the inference easily available. So the minimalist hypothesis predicts that the inference about death will not be included automatically in the mental representation. McKoon and Ratcliff (1992: 457-8)

The advantage of such a *lexical decision task*, where only *one* word cue is employed, is that it reduces the chances of an inference being constructed at the time of the test. This is in contrast to the memory-recognition test of Bransford et al. (1972) where, in the test, subjects were presented with a *whole* sentence. While in essence the above experiment is a kind of memory recognition test, for McKoon and Ratcliff (1992: 458):

The delay between test word and signal was short enough that slow, strategic processes (that might construct inferences at the time of the test) were eliminated.

Hence for McKoon and Ratcliff (1992: 458) this '*speeded* recognition memory test' avoids the problems of inferences constructed at the time of the test. The experiments of McKoon and Ratcliff (1986, 1989a) demonstrate that elaborative inferences are not inferrable when

they rely on information which is not very familiar. The elaborative inference under analysis above is actually a particular kind of elaborative inference - a *causal consequent* inference.

We shall see below that this type of inference is not usually constructed (6.2.4).

As a final point, some of the experiments in McKoon and Ratcliff (1986) were followed up in Potts et al. (1988). Recall from 6.2.1 the texts I provided to show differences between a coherence inference and an elaborative inference. The experimental evidence of Potts et al. (1988), goes even further than of McKoon and Ratcliff's, indicating that the elaborative or predictive inference that the vase broke in the text:

ii) No longer able to control his anger, the husband threw the delicate porcelain vase against the wall. He had been feeling angry for weeks, but had refused to seek help

was not generated on-line at all, not even generated weakly.

### *Automatic vs Strategic Inferences*

I have indicated that the minimalist hypothesis is concerned with predicting whether inferences will be *automatic* or not for a reader with no particular reading goals. Of course, if a reader *is* goal-oriented, being willing, then, to invest more cognitive effort in reading, then some elaborative inferences generated may not be automatic. McKoon and Ratcliff term these non-automatic inferences *strategic* inferences. Examples of strategic inferences are the 'implications' drawn by the journalist in the following news text:

The Guardian Friday April 30 1999 p.13

**HESELTINE SAYS TORIES NOW LIKE LABOUR IN 1980s**

Michael White Political Editor

Michael Heseltine, the former deputy prime minister, last night likened the Conservative party's present state to the Labour plight of the 1980s - *implying that the party would ultimately recover support but casting a shadow over the future of William Hague [strategic inference].*

In an analysis which will have brought little cheer to the embattled Tory leader, Mr Heseltine told BBC1's Question Time that public opinion was 'febrile' and would swing back to the Tories.

'It's not too far from reality to say that the Labour Party in the 1980s did go to very low levels of electability, but they recovered. They recovered when the mood changed against the government of the day. It will turn against the government of this day and when it does the Conservatives will be back in power.'

The analogy, however, carried the *clear implication that Mr Hague is a Neil Kinnock figure, who may reshape his ailing party but will ultimately be jettisoned before it takes power [strategic inference].* Alternatively, as some believe, he may even be his party's Michael Foot, battling to keep factions in check and drifting even further from electability.' [my italics]

The above inferences are elaborative inferences since they elaborate upon the text and so are not necessary for coherence. However, they are less likely to be drawn by a reader who is not goal-oriented, one with little vested interest in the text. In other words, they are unlikely to be *automatic* elaborative inferences. Presumably, the strategic elaborative inferences above are in line with the vested interest of the Guardian journalist to show the weakness of William Hague's position as leader of the Tory party.

#### *Vonk and Noordman (1990)*

I have outlined psycholinguistic evidence to the effect that *elaborative* inferences are *not* automatic when the scenario is unfamiliar to a reader and the reader has no specific goal.

I have also indicated how *elaborative* inferences may be generated *strategically* if the reader has a vested interest in the text and is, thus, willing to invest cognitive effort. I now want to profile reasons why even *coherence inferences* are not necessarily constructed. The experimental work of Vonk and Noordman (1990) presupposes that generation of coherence inferences is to a large extent dependent on the reader's familiarity with the material,



regardless of whether explicit cohesive markers are present. Vonk and Noordman (1990) examined the processing of sentences containing the conjunctions *because* and *but*. They considered sentences such as:

Chlorine compounds are frequently used as propellants *because* they do not react with other substances.

In order to establish coherence for this sentence, the inference that needs to be generated is that good propellants do not react with other substances. Vonk and Noordman wanted to know whether this inference would be produced on-line. Using another sentence:

John is a linguist, *but* he knows a lot about statistics.

Vonk and Noordman were interested in knowing whether the inference necessary for coherence, that linguists do not know much about statistics, would also be generated. They used a combination of on-line reading time and question-answering time analyses. Inferences were presented as statements to be verified *after* reading sentences like the ones above. In some cases, inferences were *already* explicitly stated in the material to be read by subjects, and in other cases not. It was found that subjects answered more quickly in the verification stage when the inferences had been explicitly present in the reading material. The significance of all this? Vonk and Noordman concluded that when the material of the text is unfamiliar even *coherence* inferences are not always constructed. With unusual material, readers have a 'tendency to satisfy themselves with rather shallow processing' (Vonk and Noordman, 1990: 462), especially if the reader has no particular reading purpose.

The reasoning behind this was that readers may satisfy themselves that the text is to some extent *coherent* because *cohesive* markers are present (as in the above examples)<sup>2</sup>. In more detail, here are Vonk and Noordman (1990: 462-3):

The control of inferences depends to a considerable extent on the reader's purpose and the reader's knowledge. Inferences are made on-line if they are related to information that is relevant to the reader's purpose, and inferences are more likely to be made if they deal with familiar topics.

These results suggest that reading is a process in which a balance between costs and benefits is achieved. The benefits consist of the information extracted from the text; the costs are related to the extra mental processes that this requires. **The reader seems to be rather parsimonious in processing.** This is indicated by the absence of the inferences in normal reading expository text. [Experiments 1 and 2]. **If the information is more relevant to the reader's purpose, the benefits are higher and inferences are made on-line [Experiments 3 and 4]. If the inferences are related to information that is familiar to the reader, the costs may be lower, which enables on-line inferences [Experiment 5].** It should be noted that the interpretation of the latter result is not that these inferences are made because they do not require time. They do, in fact. Readers engage in inference processes because the costs are relatively low when the inferences are related to available knowledge... [my bold]

And summing up here are Sanford and Garrod (1994: 705):

Vonk and Noordman suggest that in normal reading, processing may be SHALLOW if the material is unfamiliar and the reader does not have a vested interest in the material.

I have dealt very generally with the likelihood of coherence inferences and elaborative inferences being produced. When I introduced both types of inference in 6.2.1 and 6.2.2, I was actually dealing with a specific kind of *coherence* inference and a specific kind of *elaborative* inference, respectively *causal antecedence* and *causal consequence*. In the next two sections, I want to refer in more detail to the nature of these causal inferences particularly in relation to the likelihood of their generation.

### *6.2.3 Antecedent (Backward) Causal Inferences and the Likelihood of Their Generation*

I follow van den Broek (1994: 561-573), again in Gernsbacher (1994), breaking down causal antecedent inferences into three types:

i) *connecting inference* - the reader establishes a link between the current focal event and prior information thus instituting a **causal antecedent** for the new event.

ii) *reinstatement* - information from prior text which is currently not activated can be reactivated by the reader in order to establish a **causal antecedent**.

iii) *elaborative* - the reader utilises background knowledge to establish a likely but unmentioned **causal antecedent**.

[Notice iii). Although causal antecedent inferences are necessary for coherence, and so are also *backward* inferences, they may also be *elaborative*. In other words, elaborative and forward inferences are not necessarily equivalent descriptions.] A hypothesis of Keenan et al. (1984) was that if readers do institute *causal antecedent inferences*, (i.e., identifying causal relations between adjacent sentences), then the stronger a causal relation in memory, the easier it would be for the reader to recognise it (see also Myers et al., 1987). In Keenan et al. (1984), subjects were shown sentence pairs that Keenan et al. regarded as differing according to the strength of the causal relation in memory. For Keenan et al., causal strength decreased from 1 to 4:

level

- 1      Joey's big brother punched him again and again.  
The next day his body was covered with bruises.
- 2      Racing down the hill, Joey fell off his bike.  
The next day his body was covered with bruises.
- 3      Joey's crazy mother became furiously angry with him.  
The next day his body was covered with bruises.
- 4      Joey went to a neighbor's house to play.  
The next day his body was covered with bruises.

Reading times for the second sentence were recorded and the hypothesis of Keenan et al. (1984) was confirmed since reading times for the second sentence increased from scenarios 1 to 4. Further investigation found that a stronger memory connection was established for the high-causal relation sentence compared to low-causal pairs. Taken together, these

investigations suggest that readers *automatically* make causal antecedent inferences between adjacent sentences when the causal relation is strong in memory. When the causal relation is weak, the processing becomes *shallow*. Looking at the results of Keenan et al. (1984) from the perspective of van den Broek (1994) we can say the following. Readers have little problem instituting *connecting* causal antecedent inferences across sentences. However, the strength or *shallowness* of the *elaborative* causal antecedent inference is dependent on the strength of the causal relation in memory. This is in line with Vonk and Noordman (1990) above - i.e., coherence inferences become *shallower* the less the material leads to accessing of well-known encyclopaedic knowledge. An important corollary follows from all of this. Earlier I said that coherence inferences are necessary inferences. But the evidence of Vonk and Noordman (1990) indicates that coherence inferences are not *always* necessary. While one coherence inference - *connecting causal antecedence* - is automatic, another type of coherence inference - *elaborative causal antecedence* - depends on its automaticity on familiar information being accessed in memory.

#### 6.2.4 Causal Consequent Inferences and the Likelihood of Their Generation

In 6.2.2, we saw that the inference that the actress had died from the sentence:

The director and the cameraman were ready to start shooting when suddenly the actress fell from the 14th storey.

was only weakly identified or, put another way, was produced *shallowly*. This was in line with the predictions of McKoon and Ratcliff's minimalist hypothesis since the event of falling from a 14-storey building is not familiar enough to facilitate the production of the inference and is also not essential for coherence. The above inference is, more specifically

speaking, a *causal consequent* inference. McKoon and Ratcliff's (passim) position that causal consequent inferences are not usually generated if the situation is unfamiliar is in consensus with the position of the *later* constructionists - Graesser, Singer and Trabasso (1994), a constructionist response to McKoon and Racliff's minimalist hypothesis. The following is from Graesser et al. (1994: 382) on why causal consequents are *not* usually constructed:

...because there are too many alternative hypothetical plots that could potentially be forecasted, because most of these alternatives would end up being erroneous when the full story is known **or because it takes a large amount of cognitive resources** to forecast a single hypothetical plot (Graesser and Clark, 1985; Johnson-Laird, 1983; Kintsch, 1988; Potts, Keenan, and Golding, 1988; Reiger, 1975). [my bold]

I have indicated that causal consequent inferences are not normally generated by readers unfamiliar with subject matter. But, it is further the case that causal consequent inferences are very *shallowly* generated at best, for a reader with *little vested interest* in a text, i.e. a reader who is not prepared to invest 'a large amount of cognitive resources' [see also the quote from Sanford and Garrod (1994: 705) in 6.2.2]. Of course, a reader with more vested interest, and thus one more willing to invest cognitive effort, can generate a causal consequent inference *strategically*. But, there are other conditions that make the automatic generation of causal consequent inferences more likely. The next two sections on 'topicalisation' and 'the degree of contextual constraints' deal with this issue.

### *Topicalisation*

Sanford (1990: 519-20) draws attention to how sentences with different attentional foci can affect the likelihood that automatic elaborative inferencing (causal consequence below) can

occur. He considers the following sentence (similar to that used by Potts et al. (1988) - see 6.2.1 above):

i) Unable to control his anger, the husband threw the delicate porcelain vase against the wall

Sanford highlights the topicalising adverbial [reason] clause; the sentence is about the inability of the husband to control his anger and less about the vase and so he conjectures that we are less likely to form the causal consequent inference that the vase broke. By reconfiguring the sentence below, Sanford tries to highlight the breaking of the vase through the downplaying of anger control:

ii) The husband had been unable to control his anger, and he hurled the extremely delicate and very valuable antique porcelain vase at the brick wall.

Throwing the vase becomes a second conjoined sentence, the initial adverbial having been removed. The verb is now *hurled*, suggesting more intensity, and the richness of the modification 'up-plays' the significance of the object (see also: Sanford and Garrod, 1981: 171-185). The inclusion of the material of the wall may also contribute to the possibility that the causal consequence inference will be automatically constructed. Sanford (1990: 520) concludes as to the significance of topicalisation when considering the likelihood of an inference:

...the point is that unless the attentional focus initiated by a sentence is considered, just because an inference is plausible does not mean that it will be made or that it will be fully developed...

Basically, inferential activity should be a function of the structure of the text, of the choice of words, and of the topicalization devices used by the writer. It should not surprise anyone that elaborative inferences can be made; the questions are how the state of an inference relates to focus and how focus relates to language input.

### *Highly Constrained Context*

Both McKoon and Ratcliff (1992: 458) [minimalists] and Graesser et al. (1994: 382) [constructionists] aver that causal consequents can be automatically generated if the context is *highly constrained*, thus making background knowledge more ‘easily available’. As an example of a highly constrained context consider the following from Keefe and McDaniel (1993: 454):

After standing through the three-hour debate, the tired speaker walked over to his chair. He realized that his valiant effort was probably in vain.

Keefe and McDaniel (1993) found in experiments that the causal consequent inference - *the speaker sat down* - was inferred a significant number of times. This can be put down to the fact that subjects were drawing on knowledge of a familiar scenario. But the high degree of semantic associative richness between ‘chair’ and ‘sitting’ also constrains context and contributes to production of the elaborative inference.

In a related experiment, McKoon and Ratcliff (1989a) also indicated that the degree of semantic associative richness present affects the likelihood of automatic elaborative inference generation. For example, in a lexical decision task ‘sew’ was readily recognised after subjects had read the sentence:

The housewife was learning to be a seamstress and needed practice so she got out the skirt she was making and threaded her needle.

suggesting the inference had been made on-line. In comparison with their earlier experiment, McKoon and Ratcliff point out that ‘dead’ is weakly associative with the

‘actress fell from the 14th storey’ sentence, explaining why it is not generated on-line.

To sum up. The possibility of a causal consequent inferences being automatically generated seems to depend upon how constrained the context is and how familiar the scenario is to the reader. A high constraint is created if there is a high degree of semantic association, if the situation is well-known, and there is only one obvious consequence. Topicalisation also needs to be considered as a factor in automatic causal consequence generation.

Let me now move to consider another type of elaborative inference - instantiation.

### *6.2.5 Instantiation and Likelihood of Generation*

Instantiation transpires when a superordinate / general category (e.g. vehicle) is processed more specifically from its context. ‘Vehicle’ in:

The vehicle hovered over the crowd

is most probably processed as a helicopter, balloon etc and is unlikely to be processed as a car, a bus etc (Garnham, 1985: 162). Investigators into instantiation inferences have commonly constructed sentences with general and specific categories:

1. Julie was convinced that spring was near when she saw a cute red-breasted **bird** in her yard
2. Julie was convinced that spring was near when she saw a cute red-breasted **robin** in her yard

Reading time for an ensuing sentence such as ‘the robin pecked the ground’ has been shown to be about the same when it followed sentences like 1 and 2 (Garrod, O’Brien, Morris, and Rayner, 1990, McKoon and Ratcliff, 1989b, O’Brien, Shank, Myers, and Rayner, 1988),



suggesting that 'bird' is automatically instantiable on-line in a sentence like 1. Singer's (1994) synoptic treatment of recent work in 'discourse inference processes' in Gernsbacher's (1994) 'Handbook of Psycholinguistics', a standard reference book, also avers that there is strong support for automatic instantiation of general categories on-line. For McKoon and Ratcliff (1989b: 1143), that general object categories are automatically instantiated on-line is in line with the minimalist hypothesis:

If a specific inference is provided by easily available general knowledge from long-term memory, then it will be constructed even if it is not required for coherence.

In other words, if features of the meanings of words are automatically encoded into memory, then according to the minimalist hypothesis, they must have been easily available during comprehension. As another example, McKoon and Ratcliff (1989b: 1145) found that the instantiation inference 'piano stool / organ stool' was automatically produced by subjects in the following:

Susie pulled out the bench, sat down, and called to everyone in the family to gather around: 'It's Christmas – I'll play and we'll all sing carols'.

In the above, instantiation depended on the readily available encyclopaedic knowledge of Christmas (family gatherings in the home etc), that people sit at pianos, carols are often accompanied by pianos in the home etc:

### *Processing and the Order of General Category / Instance*

Finally, in this section, I want to consider how the order in which an instance and a general category noun are expressed in a text affects processing time. In an investigation of

superordinate category instantiation, Garnham (1981) used a measure known as *self-paced reading time* where:

If subjects are allowed to control the rate at which text is presented to them, the time which they spend reading a particular portion of that text can be used as a measure of on-line processing load.

Garnham (1981: 377)

Consider the following:

1. a) The denim would fade in the sun. b) The cloth was kept in big rolls.
2. b) The cloth was kept in big rolls. a) The denim would fade in the sun.

Garnham found that there was no difficulty in processing in 1 [when b) followed a)] but that, in 2, a) was read much more slowly in following b). Garnham (1981: 383) concluded that subjects take a longer time reading a sentence with an instance noun, if it follows a sentence in which its referent was only described as belonging to a more general category. Garnham does qualify these results by indicating that they only hold if co-text for the first sentence does not hint what type of instance the category member is. So for example, if in 2 above, b) was replaced by 'the cloth was made into jeans', the general category 'cloth' receives some instantiation towards the specific category of 'denim'.

To sum up, the results of Garnham / McKoon and Ratcliff etc indicate that instantiation is for the most part an automatic process when appropriate knowledge is readily accessible, appropriate contextual information is present, and when instances precede superordinates.

As a corollary, text which introduces a superordinate term which is then followed by an instance (and in the absence of appropriate encyclopaedic information) could then be said

to be *inconsiderate* (Sanford and Garrod, 1981: 196-8). I move onto discussion of one more elaborative inference - instrument inferences.

#### *6.2.6 Instrument Inferences and the Likelihood of Their Generation*

Consider the following:

The accountant dried his hands

A reader may infer that the accountant was using a towel to dry his hands. In other words, 'towel' is an *instrument inference*. Early investigations on instrument inferences suggested that instrument inferences were not automatically produced during reading. In the experiments of Singer (1979a; 1979b), verification times for the instrument word were longer with an implied rather than explicit instrument, indicating that 'instruments' are not inferred automatically during reading. Later research suggested that instrument inferences are automatically produced if the context is sufficiently rich or constrains the instrument.

When it is the case that the instrument is necessary for understanding, verification times for instrument words in some findings (Singer, 1980) were similar to verification times for explicitly indicated words. For example, with sentences such as 'The worker drove the nail' and 'The tool was too small for the task', the word 'hammer' was generated in equal proportion when mentioned or not. Other research has also indicated the effects of context on instrument inference production. An instrument has a higher probability of being inferred when it has been referred to already in the text (Lucas et al., 1990) or when context constrains towards a specific instrument. O'Brien et al. (1988) employed passages to determine whether readers automatically made any type of instrument inferences:

- 1) All the mugger wanted was to steal the woman's money.
- 2a) But when she screamed, he **stabbed** her with his **weapon** in an attempt to quiet her.
- 3) He looked to see if anyone had seen him.
- 4) He threw the **knife** into the bushes, took her money, and ran away.

[2b) But when she screamed, he **assaulted** her with his **weapon** in an attempt to quiet her.]

[my bold]

O'Brien et al. (1988) considered reading times for the *last* sentence, 4) when the passage was read by different subjects with either 2a) or 2b). When the passage was read with 2b), the last sentence took longer to read. O'Brien et al. (1988) deduced that this was because the inference that the weapon was a knife was produced only while 4) was being read. In contrast, after finding a quicker reading time for 4) when 2a) was used, they deduced that this was because 'knife' had already been generated *before* 4) was read. Processing of 1) and 2b) was, then, more *shallow* than with 1) and 2a).

Garrod and Sanford (1982) also produced evidence that implied instruments are automatically generated on-line. Sanford (1990: 518-9) attempts to answer the question why the results of Singer (1979a) differ from those of Garrod and Sanford (1982). Sanford (1990) draws attention to the fact that a subsequent experiment by Cotter (1984), using both sets of materials, showed that the Garrod-Sanford and Singer results were replicable within the same study, ruling out spurious explanations. Sanford reasons that the difference appears to be that with the Garrod-Sanford set, the instruments are 'part of the meaning' of the verbs (e.g., *key* is 'part of the meaning' of *unlock*), whereas for Singer's verbs, this is not the case. [A similar set of reasoning is provided in Sanford and Garrod (1994: 703), a review of inference generation in the reference book, Gernsbacher (1994).] 'Part of the meaning' is rather clumsy from a semantic perspective. A better explanation would be to see the

evidence in terms of a degree of semantic association. The instrument 'key' is more readily generated from a scenario with the verb 'unlock' because of the strong semantic association between 'key' and 'unlock'. Strong associations mean knowledge is automatically available so automatic instrument generation on this basis is in line with the minimalist hypothesis.

### *6.2.7 Summing-Up of the Psycholinguistic Research on Inference Generation*

I have drawn attention to the following inferences: causal antecedent, causal consequent, instantiation, and instrument. I have also drawn upon the minimalist hypothesis that states in the absence of specific goal-directed strategic purposes only two types of *automatic* inferences are produced:

- i) local coherence inferences
- ii) elaborative inferences that rely on information that is readily and easily available

When a reader is goal-directed and willing to spend more time and effort in processing, then other elaborative inferences generated are not automatic but *strategic*. There is a psycholinguistic consensus that *connecting* causal antecedent inferences are normally constructed since they are coherence inferences. But the strength of *elaborative* causal antecedent inferences is dependent on the degree of causal relation in memory. A weak causal relation leads to *shallow* production of an elaborative causal antecedent inference.

Elaborative causal antecedent inferences are not always necessary then. Instrument inferences, instantiations and causal consequent inferences are elaborative inferences, falling in to the second category above. Instantiations are automatically generated if they depend upon well-known knowledge. The same applies for instrument inferences, particularly when the scenario is constrained by a high degree of semantic association. There is also a

consensus on the generation of causal consequents that they are not normally constructed or only *shallowly* at best. Part of the reason for this is because readers, without a specific goal, are parsimonious in processing and so refrain from the labour-intensive forecasting of consequents. The generation of causal consequents is much more likely to be *strategic* rather than automatic. The inferences I have referred to will inform my framework in chapter 8 for the analysis of mystifying discourse, providing a more consistent basis than found in CDA for the appreciation of inference generation in text comprehension and how inference generation relates to mystifying discourse. For the sake of clarity, I list in tabular form the inferences that will inform my framework in chapter 8.

#### INFERENCE GENERATION IN TEXT COMPREHENSION RELEVANT TO THIS THESIS

AUTOMATICALLY GENERATED:  NECESSARY FOR LOCAL COHERENCE	AUTOMATICALLY GENERATED:  WHEN INFORMATION IS QUICKLY AND EASILY AVAILABLE	USUALLY HAVE TO BE STRATEGICALLY GENERATED
1. connecting causal antecedents	1. elaborative causal antecedents 2. instantiations 3. instruments  <i>strength of 1, 2 and 3 relies on ready availability of information.</i>	1. causal consequents

I have highlighted the issue of *shallow* generation of causal *consequent* and elaborative causal *antecedent* inferences. The issue of shallow processing is continued into the next section where instead of inferences, I look at how *top-down* processing can lead to shallow processing of lexical items.

### 6.3 Shallow Processing of Text through Top-Down Processing

#### 6.3.1 Sanford and Garrod's (1981 / 1994) Primary Processing Principle

Sanford and Garrod (1981) argue that a central aspect of comprehension is that when it is possible, and as soon as possible, linguistic input will be related to background knowledge, the resulting representation being used to assist the comprehension of subsequent text. This is known as the *primary processing principle*. The primary processing principle, then, has much in common with the minimalist hypothesis in this respect.<sup>3</sup> For Sanford and Garrod (1994: 710), the need to link linguistic input quickly to background knowledge accounts for the fact that it is common for top-down processing to prevail over the processing of syntax, that 'pragmatic mapping can override syntax'. As examples of this phenomenon, Sanford and Garrod offer the common misperceptions of 'Ask not what you can do for your country; ask what your country can do for you' and 'can a man marry his widow's sister?'. Sanford and Garrod (1994: 710-1) conclude:

...the problem seems to be that existing stereotyped knowledge of the situations seems to override sufficient processing detail to enable the anomaly to be detected...processes necessary for true coherence can be incomplete, and that a sense of coherence is achieved on the basis of a good fit between some elements of the sentences concerned and stereotyped knowledge. This is consistent with the primary processing principle of mapping...

*Shallow* processing can occur with the sentences given above because extant stereotyped knowledge of the particular situations for both prevails over the processing necessary to be alerted to the aberration.

Sanford (1990) / Barton and Sanford (1993) make use of a well-known puzzle to illustrate

how top-down processing prevails over lexical registration. The puzzle is as follows:

There was a tourist flight on its way from Vienna to Barcelona. On the last leg of the journey, it developed severe engine trouble over the Pyrenees. The pilot lost control, and it crashed, right on the border. Wreckage was equally strewn in France and Spain, and one question facing the authorities was where the survivors should be buried. What was the solution?

In Barton and Sanford's (1993) experiment, a two-thirds majority (67%) did *not* detect the anomaly that 'survivors should be buried'. This could be explained by invoking the potency of top-down driven expectation, in this case that dead people follow crashes. So since 'survivors' is part of the semantic field of air crash, readers might admit it as a legitimate filler for the patient of 'bury'. To test the argument that the anomaly is *not* detected because of the potency of top-down-driven expectation, Barton and Sanford (1993) conducted a similar 'detection experiment' except this time using the scenario of a bicycle accident instead of an air crash. Detection rates for 'survivors being buried' following bicycle accidents averaged 80% compared with 33% for the air crash. It seemed, then, that top-down processing influences detection rate since one would not usually expect death to accompany a bicycle accident, while one would in a plane crash. In a further study of the *air-plane* crash text, 'survivors' was replaced by 'surviving dead', clearly in this context contradictory. Since this phrase is irregular anyway, if the meaning of 'surviving dead' is fully analysed by subjects when they read the text, then the detection rate should be high. Barton and Sanford (1993) found the opposite, however, only 23 % detecting the anomaly. The result seems to suggest that if one element of the noun phrase, in this case 'dead', is able to fit expectation to an adequate degree, then the noun phrase is in effect only analysed in a shallow manner. A possible objection to this experiment is that it is not so much that readers fail to notice that survivors would not be buried or that they cannot be dead, but that readers edit such phenomena in real discourse, treating them as mistakes; [see, though, the



section on ‘Reder and Kusbit, 1991’ in 6.3.2 below]. However, the strength of the objection still rests on the fact that top-down has precedence over bottom-up processing.

### 6.3.2 ‘Moses Illusion’

#### *Erickson and Mattson (1981)*

Erickson and Mattson’s (1981) investigations into partial processing led to the coinage of the ‘Moses Illusion’. They asked their subjects to respond to:

How many animals of each sort did Moses put on the ark?

Erickson and Mattson found that a large number of subjects processed this in a shallow manner since they did not detect the anomaly. Erickson and Mattson argued that this was because of the closeness between Moses and Noah in memory. From the perspective of Sanford and Garrod’s primary processing principle, this would be accounted for as follows.

Readers make an effort to immediately relate textual information to the familiar (familiar, in this case, to those with some knowledge of the Old Testament); the necessary partiality that ensues will mean that anomalies will often go undetected. Indeed, Erickson and Mattson go on to state that the Moses Illusion suggests that natural language processing involves incompleteness on a more general level. In other words, readers will embrace a particular term providing there is a *high degree of fit* with their expectations in that context, and a full analysis of the meanings of terms usually does not occur.

*Reder and Kusbit (1991)*

Reder and Kusbit performed a series of experiments in order to try and ascertain explanations for the Moses illusion. One explanation might be that readers do notice the anomaly but do not report it, assuming the text is *co-operative* in the sense of not violating Gricean maxims. To test this, Reder and Kusbit asked one group of respondents to spot irregular items in sentences, and another group to ignore anomalies and answer questions as though they were semantically adequate. However, subjects who had been requested to detect anomalies *still* erred, overlooking many of them. Reder and Kusbit noted that the detection tasks were time consuming, suggesting the difficulty of such a process. This in turn suggested the high degree of overriding capacity of top-down processing over bottom-up analysis. Overall, for Reder and Kusbit, these results lent support for the following: subjects who originally did not report the 'Moses Illusion' most probably had *not* noticed the anomaly.

*Topicalisation Leading to Less Shallow Processing*

In another experiment related to the 'Moses Illusion', Bredart and Modolo (1988) showed that error detection in the Moses Illusion depends on whether the crucial name is a syntactic focal point. In contrast to the sentence:

Moses put two of each kind of animal on the Ark.

which Erickson and Mattson had used, Bredart and Modolo (1988) used:

It was Moses who put two of each kind of animal on the Ark.

In the immediate sentence above, the cleft transformation highlighted ‘Moses’ and led to higher detection rates. In other words, the topicalisation of ‘Moses’ leads to less shallow processing.

### 6.3.3 *Flores d’Arcais*

Other evidence that *registration* of syntax is not necessarily prominent in processing comes from Flores d’Arcais (1987). He makes a distinction between *computation* and *use of the results of computation*. On the basis of findings, he argues that while readers make full computation of syntax, the results of this computation are not necessarily used. This is because:

...what makes good readers good is their ability to deal more efficiently with ‘higher order’ kind of strategies, thus relying less on the results of syntactic computation. Flores d’Arcais (1987: 632)

For Flores d’Arcais, readers rely firstly on top-down strategies and use bottom-up strategies when top-down sources are impoverished, etc. Indeed, from the results of one experiment, Flores d’Arcais (1987: 632) suggests that:

...readers who are good at extracting information from rapidly presented texts seem to be less sensitive to syntactic violations than less efficient readers.

This supports the results of experiments we have seen above. On a similar point, evidence cited in Garnham (1985: 138-141) suggests that what people remember about a sentence is not its syntactic structure, nor its semantic meaning verbatim, but its *message* constructed through an interaction between the sentence and the reader’s cognitive resource.<sup>4</sup>

#### 6.3.4 *Summing-Up*

From the evidence discussed in 6.3, top down processing *can* prevail over the registration of syntax and lexical items such that they are processed in a shallow or partial manner. So, we should also be sceptical about regarding the mental representation of lexical items in a sentence as being necessarily *compositional*. In the Barton and Sanford (1993) experiment, while a two-thirds majority (67%) did *not* detect the anomaly ‘burying survivors’ in the air crash scenario, of course 33% did detect this irregularity. We cannot, then, take it as an absolute that top-down expectancy *always* overrides the semantic meaning of a lexical item for every reader. However, to detect the anomaly, the reader has to be able to analyse ‘survivors’ and ‘buried’ compositionally or ‘fully’, something that the majority patently did not do. We saw also that Reder and Kusbit (1991) reported that detection of anomalies, which would necessarily involve compositional *analysis*, was time consuming. To sum up: i) it would appear that it is more usual for the mental representation of a reader to be non-compositional with regard to lexical items in a text and ii) to detect anomalies, the necessary *compositional* analysis is in line with *extra cognitive effort*.

Returning to the Erickson and Mattson (1981) experiment, the evidence is generalised to apply to all readers when of course the processing presupposed a particular kind of reader - i.e. one versed, at least a little, in the Old Testament. The same goes for the Christmas scenario of McKoon and Ratcliff (1989b: 1145) that we saw in 6.2.5. The danger, of course, with generalising the results of such experiments is that inter-individual variation as well as intra-individual variation is downplayed. For example, if Reder and Kusbit (1991) had offered large cash awards for all the anomalies that their subjects were able to spot, these subjects would have had a greater vested interest and so would more likely have engaged with the text with greater cognitive effort. It is a basic assumption, though, of this thesis that

degrees of reading scrutiny and vested interest in texts vary both inter and intra-personally.

As we shall see in the last section of this chapter, I argue that a CD analyst *has* a vested interest in the text they examine. This leads them to derive a different discourse from a reader with no specific reading goal for the same text because the former derives *strategic* inferences (as well as automatic inferences) and the latter only *automatic* inferences.

I want now to deal with the issue of shallow registration of form and structure in the context of *classical (logical)* processing. In doing so, I will outline the likelihood of classical (logical) processing in readers.

## 6.4 Classical Processing and Shallow Processing

### 6.4.1 Semantic Value

Consider the following from Sanford (1990: 522) on the partial processing of the ‘burying the survivors’ text we saw above:

If the processing of inferences is incomplete, in what way is it incomplete? Linguistic elements can fill roles that they do not really fit on the basis of partial matches. What kind of relation can be said to result? In terms of conventional AI, the role could be said to be an **IS-AN-INSTANCE-OF relation**. Yet this may not make much sense. With the airplane scenario, if one asks explicitly, ‘do you bury survivors?’ subjects recognise how silly the problem was. An alternative to the full relation is that the pattern match produces the equivalent of a link, but one that does not have a semantic value (i.e. *survivors* is simply *associated* with the slot for patient). **This is very different from the classical view of inference, of course, which would require the derivation of a predicate with a semantic value.** [my bold]

Evidence from Sanford (1990) and Barton and Sanford (1993) indicates that readers are often shallow with regard to the construction of *semantic value*. Since on the classical

picture, the derivation of a predicate requires ‘whole’ semantic value, the classical theory is directly confronted by Barton and Sanford’s (1993) experimental evidence. For Sanford (1990: 526-7):

...it seems likely that the brain is poor at inference in the classic sense of making full-blown inferences but is well adapted to making simpler mappings between things. In the discourse context, the brain is good for mapping segments of discourse onto existing knowledge but is perhaps less good for generating novel inferential outcomes.

#### *6.4.2 Formal Deductive Inferences*

Sanford (1990) illustrates the above distinction he makes as follows. In a problem where a subject is informed that *Fred is taller than Bill, Bill is shorter than John*, and so on, subjects may be able to recognise this as a transitive inference problem. But unless a subject was expecting to be questioned, there would be little inference construction as to the transitive relations.

I would argue that such a case shows that, for what is only a little more effort in terms of inference operations, there is a large difference in difficulty between having a representation for a transitive inference problem and knowing that a problem belongs to the transitive inference class. Sanford (1990: 527)

Instituting logical relations on Sanford’s account is something that requires extra processing.

This view is supported by a well-known experiment by Wason (1966) which I outline below.

We saw in chapter two that on the classical account, mental computation is the manipulation of symbols according to a structured rule system. If this were the case, then trivial deductive inferences should present no significant problem for humans. However, evidence as to how humans deduce counters this assumption. This is particularly the case with reasoning

involving *material implication* or the *if...then* [ $p \rightarrow q$ ] construction. One of the most famous demonstrations of this is the ‘Wason selection task’ (1966). Subjects were given four cards, each with a number on one side and a letter on the other. They were presented with the rule - ‘**if there is a vowel on one side, then there is an even number on the other**’ - and four cards showing an ‘A’ ( $p$ ), a ‘K’ ( $\sim p$ ), a ‘2’ ( $q$ ) and a ‘7’ ( $\sim q$ ), (where  $p$  is a vowel and  $q$  an even number). The problem: *which of the cards must you turn over to see whether the rule is true or false?* From the classicist position, processing would be consistent with the truth functionality of  $p \rightarrow q$ , i.e:

$p \rightarrow q$	$p \rightarrow q$
T T T	V T E
T F F	V F O
F T T	C T E
F T F	C T O

Most subjects stated that the A and the 2 need to be turned over. But the logical response, the A and the 7, was produced in only 4 % of subject responses. Substituting the symbols V for vowel (T) and C for consonant (F), E for even (T) and O for odd (F), we can see more clearly which cards have to be overturned to disprove the rule. Material implication is true when  $p$  is true and  $q$  is true, i.e. when there is a vowel on one side and an even number on the other. Clearly, A needs to be turned over. This of course confirms the rule, but we need to turn over another card to see if the rule can be disproved. Material implication is false when  $p$  is true and  $q$  is false, that is, when there is a vowel on one side and there turns out to be an odd number on the other. So the 7 needs to be turned over since if there were a vowel on the other side, then this would disprove the rule. Most people do not select the K since the rule says: ‘if a card has a *vowel* on one side’, and since this card does not have a vowel on one side, it makes no difference what is on the other. But why do so many people erroneously choose the 2 when they commit the logical offence of *affirming the*

*consequent*?<sup>5</sup> In other words:  $p \rightarrow q \neq q \rightarrow p$ .

When more realistic and ‘contentful’ materials were used, reasoning scores drastically improved. Wason and Shapiro (1971) employed the following rule: ‘Every time I go to Manchester I travel by car’. When subjects were given four cards indicating travel destinations on one side and modes of transport on the other, and instructions akin to the above, 63% of subjects produced the logical response. It seems, then, that the more sense of the materials subjects can establish, the more likely they are to perform logically. However, this is anomalous from the classical position, since logical inferencing is meant to depend on *form* and not *content*. Wason’s experiments, then, mount a challenge to the classicist view of cognition as autonomous symbol manipulation. Reasoning seems to be very much dependent on assimilating the problem to encyclopaedic knowledge rather than penetrating to the formal structure of the problem sentence. Common-sense reasoning strategies involving induction and default inference are hardly formalisable in logic (from the philosophy of logic: Harman, 1986; from the philosophy of science: Kuhn 1970; from cognitive science: Oaksford and Chater 1991). In a sense then, common-sense reasoning is *shallow* with regard to formal logical structure when the scenario is unfamiliar. As a coda to the above, I am not suggesting that the majority of people have difficulty in processing logically. Rather, this kind of processing will incur a higher than normal degree of cognitive labour if the reader does not have familiar knowledge to draw upon. By the same token, this processing labour may not be forthcoming from a reader who has little interest in a text. Thus, logical processing, and so the detection of logical offences, is facilitated by a high level of interest in the text, as well as familiarity of context. This ties in with the possibility of anomaly detection (Moses Illusion / ‘survivors buried’) since this too can be dependent on a vested interest in the text, in turn in line with extra processing effort.



I now go on to outline evidence for the selectional processing of narrative with regard to main characters and thus the shallow processing of secondary characters.

## 6.5 Selectional Processing of Narratives

### 6.5.1 Main vs Secondary Characters

There have been various studies that have shown how references to *main* and *secondary* characters are treated differently in reading. Anderson, Garrod and Sanford (1983) investigated the differences in processing of 'main' and 'scenario-dependent' characters.

The latter were characters which depend upon a particular scenario. So for example, if John visits the cinema and is shown to his seat by the usherette, then John would constitute a main character, since his action is outside the limits of the cinema whereas the usherette is bound to the scenario of the cinema. Using a self-paced reading approach, Anderson et al. ascertained that sentences containing pronominal anaphoric references to main characters were read more quickly than those with references to scenario-bound characters. This suggests that readers are more 'involved' in processing with main characters than with secondary characters. In a similar experiment, Morrow (1985) explored the processing of a main character, a character which was the theme of the first three sentences of a short narrative. Below is the narrative used by Morrow (1985) where different subjects were presented with the text in its different variations [(A) or (B) or (C)] :

Paul caught the flu and was feeling pretty awful. He told his eldest son Ben to keep the house quiet. He got up from bed to go to the bathroom, irritated by the noise. Traffic was rushing by the house. The kids were arguing in the den.

(A) That noisy Ben was messing up the kitchen.

or

(B) Noisy Ben was tramping around in the kitchen.

or

(C) Ben was wondering when his father would feel better as he ate in the kitchen.

The floor felt cold on his feet.

Whose feet are referred to?

The narrative presents the main character (Paul) and the subsidiary character (Ben). (A) introduces Ben from Paul's attitude, (B) is similar to (A) but presents the subsidiary character as being more active, but (C) allows Ben to have a perspective. So with (A) and (B), the subsidiary character is backgrounded while it is more foregrounded in (C). In response to the probe question, 'Whose feet are referred to?', Morrow found that choice of main character (Paul) as referent was high in (A) (97%) and (B) (84%), but low in (C) (34%). Morrow concluded that when the subsidiary character is backgrounded from the perspective of the main character that the reader's perception of states is bound up with the perspective of the main character. A straightforward objection to this experiment is that the reader's perception of states is generated after the probe question and so we cannot infer that the experiments indicates what transpires normally in the heads of readers. An experiment where this objection does not apply, because it is *on-line reading* which is being gauged by measurement of reading time, is included in the following section.

### 6.5.2 Proper Names vs Role-Descriptions

Garrod and Sanford (1988) performed a reading time investigation into how setting information or atmosphere affects the processing of characters [see also the experimental evidence of Sanford, Moar and Garrod (1988)]. Below is an example of a text used by Garrod and Sanford (1988):

*Lunch at the cafeteria*

Alistair hung up his coat and picked a tray.

The waitress smiled as she poured the coffee.

The atmosphere was hot and sticky. (optional)

He took the cup or *She* offered the cup. (character mention)

He mopped his brow or *She* mopped *her* brow. (target sentences)

The main character is Alistair since he has a proper name, whereas the subsidiary character is 'the waitress' since it is a merely role description. The critical target sentence can refer to either the main or subsidiary character. In order that the referent of the target sentence does not surprise the reader, the same referent is mentioned in the previous sentence. So when the character mention sentence is 'He took the cup', the target sentence is 'He mopped his brow'. The reading times of subjects were ascertained for the target sentence 'he mopped his brow' referring to the main character (Alistair) and the target sentence 'she mopped her brow' referring to the secondary character (waitress). This was done with the atmosphere sentence either present or absent. Consequently, there were 4 sets of reading times:

*Reading Times (ms) for the TARGET SENTENCES*

	Main Character ( <i>Alistair</i> ) Target: He mopped his brow	Secondary Character ( <i>waitress</i> ) Target: She mopped her brow
With <i>Atmosphere</i> Statement	1379	1430
Without <i>Atmosphere</i> Statement	1650	1463

The results show that when the atmosphere sentence is absent, reading times are significantly slower when the main character (Alistair) is referred to in the target sentence but this is not the case when the subsidiary character (waitress) is referred to in the target sentence. Here are Sanford and Garrod (1994: 709):

The results thus fit intuitions that the behaviour of the main character needs causal explanation, not that of the secondary characters. Certainly, the actions of main characters are typically either explained or motivated

directly in stories and narratives, whereas those of secondary characters are not. The present results suggest that this goes hand in hand with an **automatic selective process** which seeks the formation of richer structures around main characters. [my bold]

For Sanford and Garrod (1994), then, the experiment showed that subjects focused more on the main character since they were actively seeking to supply a causal explanation for *why* he mopped his brow in the absence of information that the atmosphere was hot and sticky.

When the target sentence referred to the waitress, the absence of the 'atmosphere sentence' did not affect processing time, indicating that the behaviour of secondary characters do not usually require causal explanation.

The expectation that the actions of main characters should be explained is in line with other experiments cited in this section in the sense that main characters are an *automatic* focus for readers (see Sanford and Garrod, 1994 above). The corollary of this is that inferences surrounding *secondary* characters are normally generated in a more *shallow* manner compared to that of main characters since causal explanations are not normally sought for secondary characters. This is not to say that a reader with a vested interest is not able to seek causal explanations for secondary characters. However, these processes would be *strategic* rather than automatic and would thus require more effort. (See also Stevenson (1993: 119) on the issue of reader-focus on characters in narrative).

In the last section, I compare the psycholinguistic evidence I have produced for shallow processing of text with certain cognitive assumptions in CDA (many of which were outlined in chapter 1), teasing out the implications of this chapter for CDA.

## 6.6 Implications of This Chapter for CDA

### 6.6.1 Orientation

In order to facilitate comparison of psycholinguistic evidence for shallow processing with processing assumptions in CDA, I will use what I call the idealised reader (IR), a reader based on the psycholinguistic evidence outlined above. This is a reader who has little vested interest in a text and is not particularly goal-driven. This reader is, then, one who would not make *strategic* inferences, but only *automatic* inferences. I also assume this reader would be largely unfamiliar with the subject matter of the texts below although I do assume IR to have encyclopaedic knowledge that would be expected of an adult (unless otherwise stated).

Such a reader, then, is in line with the formulation of the minimalist hypothesis but also with the results of experiments where top-down processing overrides bottom-up factors and where there is a lack of detection of logical offences. IR is also in line with the automatic seeking of causal explanations for main characters. Below, I use IR to focus more keenly than CDA on the micro-context of interpretation. I am concerned broadly speaking with the following:

1) In chapters 4 and 5, I showed that what *symbolic / logical empiricist / classical* CDA regards as mystifying text was not from the perspective of *connectionism* and *cognitive linguistics*. Below, I show that what CDA regards as mystifying text is also *not* the case from the perspective of IR - i.e. the psycholinguistic evidence for shallow processing I have amassed in this chapter also conflicts with *symbolic* etc assumptions in CDA. This is the topic of 6.6.2.

2) I make the assumption that a 'resistant' reader (in the parlance of CDA: 1.2.2) makes a

good deal of cognitive effort. In CDA, ‘non-resistant’ readers allow texts to position them into particular interpretations. I assume, then, that ‘non-resistant’ readers for CDA make only minimum cognitive effort. Since IR invests minimum cognitive effort, IR could be construed as a non-resistant (or non-analytical) reader. I show below that CDA explanation (interpretation-2) of how a non-resistant reader is constrained by the macro-context into accepting the subject position of the text is *not* coincident with IR’s discourse derivation from the same text. This is because CD analysis done by proxy for a non-analytical reader includes the generation of *strategic* inferences (in line with a large degree of cognitive effort), when non-analytical IR only generates *automatic* inferences, (in line with only minimum cognitive effort). This is the subject of 6.6.3.

### *6.6.2 How the Idealised Reader Problematizes what CDA Regards as Mystifying Text*

#### *Cause-Effect*

Recall the extracts from Trew (1979), from 1.4.1, and Simpson (1993), from 3.6.2, as to how CDA isolates text which is mystifying of causality. Their focus was sentential. Causality was transparent if the (transactive) AGENT- PROCESS-PATIENT construction was used. Causality that had to be inferred across clauses or sentences was somehow weaker than an AGENT-PROCESS-PATIENT clausal construction. Now, consider again the following from Keenan et al. (1984) that I outlined in 6.2.3:

level

- 1     Joey’s big brother punched him again and again.  
      The next day his body was covered with bruises.
- 2     Racing down the hill, Joey fell off his bike.  
      The next day his body was covered with bruises.
- 3     Joey’s crazy mother became furiously angry with him.  
      The next day his body was covered with bruises.

- 4        Joey went to a neighbor's house to play.  
           The next day his body was covered with bruises.

The form of 'the next day his body was covered with bruises' is similar to that of the passive voice although in this context 'covered' is functioning as a stative adjective<sup>6</sup>. [I treat 'covered' as a stative adjective rather than the past participle of a passive partly because of the time interval in this context; the fact that one cannot cover in real time another body with bruises. In contrast, in a sentence such as 'In the initiation ceremony, the fresher was pelted with tomatoes, covered in gunge, hosed down with water and chained to a lampost for 4 hours' where a succession of actions is highlighted, I understand '[was] covered' in terms of an action and thus an agentless passive and not a state.]

The agentless passive is a feature which CDA has often isolated as being salient since agents can be deleted with passives, with ensuing mystification of agency. Even though 'was covered', in Keenan's et al.'s (1984) texts, describes a state and not an action, because equally the agent of the bruising is absent from the sentence it enables us to say more precisely (as I will show) why use of the passive *can* be 'mystifying' and sometimes not.

The issue relates to the strength of causal relation in memory. In Keenan et al.'s (1984) experiments, the 'agent' was readily inferable *across sentences* in 1) above because the causal relation was strong in memory. That is, not only could a *connecting* causal antecedent inference be instituted (which we saw earlier are usually effortless for readers) but a strong *elaborative* causal antecedent inference could also be inferred since background knowledge was strong here. However, an agent was less sharply inferred in 4) since the causal relation was not so strong. That is, while the *connecting* causal antecedent inference was inferred, the *elaborative* causal antecedent was weak since causal relation in memory was weak. So what do we conclude? By extension, *passives* accompanied by agent deletion

can be *mystifying* (to varying degrees) at least to the identity of causal antecedence if the causal relation in memory is *not so* strong. This possible mystification has, then, nothing to do with the fact that agency has to be inferred *across* clauses or even sentences, *pace* CDA.

*Trew (1979: 98-9)*

Consider again Trew's (1979: 98-9) example, backed up by Lee (1992: 100) and Montgomery (1995: 240), that the identity of the agent of the killing is only weakly implicated in:

Eleven Africans were shot dead and 15 wounded when Rhodesian Police opened fire on a rioting crowd of about 2,000.

Not only is it [The Times report] in the passive, but the syntactic agent is deleted...and is identified only weakly by implication through the temporal conjunction with the police opening fire... Trew (1979: 98-9)

The Times uses passives (*Rioting blacks shot dead by police, Eleven Africans were shot dead*). The effect of the passives is to further attenuate the agentivity of the police... Lee (1992: 100)

...although the police are clearly the agent in an active construction, it is one in which they 'open fire on', a process which is significantly more neutral as to its consequences than 'shooting dead'.

Montgomery (1995: 240)

[Simpson (1993: 170) rests on the same assumption that inferences generated across clauses are weak representations. This was outlined in 3.6.2]. Two points. Firstly, as Keenan et al. (1984) showed, agency is inferable across sentences and so by extension across clauses. Secondly, the causal relation between 'Police' or anybody else 'opening fire' and people being 'shot dead' is, I would assert, a fairly strong causal relation in the memory of many people. Although IR is largely unfamiliar with the subject matter of a text, as I have said,



IR has ‘normal’ encyclopaedic knowledge of an adult. Thus, in the discourse IR derives from the text, the causal antecedent is *not*, then, ‘identified weakly’. A *connecting* causal antecedent inference is not only instituted by IR but a *strong* elaborative causal antecedent inference also. In other words, what Trew regards as being mystifying of agency is not the case from the perspective of IR.

Trew’s analysis, as we saw (3.7.1) (and Simpson’s (1993)), is predicated upon a *symbolic / logical empiricist* over-emphasis on sentential *structure*, (as well as internalist assumptions of mental representation), over consideration of the contribution of the human understander. On this assumption, then, Trew, and the other authors cited, misleadingly ignored the contribution of the human understander, given that such a contribution is supported by psycholinguistic evidence. In this particular case, the contribution of the human understander is their strength of causal relation in *memory* and their capacity to generate inferences across clauses and sentences. Indeed, if as Flores d’Arcais (6.3.3) says:

...what makes good readers good is their ability to deal more efficiently with ‘higher order’ kind of strategies, thus relying less on the results of syntactic computation. Flores d’Arcais (1987: 632)

CDA are misguided in focusing so much on sentential structure. Full use of syntactic computation is not the norm for someone reading quickly for general information, with otherwise little vested interest in the text, since they are relying much more on ‘higher-order’ or top-down strategies.

*Hodge and Kress (1993)*

As another example, consider the following from Hodge and Kress (1993: 48) which I

outlined in 3.2.5:

As time went by, the rain fell on the seeds and the sun shone down on them, and the turnips began to grow.

Hodge and Kress (1993: 50) then compare both transactive and non-transactive descriptions:

non-transactive: The rain fell on the seeds.

transactive: The rain (water) moistened the seeds.

non-transactive: The sun shone down on them.

transactive: The sun warmed the soil.

They argue that the non-transactives used in the story mystify the nature of the causality whereas the transactives do reveal the causality. From the perspective of the psycholinguistic evidence provided here, the richness of the causal antecedent inferences would depend on the strength of causal relation in memory. [I highlighted in 3.2.5 how Hodge and Kress (1993) neglected the issue of how language *cues* encyclopaedic knowledge]. Although IR is an adult, the same basic principles apply to children (though of course children do not have what constitutes normal adult knowledge). A young child, I presume would *not* have rich knowledge of the cause-effect involved in respiration, germination, photosynthesis etc. We can suppose that in discourse they would not make rich backward *elaborative* causal antecedent inferences but merely make *connecting* causal antecedent inferences, i.e., connecting turnip growth to the rain and the sun but not really knowing *why*. Hence, Hodge and Kress's (1993) notion that transactives demystify the causal relations is incorrect. Again, the psycholinguistic evidence conflicts with the *symbolic / logical empiricist* over-emphasis on sentential structure in CDA, as well as internalist mental representation, over considerations of the contribution of the human understander. In this particular case the contribution of the human understander is their

strength of causal relation in *memory*. That is, Hodge and Kress (1993) neglect how mental representation is the *output* of what the sentences *cue*, i.e., weak causal relations in children's memories in the case of germination, respiration, photosynthesis etc. Mental representation necessarily goes *beyond* the semantic-syntactic structure of a sentence and so is not, as implied in Hodge and Kress (1993), a facsimile of semantic-syntactic structure.<sup>7</sup>

*Simpson (1993: 170-1) / Clark (1992)*

In chapter 3, I showed how Clark (1992) and Simpson's (1993) analyses of a news text were underpinned by symbolic postulates. In chapter 4 and 5, I showed how these symbolic postulates were problematised by connectionism and cognitive linguistics respectively and thus what these authors regarded as mystifying text. In the last section, the psycholinguistic evidence for causal antecedence inference generation not only conflicted with Trew (1979: 98-9) but also with Simpson (1993: 170). Below, I show how psycholinguistic evidence - but this time relating to *instantiation* inferences and *selective processing of main characters* - also conflicts with Simpson's (1993), and Clark's (1992) *symbolic* analyses and again what they regard as being mystifying text.

*Selective Processing of Main Characters*

Consider the following from Clark (1992: 215) again:

- 1) Two of Steed's rape victims - aged 20 and 19 - had a screwdriver held at their throats as they were forced to submit.
- 2) His third victim, a 39-year-old mother of three, was attacked at gunpoint after Steed forced her car off the M4.
- 3) Two days later, he gunned down call-girl, Jacqueline Murray, 23, after picking her up in London's Park Lane.

Clark alleges that ‘...the perception of Steed as rapist is reduced by making the sentences passive and deleting him as agent’. That is, the semantic-syntactic encoding diminishes Steed’s responsibility. We saw in 3.6.2 that what underlay Clark / Simpson’s analyses was the symbolic / logical empiricist assumption that a sentence should fully represent events and independently of the contribution of the human understander. In the next paragraph, I take into account the contribution of the human understander with regard to processing of main characters.

In the above, Steed is a proper name and the other referents, apart from Jacqueline Murray, are role descriptors, ‘rape victims’ and ‘mother of three’. If we go back to the psycholinguistic evidence for *name / role descriptor*, we see that Steed is, then, the main character. IR *automatically* seeks causal explanation of Steed’s actions in discourse derivation from the text. Steed, then, is unlikely to be seen as absent from the discourse perspective of IR when sentence 2 is processed, *pace* Clark and Simpson. That is, again, the psycholinguistic evidence from this chapter conflicts with the symbolic / logical empiricist assumptions of CDA.

Now, Simpson (1993: 170) offers reinforcement of Clark’s analysis:

...so obscured is the relationship between attacker and victim that it allows a possible reading wherein **someone else** attacks the woman at gunpoint while Steed only forces her car off the road. [my bold]

Since ‘someone else’ (see bold above) is not given a name, ‘someone else’ would have to be a *secondary* character. But IR, with little vested interest in the text, only *automatically* seeks causal explanation around *main* characters. It follows that the alternative

interpretation that Steed forced the ‘victim’ off the road while ‘someone else’ attacked her is a purely *strategic* elaborative inference, generated by an *analyst* investing time and effort in the text. Not only is there little textual warrant for this ‘alternative reading’, but *non-analytical* IR would not produce such a reading.

One further remark. CDA makes the point that since nominalisations remove participants, events are made more distant. CDA tend, however, to conflate *ideational distance* with *interpersonal distance*. I have no problem with nominalisations producing *interpersonal* distance since formal and so ‘distant’ letters are characterised by a higher degree of nominalisation, and thus deletion of participants, than informal letters. In other words, I have no problem with the fact that *interpersonal* distance is created through the absence of participants. But this does not mean, as I showed in chapter 4 and 5, that *ideational* distance is necessarily created by nominalisations through the absence of participants and the presence of ‘object-like’ nominals. All the same, in contrast to CDA, what I have tried to show in this chapter is that ideational distance *can* be related to a particular discourse derivation *even if participants are present*. That is, just because participants are present in a text does not necessarily mean that there are no differences in the ‘depth’ to which different participants are processed in *discourse*. Secondary characters are *automatically* read in a more shallow way, than primary ones, for a reader who has little vested interest in a text. Or put another way, in the discourse of such a non-analytical reader, secondary characters are *ideationally distant*.

### *Instantiation*

Consider again (from 3.7.2) the following from Clark (1992: 215), which comments on the Steed text above:

In both descriptions of the rapes (1) and (2), the perception of Steed as rapist is reduced by making the sentences passive and deleting him as Agent. This perception is further reduced by using the euphemism 'attacked' to mask the terrible details of abduction, repeated rape, and death threats (not mentioned at all in this newspaper).

I pointed out, in 3.7.2, that there was no mention of how 'attack' might be *instantiated* by the concrete details in the text. This is partly because of logical empiricist / symbolism assumptions common in CDA that all the information should be 'in' the sentence, i.e., the structure of the sentences should mirror reality regardless of the processor's (inferential) input. But we have seen that instantiation takes place automatically when it is based on well-known encyclopaedic knowledge. In the minimalist hypothesis, this includes information already established in the text (McKoon and Ratcliff, 1992: 440). Thus, IR's discourse would readily include instantiation of 'attack' (sentence 2) via the more specific 'rape' information (from sentence 1), especially given that when Steed 'attacks', he is attacking his *third* victim, having raped his *first* and *second*. [This also applies to Fairclough (1992a: 170-1) and what I said about instantiation of the category 'abnormalities' (3.6.1)].

From all of this, we can see that Simpson's (1993: 171) second alternative sentence for the Sun text:

2) Steed attacked at gunpoint his third victim, a 39-year-old mother of three, after he had forced her car off the M4

in containing the more general 'attack', but lacking a reference to 'rape', is actually 'euphemistic' in Clark's sense; or in terms of inference generation, 2) lacks enough information to instantiate 'attack' and so ironically mystifies in *discourse* as to the nature of the 'attack'.

### 6.6.3 Conflict between the Discourse of the Idealised Reader and CDA Explanation (Interpretation-2) for a Non-Resistant Reader

*Fairclough (1995a: 122-3)*

Recall from 1.4 Fairclough's line that if the textual material is unfamiliar, then, readers would commit themselves to inferential labour to 'fill in' 'implicit links'. Now, let me reproduce part of a text which Fairclough (1995a: 122-3) analyses and then some of Fairclough's analysis:

[*MIX to pipes in slum area of Manila, pan to wide-shot slums*]

[Narration]

1) Everywhere in the Third World life in rural areas gets harder - and poor people flock to the city.

2) The urban poor get poorer.

[*Close-up child standing in pipe; Slum area, mother and child*]

3) When rice prices go up, hunger and unrest grows.

4) In the city, the people can usually be kept in their place.

Cohesion relations are largely implicit in this sequence. For example, I interpret the clauses of the first three sentences as in relation of enhancement, and more specifically cause-effect relations, but they lack markers of causal cohesion. In sentence 1, the two clauses are linked by the all-purpose conjunction *and*, which leaves implicit the cause-effect relation (poor people flock to the city *because* life in rural areas gets harder). I also see an unmarked cause-effect relation between the second clause of sentence 1 and sentence 2 (the urban poor get poorer *because* so many people flock to the city). Again, although the first clause of sentence 3 is marked with a temporal conjunction (*when*), there is an implicit causal relation between the two clauses (hunger and unrest grow *because* rice prices go up). **It takes quite an inferential leap on the interpreter's part to establish a coherent meaning relation between sentences 3 and 4.** I interpret this an extension-type relation of an adversative type (unrest grows, *but* the people can usually be kept in their place; or *although* unrest grows, the people can usually be kept in their place). The connection between these sentences rests upon a 'bridging assumption' (Brown and Yule 1983, Fairclough 1992a): that popular unrest gives rise to a problem of order, and the need for official action to try to contain it. [my bold]

Before I continue with quotation from Fairclough, let me highlight Fairclough's point above that 'it takes quite an inferential leap on the interpreter's part to establish a coherent meaning

relation between sentences 3 and 4'. I will come to the significance of this highlighting shortly. Fairclough continues:

Overall, this part of the extract addresses an ideal interpreter who is familiar with a particular preconstructed 'script' (Montgomery *et al.* 1989) that is being evoked here: a predictable sequence of events leading from rural poverty to urban squalor and unrest and consequential problems of order. **The ideal interpreter is relied upon to fill in the gaps, make explicit what is left implicit, and construct a coherent, preferred, meaning for the text.**

But this is not just a matter of textual economy, not bothering to spell out what can be taken for granted. **It is a moot point how many real audience members might, if asked, actually agree with the stereotypical narrative of Third World urban problems which constitutes the script. But the text takes the script as universally given for its audience, and so positions audience members that they are induced to draw upon it to arrive at a coherent interpretation...**Local coherence relations in cases of this sort can therefore contribute significantly to textual processes of ideological interpolation (Althusser 1971): audience members are, so to speak, called upon to acknowledge the framework of ideological common sense (in this case, the Third World script) within which they are positioned. Such texts can cumulatively shape the knowledge, beliefs and values of audience members. [my bold]

Examining the above, we can suppose that Fairclough has a non-resistant or non-analytical reader in mind, gleaned from what Fairclough's says about it being a 'moot point how many real audience members might, if asked, actually agree with the stereotypical narrative of Third World urban problems which constitutes the script.' That is, if such a reader were *prompted* to think about the stereotypical narrative of Third World urban problems, they might not agree with it, given the extra cognitive effort incurred by the prompting. But without such prompting they will be *compliant* with the way in which the text 'positions' them.

For Fairclough, a reader ('the ideal interpreter') will invest the necessary cognitive labour to make the 'bridging assumption' he outlines between sentences 3 and 4. As I have highlighted, Fairclough is explicit that this involves a fair amount of cognitive labour ('quite



an inferential leap'). But we saw from the experiments of Vonk and Noordman (1990) (see 6.2.2) that even coherence inferences (i.e., 'bridging assumptions') can be *shallow* if the textual material is not so familiar. And indeed, I would argue that for many readers, 'problems of order following popular unrest and the need for official action to try to contain it' is hardly an everyday, familiar scenario. So Fairclough over-estimates the amount of cognitive work readers are prepared to invest since the 'inferential leap' would not be made, at least, by a reader with no specific goal, with little vested interest in the text, and one *largely unfamiliar with subject matter* (i.e. IR).

For Fairclough, the absence of *cohesive* markers is significant since readers are, then, induced to draw upon a particular script of the third-world in order to 'arrive at a coherent interpretation'. But as we saw in the experimental work of Vonk and Noordman (1990) (6.2.2), even if cohesive markers *are* present, if the reader is *unfamiliar* with the textual material, then, coherence will inevitably be shallow, readers satisfying themselves that the text is coherent because it is *cohesive*.<sup>8</sup> IR's discourse, then, conflicts with Fairclough's *explanation* (interpretation-2). But Fairclough's explanation itself, it could be argued, rests on a conceptual tension. In being willing to make the cognitive effort, to make 'quite an inferential leap', one might assert that Fairclough's 'ideal interpreter' (see bold above) is someone *with* a vested interest in the text. But at the same time, for Fairclough, the 'ideal interpreter' is a non-resistant reader, someone who passively accepts the positioning of the text and thus someone who presumably does not make a great deal of cognitive effort, i.e. one who would *not* make 'quite an inferential leap'.

*Gough and Talbot (1996)*

Now recall from 1.4 the problem page advice commented upon by Gough and Talbot (1996:

226):

Many heterosexual men have a passing curiosity about homosexuality, and that isn't such a bad thing. It compels you to make choices.

...the **causal link** which is needed to coherently combine these two sentences is not cued by any formal element, and this is a point where **the reader's complicity** is required if the two sentences as they stand together are to make sense. **The 'missing link' we need to supply is that heterosexuality and homosexuality are separate sexualities and that interest in homosexuality is useful inasmuch as it reinforces this separate heterosexual identity.** For some readers it may require inferential work.

...Following Fairclough's [1989] approach, this interpretation would be accounted for using the notion of automatic 'gap-filling' between explicit propositions. A reader who is unfamiliar with problem pages...would need to engage in a good deal of inferential work to make this connection. [my bold]

The notion of a non-resistant reader here is more implicit than in Fairclough's text above but, all the same, the phrase 'the reader's complicity' indicates Gough and Talbot do have a non-resistant reader in mind. The assumption in the above, as we saw in 1.4.1, is that readers will 'work', even if they are 'unfamiliar with problem pages', to produce the inference that 'interest in homosexuality (in this context, by adolescent males) is useful in as much as it reinforces this separate heterosexual identity'. In doing so, the reader becomes ideologically positioned. But the inference that interest in homosexuality reinforces heterosexual identity is a *causal consequent* inference. As we saw (6.2.4), causal consequents are normally weakly generated since they are not required for coherence. This would be especially the case for IR, a reader with little vested interest in a text and largely unfamiliar with subject matter, or in this case a reader 'unfamiliar with problem pages'. Furthermore, like 'popular unrest leading to problems of order etc' the above inference is a rather non-everyday, abstract one that the minimalist hypothesis would discount as likely to be generated unless the reader had quite a specific goal in mind. The psycholinguistic evidence, then, contradicts Gough and Talbot's line.

Thus, what Gough and Talbot have produced is a *strategic* inference and one not generated by IR since only *automatic* inferences are a part of their discourse derivation. This strategic inference generation most probably reflects Gough and Talbot's (1996) *own vested interest* in the text, leading them to derive a different discourse from IR in their *explanation* of the text for a non-resistant reader. As a final point, Gough and Talbot's position (as well as Fairclough's earlier) seems somewhat akin to that of the *early* constructionists (6.2.2) where there are few restrictions placed around the types of elaborative inference produced, a position which the *later* constructionist position has moved away from (6.2.2)<sup>9</sup>.

*Fairclough (1995a: 113)*

Recall from 3.4.2 and 4.3.3 the following excerpt from Fairclough (1995a: 113):

Everywhere in the Third World life in rural areas gets harder - and poor people flock to the city. The urban poor get poorer.

Fairclough argued that 'flock' is 'associated with 'sheep - notoriously passive', thus reinforcing the perspective in the rest of the text that the 'poor' are not agents of their circumstances. There is no explicit positing of the perspective of a non-resistant reader. The perspective, however, is implicit - Fairclough is analysing the text presumably to indicate the reading of someone who does not oppose the text's 'view' of the poor. We saw, in 3.4.2, that an assumption which enables him to do this is that of compositionality where symbols are *discrete* and *enduring*. Compositionality, then, facilitates his choice of the ovine sense of 'flock' so as to suit his line. However, in 4.3.3, I showed how a connectionist perspective conflicted with Fairclough's explanation.

But how does all of this fit with IR? The low detection rates of lexical anomalies in the

experiments of Reder and Kusbit (1991) (6.3.2) and Barton and Sanford (1993) (6.3.1) suggest that it is *automatic* for the mental representation of a reader to be *non-compositional* with regard to lexical items in a text. This was because of top-down processing prevailing over the processing of discrete lexical items. I highlighted previously that to detect anomalies such as ‘surviving dead’, the necessary extra cognitive effort was in line with *compositional* scrutiny of the text. That is, compositional analysis and processing are *strategic* and not automatic. However, since IR invests minimum cognitive effort, IR’s processing is *automatic*, thus consisting of *non-compositional* mental representation. So, for IR, the mental representation of ‘flock’ has *partial* value and not the ‘full’ semantic value of ‘collection of sheep’. This, then, conflicts with Fairclough’s analysis of ‘flock’, which presupposes compositional mental representation in the mind of a non-resistant reader. In conclusion, Fairclough has produced a *strategic* inference to suit his line and so Fairclough’s discourse is at odds with the discourse derived from the text by IR. As a final point, since the text is accompanied by pictures [*MLX to pipes in slum area of Manila, pan to wide-shot slums*], presumably the pictures of slums (which also presumably show people living in them) will be involved in the top-down processing, once more making the ovine reading less likely for non-strategic IR.

## 6.7 Summary

There is not much citation of psycholinguistic work on inferencing in the CDA literature, so I have tried to broaden the discussion and prepare the ground for chapters 7 and 8. In the last section, broadly speaking I have shown three things:

i) psycholinguistic data from this chapter conflicts with *symbolic / logical empiricist* inspired

CDA. IR is a reader based on this psycholinguistic data. Thus, what *symbolic / logical empiricist* CDA regards as ‘mystifying’ text is *not* from the perspective of IR - i.e., does not lead to a mystifying discourse for IR.

ii) different levels of cognitive investment in a text lead to different discourses being derived from the text. In the last section, the different discourses were the non-shallow discourse generated by proxy for non-analytical readers by CD analysts, and the shallow discourse generated by proxy for IR by myself.

iii) CD analysts wrongly assume that the high cognitive effort they invest in interpretation by proxy for a non-resistant reader is replicated by such a reader who may have little vested interest in a text and so invest much less cognitive effort. By exploring the micro-context of interpretation (in contrast to CDA), I have shown that a non-analytical reader (i.e. IR) is *not* positioned by the above texts into making a particular interpretation as deemed by the CD analysts cited above. In other words, for the texts discussed in 6.6, (where readers are not coerced into interpretations), the assumption in CDA that the macro-context governs the micro-context of interpretation of a non-analytical reader is problematised.

Other similar things follow from the above:

- Since different levels of vested interest and thus different levels of cognitive labour lead to different (by proxy) interpretations, the principle of ‘partiality of interpretation’ from the same text has thus been demonstrated.
- CDA ‘over-interpret’ by proxy for the non-resistant reader.
- Since the discourse of *non-analytical* IR (i.e one in line with empirical psycholinguistic

evidence) conflicts with the discourse of the CDA *non-analytical* reader (i.e, a reader not supported by empirical psycholinguistic evidence), the notion that there is critical hermeneutic exegetic privilege in the above CDA explanations is considerably weakened.

In chapter 7, I shall highlight compatible elements in the psycholinguistic evidence in this chapter for shallow processing and the enterprises of connectionism and cognitive linguistics that were discussed in chapters 4 and 5 respectively. In chapter 8, these compatibilities will augment the basis of the idealised reader, creating a more systematic and comprehensive *alternative* framework for the analysis of shallow discourse. While I have shown in this chapter that the texts analysed by CD analysts are *not* mystifying of subject matter for IR, I will use this framework in chapter 8 to highlight text which *is* mystifying of subject matter for IR, but text which would not be highlighted as such by CDA.

#### Notes:

1. Although I do not deal with anaphoric references in this chapter since they do not inform my framework in chapter 8, there is ample evidence that coherence anaphoric inferences are made on-line (e.g. Dell, McKoon and Ratcliff, 1983).

2. Vonk and Noordman's (1990) experiments highlight one essential difference between *cohesion* and *coherence*.

3. Sanford and Garrod (1994: 704) point out that the 'primary processing principle' differs from the minimalist hypothesis 'in philosophy' since it makes:

'...the assumption that grounding incoming text in background knowledge is necessary to achieve a sense of understanding and to achieve coherence in many instances.'

However, Sanford and Garrod (1994: 704) also acknowledge that:

'...once possible elaborative inferences are based on the availability or accessibility of general knowledge, the

minimalist position and the primary processing accounts are not so easy to distinguish.’

4. Of course there are occasions when sections of texts are remembered verbatim: e.g. texts of personal significance, or statements of high interactional value (see: Keenan, MacWhinney and Mayhew: 1977). This can also apply to highly-patterned language read for the first time, e.g. lines from a poem or a song, some newspaper headlines, etc. But with regard to the last section of this chapter, since CDA has tended to deal with news texts of a reasonable length, rather than just headlines, I make the assumption that such texts will usually hold little personal significance for the reader or are so unpatterned and of such length that they would not ordinarily be remembered verbatim.

5. One of the conclusions that Wason drew from this research was that humans find it much more natural to search for proof than to search for disconfirmation / falsification, conflicting with Popperian methodology in the philosophy of science.

6. Take the *subject + [be] + past participle* form - *the window was broken*. How are we to analyse this? Is this a description of an *action* - an *agentless passive* where the past participle is the 3rd form of a verb and [be] an auxiliary verb? Or is it a description of a *state* where the past participle is an *adjectival complement* and [be] a copula? The answer is that analysis must take into account context and human motivation, i.e. *discourse*. For example, for ‘the window was broken’ where the context is a burglary, a detective may be more interested in the person who was responsible for the breakage rather than the state of the window and so see the form ‘was broken’ in terms of an agentless passive. Alternatively, imagine the context of a car safety experiment which is testing whether a particular windscreen glass is strong enough for a certain impact. In a report that ‘the window was broken’, ‘broken’ would be understood in terms of a stative adjective since the state of the window is more important than the agent of the breaking.

7. For descriptions of scientific phenomena, it is often irrelevant whether sentences are transactive or non-transactive if there is an absence of sufficient knowledge to draw upon. This is well exemplified by the following folk-explanation - *The ice cooled his forehead*. The transactive status of the sentence means that for Hodge and Kress (1993), at least, the causal relation is explicit, satisfying their criterion that this is a scientific statement. But the causal relation of the folk-explanation is wrong. Ice does not cool a forehead. Instead, heat from a forehead disrupts the bonds between water molecules in ice leading to melting. It is because heat is transferred to the ice, rather than the ice doing something to the forehead, that the temperature of the forehead reduces. For conflict between folk and scientific understanding of causality, see Wolpert (1992).

8. Fairclough’s (1989) distinction between *automatic connections* and *inferences which require interpretative work* (and thus the idea that readers are willing to invest in such inferential ‘work’ even if they are unfamiliar with subject matter) is found in Brown and Yule (1983). Here are Brown and Yule (1983: 266):

‘...the more interpretative ‘work’ the reader (hearer) has to undertake in arriving at a reasonable interpretation of what the writer (speaker) intended to convey, the more likely it is that there are inferences being made.’

Indeed, Fairclough (1989: 108) cites and endorses Brown and Yule's (1983) chapter 7 as 'a helpful discussion of inferencing, and its relation to automatic gap filling'. Gough and Talbot's (1996) position on inferences is derived from Fairclough (1989) which in turn derives from Brown and Yule (1983); see Gough and Talbot (1996: 226). Brown and Yule (1983) was published before the papers associated with the minimalist hypothesis and the later constructionists, these being published in the late 80s and early 90s. Consequently, Fairclough's sources are not up-to-date and consensus thinking on inference generation. Brown and Yule's position, and in turn Fairclough's, is contradicted not only by the minimalist hypothesis and the later constructionists but by Sanford and Garrod's (1981 / 1994) primary processing principle and Vonk and Noordman (1990).

9. The later constructionists, Graesser et al. (1994: 384) indicate that the types of inference that are generated on Schank and Abelson's (1977) model include elaborative inferences such as causal consequence:

'This model asserts that reading is expectation-driven in addition to explanation-driven. That is, readers generate expectations about future occurrences in the plot, and these expectations guide the interpretation of clauses in a top-down fashion (Bower et al., 1979; DeJong, 1979; Dyer, 1983; Schank and Abelson, 1977). Expectations are formulated whenever higher order knowledge structures are activated, such as a script or a theme. For example, if a story activates a RESTAURANT script and the text mentions that two characters entered a restaurant together, then the reader would form expectations that the characters will eat, talk and be served food. If a story activates a REVENGE theme and the text specifies that character A hurts character B, then the reader would form the expectation that character B will hurt character A.'

However, the fact that Schank and Abelson's (1977) model includes fully formed causal consequent inferences is in direct conflict with both the later constructionists and the minimalists. Now, Schank and Abelson (1977) is cited in Gough and Talbot (1996: 228) which is unsurprising when we consider that Schank and Abelson script theory is drawn upon in Fairclough (1989), a source work for Gough and Talbot (1996). It is likely, then, that Fairclough (1989) and thus Gough and Talbot (1996) take for granted that causal consequents are produced since Fairclough directly and Gough and Talbot indirectly draw upon Schank and Abelson (1977).



## **CHAPTER 7: HIGHLIGHTING COMPATIBLE ELEMENTS IN CONNECTIONISM, COGNITIVE LINGUISTICS, AND PSYCHOLINGUISTIC EVIDENCE FOR SHALLOW PROCESSING**

### **7.1 Introduction**

In chapter 8, I will enhance the idealised reader (IR) framework to show how certain news text will lead to mystification for that non-analytical reader, news texts which would not be highlighted as such by CDA. As we saw in the last chapter, the IR framework is (so far) based on psycholinguistic evidence for shallow or selective processing. When I enhance the IR framework in chapter 8, it will also include compatible ‘principles’ from connectionism and cognitive linguistics which are in turn compatible with the psycholinguistic evidence for shallow processing discussed in the last chapter. The main purpose of this chapter is to prepare for chapter 8 by indicating compatibilities between *all* these areas. However, in 7.5, I outline *tensions* between connectionism and cognitive linguistics with regard to treatment of metaphor and explore the ramifications of this for CDA’s use of Lakoff and Johnson’s (1980) approach to metaphor.

### **7.2 Compatibility of Connectionism with Cognitive Linguistics**

#### *7.2.1 General Compatibilities*

There are many parallels between cognitive linguistics and connectionism. The cognitive linguist, Langacker (1987a), argues that it is unhelpful to treat syntax as autonomous and thus separate from semantics, and as we have seen already (chapter 4), connectionist models

do not treat syntax and semantics in this way. Cognitive linguistics and connectionism also intersect in their eschewing of logical reasoning and rule application as being central to an understanding of language performance (cf. Wason, 1966 and in 6.4.2). On this comparison and on the shunning of ‘propositional formats’ and transformational derivation (such as inherent in the derivational theory of complexity, a performance theory similar to that used by CDA) and the notion that linguistic systems are recurrent patterns of activations, here is the cognitive linguist, Langacker (1987b: 6):

...cognitive grammar (at least my own formulation of it) is basically compatible with the connectionist philosophy. First, cognitive grammar makes no qualitative distinction between rules and their instantiations - rules are simply schematized expressions; moreover, the ‘schemas’ in question are thought of as being ‘immanent’ to their instantiations, not as separate or discrete structures. Second, only elements with semantic and / or phonological content are permitted, and they are characterized directly in terms of such content, not in a propositional format. Third, analyses are based on the overt form of expressions; derivation from abstract, ‘underlying’ representations is precluded, as is any sort of algorithmic computation. Finally, a linguistic system is viewed as simply an inventory of ‘cognitive routines’, which are interpretable as recurrent patterns of activation that are easily elicited by virtue of connection weights; the construction of complex expressions reduces to the co-activation of appropriate routines and ‘relaxation’ into a pattern of activation that simultaneously satisfies all constraints.

So Langacker, then, explicitly recognises compatibility between cognitive linguistics and connectionism. Similar explicit recognition of these compatibilities is found in the connectionist philosophers, Churchland and Churchland (1996: 238), who draw parallels with the work of the cognitive linguists (Lakoff, Langacker, Bates, Fauconnier) and how they, like connectionists, have *not* treated syntax and semantics as being separable; see also Harris (1990), Langacker (1987a: sec. 12.3), Langacker (1997), Regier (1996), Schopman and Shawky (1996) for discussion of cognitive linguistics in the context of connectionism.

As we saw in 4.2.5, one of the pioneers of cognitive linguistics, Rosch, endorses the connectionist view that symbols are ‘approximate macrolevel descriptions of operations

whose governing principles reside at a subsymbolic level' (Varela, Thompson and Rosch, 1991: 101-2). Also for the cognitive linguist, Lakoff (1987a: 338):

Our results...do not contradict what have come to be called 'connectionist' theories...

In section, 7.5, I shall show that there are in fact some tensions between connectionism and the cognitive linguist, Lakoff, particularly with regard to approaches to metaphor.

### *7.2.2 Surface Structure (Linguistic Input) Not Equivalent To Process Output*

#### *Accommodation*

Another area of compatibility between cognitive linguistics and connectionism is in the notion of *accommodation*. For the *cognitive linguist*, Langacker (1987a: 76-7), 'accommodation' refers to the adjustment of a lexeme's meaning according to the lexical company it keeps. For example, 'ball' in 'the toddler kicked the ball' and 'the final ball of the over was a googly' are likely to be different. The connectionist Elman (1992: 170) makes explicit mention of Langacker's concept of 'accommodation', citing Langacker (1987a: 76-7) at length, which I reproduce below:

It must be emphasized that syntagmatic combination involves more than the simple addition of components. A composite structure is an integrated system formed by coordinating its components in a specific, often elaborate manner. In fact, it often has properties that go beyond what one might expect from its components alone. Two brief observations should make it clear why this is so. First, composite structures originate as targets in specific usage events. As such they are often characterized relative to particular contexts with properties not predictable from the specifications of their components as manifested in other environments. A related point is that one component may need to be adjusted in certain details when integrated with another to form a composite structure; I refer to this as **accommodation**. For example, the meaning of run as applied to humans must be adjusted in certain respects when extended to four legged animals such as horses, dogs, and cats (since the bodily motion observed in two legged running is not identical to that in four legged running); in a technical sense, this extension creates a new **semantic variant** of the lexical item.

Indeed, Elman's model was able to simulate accommodative effects (Elman, 1992: 170).<sup>1</sup>

Langacker's concept of 'accommodation' also dovetails with the capacity of McClelland and Kawamoto's (1986) network to shade meaning, which we saw in 4.3.2. For example, in the initial input, all 'balls' were assigned the microfeature 'soft.' But for the sentence 'the ball broke the vase', the network assigned the microfeature 'hard' to 'ball' since it knew from the training corpus that all things for breaking were 'hard'. This shading of meaning in a connectionist network is an *automatic* process. 'Shading of meaning' in humans intuitively appears to be an automatic process also. If connectionist networks can be treated as a simulation of brain networks, then, McClelland and Kawamoto's (1986) simulation lends support to the notion that shading of meaning or accommodation in a *brain* network *is* an automatic process.

#### *Safe-Combination Metaphor vs Building-Block Metaphor*

Elman (1990: 378) offers the 'safe-combination' metaphor as an explanatory metaphor for language processing in connectionist networks:

'A metaphor that captures some of the characteristics of this approach is the combination lock. In this metaphor, the role of words is analogous to the role played by the numbers of the combination. The numbers have causal properties; they advance the lock into different states. The effect of a number is dependent on its context. Entered in the correct sequence, the numbers move the lock into an open state...The numbers are 'present' insofar as they are responsible for the final state but not because they are still physically present.'

The safe-combination metaphor contrasts with the classical 'building-block' metaphor, (which I referred to in 2.2.4). In the building-block metaphor for language processing, process output still 'contains' the building blocks of linguistic input. However, in the safe-combination metaphor for language processing, linguistic input is responsible for the final

state and so is not ‘present’ in this final state. Linguistic input is not the same as the process output. This contrasts with how CDA often treats semantic-syntactic surface structure as equivalent to process output as we saw in chapter 3. Langacker’s (1987a: ch. 12) ‘scaffolding’ metaphor, which we saw in 5.7, has much in common with Elman’s ‘safe-combination’ metaphor. He views the (linguistic input) components of a compound as ‘scaffolding’, i.e. disposable when no longer needed. Seeing compounds in terms of the scaffolding metaphor rather than building block metaphor places a greater emphasis on how cognition *output* does not necessarily mirror linguistic *input*. For example, if one of the lexical items in a sentence is a basic-level category then it will be in cognitive interdependence with either action and object basic-level categories that are not present in the linguistic input.

### 7.2.3 *The Functionalism of Cognitive Linguistics and Connectionism*

Elman (1990: 378-9) then goes on to highlight the *functional* aspect of the connectionist approach to language comprehension:

This view of language comprehension emphasizes the functional importance of representations and is similar in spirit to the approach described in Bates and MacWhinney 1982; McClelland, St. John, and Taraban 1989; and many others who have stressed the functional nature of language. Representations of language are constructed to accomplish some behaviour. Obviously, that behaviour may range from daydreaming to verbal duels, from asking directions to composing poetry. The representations are not propositional, and their information content changes constantly over time in accord with the demands of the current task. Words serve as guideposts that help establish mental states that support this behaviour; representations are snapshots of those mental states.

As Elman says, the functional importance of language is something which other connectionists, McClelland, St. John and Taraban (1989) (see 4.4), have emphasised also.

But as we saw in chapter 5, the prominence of this aspect of language is compatible with

a similar emphasis in cognitive linguistics, where in contrast to classical semantics, it draws attention to the *interactional* properties of an object. That is, language may be used not just to institute a *conceptual* state in a reader's or listener's mind but to institute a *functional* state. Recall from 5.3.2 the example of the 'internal mouse' on a laptop computer which has different inherent properties to an external mouse. Referring to this item as a 'mouse', that is, employing a *basic-level* category, enables the hearer to realise the particular *motor-interactionality* associated with a laptop; in other words the actual *function* of the mouse. The hearer experiences no difficulty in activating the referent's function and allowing it to prevail over its form; the hearer perceives there to be no conflict.

### *Generation of Instruments*

As we saw in 6.2.6, humans have the ability to generate instrument inferences readily if the context constrains. In 4.4.2, we also saw that connectionist networks are successful in replicating this human ability. McClelland, St. John and Taraban's (1989) connectionist network was able to generate readily the instrument inference 'with a knife'. In a sense, also McClelland and Kawamoto's (1986) network (see 4.3) was able to fill in the 'missing instrument argument' for 'the boy broke the window', 'inferring' that the boy broke the window 'with something hard'.

So connectionist models can replicate human ability to generate instruments. What now about the relationship between cognitive linguistics and instrument generation? Recall the experiment from 4.2.2 of Taraban and McClelland (1988). The evidence from Taraban and McClelland's (1988) experiments was offered as refutation of the syntax first / minimal attachment perspective (2.5.4). Consider again the following which subjects were asked to read in the experiment:

The janitor cleaned the storage area with the

- a) broom
- b) solvent
- c) manager
- d) odour

because of many complaints.

For Taraban and McClelland (1988), the above sentence contains a verb which leads to the expectation that an *instrument* will be mentioned in the PP; that is, the PP will attach to the VP [inadvertently in accordance with the minimal attachment principle]. The experimental evidence bore out Taraban and McClelland's suppositions that 'with the broom' is the most expected role. This can be explained from a cognitive linguistic perspective. In 5.4.2, we saw that, for Lakoff (1987a: 54), prototypical causation is understood in terms of a cluster of 'interactional properties' since it involves a high degree of direct physical manipulation, the agent using his hands, body or some instrument. The *function* of the janitor also cues in that direction. It is not surprising, from the perspective of cognitive linguistics, then, that people choose 'broom' more readily than 'manager' or 'odour' and why Taraban and McClelland regard 'broom' as the most expected.

#### 7.2.4 Prototypes

In 5.4, I highlighted the relationship between basic-level categories and prototype effects, e.g., that prototype effects are more readily generated from basic-level categories than superordinate ones. Though the provenance of prototype theory is the cognitive linguistics of Rosch etc, connectionism and prototype theory are not incompatible. Here is Harley (1995: 196):

[Connectionism]...is not necessarily a competitor to the prototype theory; one instance of a category might

cause one pattern of activation across the semantic units, another instance will cause another similar pattern and so on. We can talk of the prototype that defines that category as the average pattern of activation of all the instances.

Let me just flesh out this relationship between connectionism and prototypes. In essence, because of the distributed encoding of a set of exemplars, the microfeatures common to the exemplars become strongly associated. In other words, common microfeatures form strong mutually excitatory links. A prototype emerges thus, an automatic consequence of connectionist networks. Interestingly, the prototype does not need to coincide with any concrete instance exposed to the system. This is because the prototype is the statistical central tendency of the various *microfeature* dimensions of the exemplars (Clark, 1993: 22). Prototype emergence goes hand in hand with *generalisation*. A net is said to be able to generalise if it can handle novel cases. For example, McClelland, Rumelhart, and the PDP Research Group (1986, chp17) were able to produce a 'dog-recognition network' which recognised a three-legged dog as a dog, something that would seem to negate against 'dogness' on the classical perspective where meaning is predicated upon necessary features. From a connectionist point of view, a triped dog is still a dog because it still shares the vast majority of *microfeatures* of the *prototypical* dog. The fact also that prototypes *emerge* in connectionist networks dovetails with the averring of the cognitive linguist, Rosch (1978) (see 5.2.1), that prototypes are not *stored* representations.

I have outlined some general compatibilities between connectionism and cognitive linguistics. In sections 7.3 and 7.4 respectively, I outline how cognitive linguistics and connectionism respectively relate to the psycholinguistic evidence I outlined in the last chapter for *shallow processing*.



### 7.3 Cognitive Linguistics and Shallow Processing

#### 7.3.1 Cognitive Linguistics, Instantiation and Shallow Processing

As we saw in 6.2.5, instantiations are elaborative inferences where a general category is made more concrete. For instance, in ‘the vehicle hovered over the crowd’, likely inferences for ‘vehicle’ are that it is a balloon or a helicopter. In line with the minimalist hypothesis of McKoon and Ratcliff (6.2.2) in the absence of specific reading goals, instantiations will be readily generated when information is quickly and readily available. Recall now from 5.6.3 the well-known (inconsiderate) text from Bransford and Johnson (1973: 400):

The procedure is actually quite simple. First you arrange things into two different groups. Of course, one pile may be sufficient depending on how much there is to do. If you have to go somewhere else due to lack of facilities, that is the next step; otherwise you are pretty well set. It is important not to overdo things. That is, it is better to do fewer things at once than too many. In the short run this might not seem important, but complications can easily arise. A mistake can be expensive as well. At first the whole procedure will seem complicated. Soon, however, it will just become just another facet of life. It is difficult to foresee any end to the necessity for this task in the immediate future, but then one never can tell. After the procedure is completed, one arranges the material into different groups again. Then they can be put into their appropriate places. Eventually they will be used once more, and the whole cycle will then have to be repeated. However, that is part of life.

In referring to the above text, Rosch (1978: 45) makes the point that:

...what Bransford and Johnson call context cues are actually names of basic-level events (e.g. washing clothes) and that one function of hearing the event name is to enable the reader to translate the superordinate terms into basic-level objects and actions.

What Rosch terms ‘translation’ is actually the generation of *instantiation* elaborative inferences via the basic-level event category, ‘washing clothes’. Now, by relating cognitive linguistic explanation to the *minimalist hypothesis*, let me go further than I did in 5.6.3 as

to why a basic-level category leads to translation (instantiation) and ready comprehension, and also why the superordinate categories (e.g. 'procedure', 'groups', 'things'...) above lead to *shallow* comprehension.

Firstly, let me just recap for a moment on the nature of basic-level categories which were examined in chapter 5. The basic level of categorisation is characterised by *cognitive economy* (Rosch, 1978). It is the level of categorisation where the largest amount of information about an item can be obtained with the *least cognitive effort*. Basic-level categories (eg. 'chair') are also associated with prototype effects. Since prototypes are by definition *familiar*, use of basic-level categories will facilitate instantiation since, from the minimalist hypothesis, instantiation is dependent on accessible knowledge. Furthermore, such instantiation can be taken as being *automatically* generated (i.e. with minimum cognitive effort) in line with the cognitive economy principle. Superordinate categories are not associated with a prototype and not characterised by cognitive economy. To yield a large amount of (familiar) information from them, a higher degree of cognitive effort is required in comparison to processing of basic-level categories. But again in line with the minimalist hypothesis, if a reader has no specific goal (i.e., does not invest high cognitive effort), in the absence of sufficient basic-level categories, superordinates do not lead to ready accessing of information. In other words, *on their own* superordinates such as in the Bransford and Johnson (1973) text are more likely to lead to *shallow* comprehension and thus mystification for a reader who does not make much cognitive effort. Of course, with the inducement of a cash prize for the 'answer' to Bransford and Johnson's (1973) text, a higher degree of cognitive effort is prompted, leading to a greater likelihood of the general categories being instantiated regardless of basic-level categories being absent. However, these would be instantiations *strategically* generated rather than automatically.

As a caveat to the above, I should say that basic-level categories are not necessarily all that is needed to instantiate a general category. The context will also need to be sufficiently constraining. But at least basic-level categories, which exist in relationships of cognitive interdependence with other basic-level categories, are more likely to contribute to instantiation. Or indeed, contexts which indicate *motor-interactivity* (i.e. associated with basic-level categories) are more likely to lead to instantiation of a general category. For example, McKoon and Ratcliff (1989b: 1145) found that the superordinate category ‘furniture’ was readily instantiated by experimental subjects in the following:

While the movers took a break, Betty went to her room for a quick nap, thankful that at least one piece of furniture had not been loaded yet.

From a cognitive linguistic perspective this can be explained by Betty’s *motor-interactive* behaviour.

### 7.3.2 Cognitive Linguistics, Instrument Inferences and Shallow Processing

Recall from chapter 6 the section on instrument inferences (6.2.6). An instrument has a higher probability of being generated when it has been referred to already in the text (Lucas et al., 1990) or when context is sufficiently constraining (O’Brien et al., 1988). We looked at how O’Brien et al. (1988) employed passages to determine whether readers made any type of elaborative inferences:

- 1) All the mugger wanted was to steal the woman’s money.
- 2a) But when she screamed, he **stabbed** her with his **weapon** in an attempt to quiet her.
- 3) He looked to see if anyone had seen him.
- 4) He threw the **knife** into the bushes, took her money, and ran away.

[2b) But when she screamed, he assaulted her with his **weapon** in an attempt to quiet her.]

[my bold]

When subjects read 1) + 2a), they confirmed O'Brien et al.'s hypothesis that 'knife' would be automatically inferred. We saw in 5.5.1 that, from the point of view of cognitive linguistics, there is a strong cognitive interdependence between action and object basic level categories. And this 'cognitive interdependence' explains the ready generation of an instrument. For example, 'stabbed' gives rise to extra information since it suggests motor-interactionality, and leads to the generation of a gestalt and therefore ease of understanding. 1) + 2b), however, did not strongly suggest 'knife' to subjects since 'weapon' and 'assault' are *superordinate* categories. That is, the superordinate categories in the above lead to shallower processing than in 1) + 2a) since the instrument inference is not generated. The same reasons as to why, outlined in 7.3.1, apply.

In chapter 6, I outlined Sanford's (1990) reasons why Garrod and Sanford's (1982) experiments supported the generation of implied instruments on line whereas Singer's (1979) experiments did not. For Sanford (1990), the difference appears to be that with the Garrod-Sanford set, the instruments are 'part of the meaning' of the verbs (e.g., *key* is 'part of the meaning' of *unlock*), whereas for Singer's verbs, this was not the case. 'Part of the meaning' can be seen in cognitive linguistic terms. The instrument 'key' is readily generated from the verb 'unlock' because 'key' and 'unlock' exist in a relationship of *cognitive interdependence*, given that 'key' is basic-level and yields high information from minimum cognitive effort.

### 7.3.3 Cognitive Linguistics, Causal Consequents and Shallow Processing

I indicated in 6.2.4 that, where a situation is not highly familiar, causal consequent inferences are likely to be only *shallowly* generated, at best, by a reader with no specific reading goal, given the extra processing effort involved in forecasting possible outcomes.

Conversely, a causal consequent inference is only likely to be generated on-line, by a reader with no specific goal, if it is *highly* constrained by a context which is very familiar, and few if any other consequences would occur. I provided an example of a highly constrained context found in Keefe and McDaniel (1993: 454):

After standing though the three-hour debate, the tired speaker walked over to his chair. He realized that his valiant effort was probably in vain.

When the above was given to subjects, Keefe and McDaniel (1993) found that the causal consequent inference - *the speaker sat down* - was inferred a significant number of times.

This was put down to the high level of semantic association between 'chair' and 'sitting'.

But a cognitive linguistic explanation can also be furnished here. In the above, there is cognitive interdependence between the basic-level category 'chair' and the verb 'sit' since 'sitting' is a motor-interactive (ontologically subjective property) of a chair. Thus, the information 'sitting' becomes available readily with the information 'chair', with only a small investment of cognitive labour (again basic-level categories being characterised by cognitive economy). Or in terms of the minimalist hypothesis, the causal consequent is generated because necessary information can become quickly available.

Imagine now the above text from Keefe and McDaniel (1993: 454) with 'his chair' replaced by the superordinate 'the furniture':

After standing through the three-hour debate, the tired speaker walked over to the *furniture*. He realized that his valiant effort was probably in vain.

In terms of the minimalist hypothesis, superordinate categories do not make necessary information available for reasons we saw in 7.3.1. The causal consequent of ‘sitting’ is unlikely to be generated *automatically* with *little effort*. So, in the absence of sufficient information to instantiate them, and for a reader who makes minimum cognitive effort, superordinates are only likely at best to lead to *shallow* generation of causal consequents. [Naturally, though, the causal consequent of ‘sitting’ above could be generated *strategically*].

## 7.4 Connectionism and Shallow Processing

### 7.4.1 Connectionism, Moses Illusion and Shallow Processing

In 6.3, invoking psycholinguistic experimental work, we saw that because of commonplace top-down driven expectation in language processing, mental representation is likely to be *non-compositional*. [Conversely, *compositional* analysis of a sentence is in line with a greater investment of cognitive effort]. Now, interestingly, in connectionist networks, sentence processing is also non-compositional. In the connectionist network of McClelland, St. John and Taraban (1989), there was a tendency for the network to be too sensitive to context, particularly with the sentence ‘the adult ate the steak with daintiness’:

...in that in fact it [the connectionist network] allows context sometimes to override the correct interpretation of a word...After the presentation of *with daintiness*, the activation of *steak* on probing for the patient is weakened. In fact, at earlier points in learning, the model actually activates soup more strongly than steak after *with daintiness* is presented.

McClelland, St. John and Taraban (1989: 322-3)

McClelland et al. (1989) point out that while this behaviour can be seen as erroneous, it is nevertheless compatible with a common type of human error where *top-down* processes of schema expectation override bottom-up processes, i.e. with the psycholinguistic evidence I outlined in 6.3. Indeed, McClelland et al. (1989: 323) explicitly cite Erickson and Mattson's (1981) evidence (which we saw in 6.3.2) for the common shallow processing of sentences such as: 'how many animals of each kind did Moses take on the ark?'. For McClelland et al., errors of the Moses Illusion type indicate that the classical theme of compositionality misconstrues the contributions of words to how sentences are processed, and the frequency of the Moses Illusion implicitly lends support for the connectionist stance which eschews output compositional representations. Similar points concerning the compatibility of connectionist processing and the routineness of shallow processing of the Moses Illusion type are made in St. John (1992).

One of the drawbacks of connectionist models is that if they are taken as simulating brain networks, then their simulation is only one of *automatic* brain processing rather than more consciously directed *strategic* brain processing. However, since I am only concerned with automatic processing in section C, this drawback is not significant for my thesis. So, irrespective of this disadvantage, the *non-compositional* processing of McClelland, Taraban and St. John's (1989) connectionist network still lends support to the psycholinguistic evidence that non-compositional language processing is also *automatic* in humans.

#### *7.4.2 Connectionism, Sanford and Garrod's Primary Processing Principle and Shallow Processing*

We saw in 4.5.2 how in connectionist systems, 'connections between actions and the facts of the world can be represented as statistical correlations' (Waltz, D. 1989: 58). There is an

emphasis on a *good fit* between language and the world rather than an *exact match*, the latter being a facet of symbolicism. In 6.3.1, we saw that Sanford and Garrod (1994: 710-1) place emphasis on coherence being achieved ‘on the basis of a good fit between some elements of the sentences concerned and stereotyped knowledge’. It is this common need of readers merely to find a ‘good fit’ which accounts for the error in the ‘air crash survivors’ puzzle (see 6.3.1). In a reference to the puzzle, Sanford and Garrod (1994: 713) state:

Because *survivors* is a word which fits with an air crash, it is accepted as a filler for the patient slot of the verb *to bury*. [my bold]

In fact, Sanford and Garrod (1994: 715) are explicit about the link between the shallow processing of the connectionist network of McClelland, Taraban, and St. John (1989), alluded to in the last section (see also 4.4), and shallow processing of the ‘survivors puzzle’:

Our interpretation of the survivors problem data is as follows. First, the data show that in the presence of sufficient expectation of accident victims who are dead, the analysis afforded to items which should be consonant with that expectation receives relatively shallow processing. Another way to think of it is that these items exert little bottom-up impact because contextual inferences are so strong. The kind of mechanism we have in mind is that described by McClelland, Taraban, and St. John (1989).<sup>2</sup> [my bold]

#### *7.4.3 Connectionism, The Minimalist Hypothesis and Shallow Generation of Forward Elaborative Inferences*

##### *Orientation*

In 6.2.2, we saw how there was a consensus in *psycholinguistics* (the minimalist hypothesis / later constructionism) that causal consequent inferences are at best only *shallowly or partially* generated for a reader with little vested interest. For minimalists and later constructionists, causal consequent inferences are much less *probably* generated than causal



antecedent inferences. In 3.7.3, I referred to Schank and Abelson (1977), where inference generation is regarded as an *all-or-nothing* process. In other words, inferences cannot be drawn partially.

Now consider the following from the connectionist authors St. John and McClelland (1992: 122):

In parallel constraint satisfaction,...as in sentence comprehension, inference-making is inherent to processing the explicit text. The combined set of constraints from the text are used as evidence to support an interpretation that best satisfies the strongest and the most constraints from the text. Each part of the text supports many aspects of the interpretation. A specific proposition, therefore, will constrain the interpretation to represent the information it contains explicitly, **and also constrain it, to varying degrees, to represent correlated information. All of this correlated information constitutes inferences.**

The reader's hope is that by the end of the story, sufficient constraints have been provided by the text for a complete interpretation to be computed. The interpretation would, at that point, contain fully activated coherence inferences, **partially activated prediction [elaborative] inference**, and resolved pronouns.

[my bold]

Because the processing of connectionist nets is based on weightings:

...the constraints in a parallel constraint satisfaction model can represent the **likelihood of information.**

Interpretations can then be activated according to their degree of support from the text and their likelihood to occur.

St. John and McClelland (1992: 123) [my bold]

Due to its capacity to capture probabilities, then, a connectionist network seems suitable for capturing the probabilistic nature of inference generation which we saw in the psycholinguistic evidence from chapter 6. Indeed, in St. John and McClelland (1992: 121), there is explicit reference to McKoon and Ratcliff (1986), the originators of the minimalist hypothesis, and their experiment where death after a 14-storey fall was only weakly inferred as a causal consequent inference (see 6.2.2).

*St. John's (1992) Short Text Processing Model and the Minimalist Hypothesis*

In what follows, I outline a model for the processing of a simple text from St. John (1992) to indicate how a connectionist model is especially suited to what the minimalist hypothesis predicts, i.e., that forward-elaborative or predictive inferences are generated in a more *shallow* manner than backward coherence inferences. Before I do so, let me highlight some preamble of St. John before he describes the model. I do this since it will bear upon how I assess the model later. After citing similar experimental evidence as I did in the last chapter (see 6.2), St. John (1992: 274) contends that:

...prediction inferences are activated only weakly according to their support from the text...

Coherence inferences are a different matter and St. John (1992) cites the same evidence I did (i.e. Potts et al., 1988; see 6.2.2) in chapter 6 for the relative ease of activation of coherence (causal antecedent) inferences:

Coherence inferences, on the other hand, are fully activated and inferred (Potts et al., 1988; Singer, 1990). Following the same line of argument, the strength of coherence inferences results from their stronger support in the text.

Both of these positions are in tandem with the minimalist hypothesis. St. John's (1992: 274) conclusion from all this is as follows:

To simulate this interpretation of the empirical results, models of text comprehension should incorporate a mechanism that uses the degree of support provided by the text to determine the activation level of inferences.

And now to the model. The connectionist model is inputted with a short text consisting of

the following three sentences:

Albert and Clement decided to go to a restaurant.

The restaurant was expensive.

Clement paid the bill.

The model uses information from the training corpus to ‘answer’ ‘questions’ such as ‘who ordered?’, ‘who paid?’ etc (see below). For example, the model ‘infers’ that Clement will order cheap wine and that this restaurant is far away since, in the corpus, expensive restaurants are usually far away. Inferencing as to tip size is more complex. Its size depends on three factors: who tips (based on who paid), whether that person is generous or not and the restaurant quality since in the training corpus, tips were not left in cheap restaurants. As in his preamble, St. John (1992: 284) makes a distinction between coherence and prediction (forward-elaborative) inferences in the model, referring the reader to the table:

Coherence inferences concern propositions that lie between explicit text propositions. Prediction inferences concern propositions that lie after the final explicit text proposition. For example, **ordering** is a coherence inference because it occurs before the explicit proposition about **paying**. **Tipping** is a prediction inference because it occurs after **paying**. In the Story Gestalt model, all inferences are processed the same way. Any information correlated with the explicit text is activated immediately as the text is read.

I would agree with St. John (1992) that *ordering* by the *agents* is a (*causal antecedent*) coherence inference (see 6.2.3) since it ‘causes’ the paying of the bill. I also agree that *tipping* is a *forward-elaborative* or *predictive* inference since it elaborates on the text and so is not necessary for coherence. Having highlighted some of the St. John (1992) preamble, let us see in detail how his particular connectionist network captures the *shallow* generation of predictive inferences. Here is St. John’s (1992: 283) simulation of inference generation:

## Inference

Input Text					
Albert and Clement decided to go to a restaurant.					
The restaurant was expensive.					
Clement paid the bill.					
Questions					
DECIDED TO GO			ORDERED		
agent:	Clement	.9	agent:	Clement	.8
	Albert	.4		Albert	.1
	and	.9			
destination:	restaurant	.9	patient:	cheap wine	.6
				expen. wine	.1
QUALITY			PAID		
patient:	restaurant	.9	agent:	Clement	.9
value:	expensive	.9		Albert	.1
			patient:	bill	.9
DISTANCE			TIPPED		
patient:	restaurant	.9	agent:	Clement	.9
value:	far	.9		Albert	.1
			patient:	waiter	.9
			manner:	small	.4
				big	.1
				not	.0

The numbers indicate the probability activation of active concepts following questions relating to the headings in capitals. In DECIDED TO GO above, there are two agents, as indicated by the *and* unit, whereas elsewhere there is only one agent.

Let me now comment on the above results. When we inspect TIPPED, the lower probabilities for *manner* do reflect the fact that predictive (forward-elaborative) inferences should have more *shallow* activation levels than coherence inferences. In other words, this is in line with the psycholinguistic evidence of the last chapter. [However, the high activation of Clement as tipper is out of step with tipping as a predictive inference. The activation level of Clement as tipper would, all the same, be greater than for Albert, as

reflected in the model, since Clement ordered and paid the bill.]

We have seen that the prospect of instrument inference generation is proportional to how readily familiar encyclopaedic knowledge can be accessed and the degree of contextual constraints on this scenario. St. John and McClelland (1992: 100) are aware that instrument inferences are related to a high degree of contextual constraint:

Psychological evidence indicates that missing constituents, when strongly related to the action, are inferred and added to the description of the event. McKoon and Ratcliff (1981) found, for example, that 'hammer' was inferred after subjects read 'Bobby pounded the boards together with nails'.

However, absent from St. John and McClelland (1992), as well as St. John (1992), is the notion that instruments are more likely to be generated also when they rely more on *familiar encyclopaedic knowledge*. [Nor do they try to simulate the fact that elaborative causal antecedent inferences are stronger when the causal relation is stronger in memory, i.e. more *familiar*, as the experiments of Keenan et al. (1984) showed; see 6.2.3]. But the probabilistic nature of connectionist models could in principle capture this. Sanford and Garrod point to the possibilities of something like this direction (1994: 704):

The primary processing account assumes that a text is easy to read if it can be mapped onto familiar background knowledge structures and that the writer should try to make such mappings possible. It does not assume that all background knowledge representations will be equally rich or detailed (Sanford and Garrod, 1981: 125-31). Furthermore, with the advent of connectionist realizations of schemata, it is easy to envisage how various aspects of a scenario may be differentially accessible at various points in processing, which is not unlike McKoon and Ratcliff's [minimalist hypothesis] views on incomplete inferences.

If with the implementation of connectionist realisations of schemata, the richness of background knowledge can be made 'differentially accessible', this lends support to the idea that differential accessibility in a connectionist network could in principle be configured

according to the nature of the cognitive categories employed. In other words, i) basic-level categories would lead to a reader accessing of well-known knowledge and familiar scenarios, so as to generate implied instrument inferences from basic-level categories such as ‘nails’ as in the above and ii) superordinates would make the accessibility of well-known knowledge and familiar scenarios more difficult.

#### *7.4.4 General Corollary*

Connectionist models, because they deal in probabilities and because they deal in constraint satisfaction, can handle a NON-all-or-nothing approach to inference generation inherent within the minimalist hypothesis. A further corollary is that from a connectionist point of view, since the difference between coherence and predictive inferences is one of probability, predictive inferences being more *shallowly* generated than coherence inferences, it is quite unrealistic to see a clear dividing line between coherence and predictive inferences. As St. John and McClelland (1992: 131) indicate:

One process, activating the story based on constraints from the input, is used to produce a complete interpretation. Representing the explicit text, **drawing prediction inferences, drawing coherence inferences,** and resolving pronouns all result from this same process. [my bold]

All this ties in with Smolensky’s (1987) connectionist notion that the term inference is merely a higher-level description, a useful label that bears ‘only approximate relations to the underlying computational structure’ (cited in Clark, 1989: 111-2). Moreover, it is in concert with the psycholinguistic work of Vonk and Noordman (1990) (see 6.2.2) who show that the coherence (necessary) inference / predictive (forward-elaborative) inference distinction does not have a readily definable boundary, a notion which is supported by Sanford (1990: 521), the co-originator of the primary processing principle.

I have indicated the compatibilities between connectionism, cognitive linguistics and the psycholinguistic evidence for shallow processing of chapter 6. In chapter 8, I will use these compatibilities to enhance the IR framework for the highlighting of text which leads to mystification in reading, text which would not necessarily be highlighted in CDA. I now want to highlight tensions between connectionism and cognitive linguistics with regard to treatment of metaphor and draw out implications for CDA.

## **7.5 Tensions between Connectionism and Cognitive Linguistics: Implications for CDA**

### *7.5.1 Tensions between Connectionism and Cognitive Linguistics in Treatment of Concepts*

#### *Orientation*

As we have seen there are many similarities between connectionism and cognitive linguistics. Indeed, Lakoff (1987a: 338) is fairly confident of this [see 7.2.1]:

Our results...do not contradict what have come to be called 'connectionist' theories...

However, with regard to metaphor, I show how Lakoff *does* contradict connectionism.

#### *Problems for How Metaphor is Treated in Cognitive Linguistics*

Recall from chapter 1 how one of the salient premises of Lakoff and Johnson (1980) is that human conceptual systems, in terms of the ways in which we think and act, are at base metaphorical.<sup>3</sup> Metaphor for these authors is not regarded then as being ornamental or superficial. Here are Lakoff and Johnson (1980: 4) again (see chapter 3 (3.4.4)):

...let us start with the concept ARGUMENT and the conceptual metaphor ARGUMENT IS WAR. This metaphor is reflected in our everyday language by a wide variety of expressions:

#### ARGUMENT IS WAR

Your claims are *indefensible*.

He *attacked every weak point* in my argument.

His criticisms were *right on target*.

I *demolished* his argument.

I've never *won* an argument with him.

You disagree? Okay, *shoot!*

If you use that *strategy*, he'll *wipe you out*.

He *shot down* all of my arguments.

It is important to see that we don't just *talk* about arguments in terms of war. We can actually win or lose arguments. We see the person we are arguing with as an opponent. We attack his positions and we defend our own. We gain and lose ground. We plan and use strategies. If we find a position indefensible, we can abandon it and take a new line of attack. Many of the things we *do* in arguing are partially structured by the concept of war. Though there is no physical battle, there is a verbal battle, and the structure of an argument - attack, defense, counterattack, etc. - reflects this. It is in this sense that the ARGUMENT IS WAR metaphor is one that we live by in this culture; it structures the actions we perform in arguing.

We saw in chapter 3 that there is an underlying *classical* assumption of compositionality here that concepts are enduringly atomic, that they are not influenced by the *accommodative* effects of lexical company (see section 7.2.2). For example, 'win' in an alternative co-text such as 'I won some money in the lottery' has little to do with *war* (cf Lakoff and Johnson (1980) 'I've never won an argument with him', above). Lakoff and Johnson (1980) seem to regard the sense of 'win' as being steadfastly connected with war, not seeming to realise that its sense is affected by its lexical company. In effect what Lakoff and Johnson (1980) have done is establish that all of the lexical items in italics can be found in the semantic field of 'war' and import their martial meanings intact into the semantic field of arguments. But these lexical items could be a part of other semantic fields: 'film-making' or 'football' for 'shoot'; 'diseases' or 'bodily disorders' for 'attack'; 'computers' for 'wipe out' etc. Lakoff and Johnson (1980) do not consider these semantic fields perhaps because such



consideration would not suit the line they take.

Consider now the following from the connectionist philosopher, Clark (1989: 110-11) who, in referring to the connectionist network of McClelland and Kawamoto (1986) (see chapter 4), sees metaphor in a different way from Lakoff and Johnson, by relating it to the accommodative effects of lexical company:

- (1) The boy kicked the ball
- (2) The ball broke the window
- (3) He felt a ball in his stomach

Sentence (1) and (2) are from McClelland and Kawamoto (1986: 315). In case (1) we may imagine a soft, toy ball. In case (2) we imagine a hard ball (a tennis or cricket ball). In case (3) we have a metaphorical use: there is no ball in his stomach, but a feeling of a localized, hard lump. Everyday talk and comprehension is full of such shading effects according to overall context. **Surely we don't want to commit ourselves to predetermining all such uses in advance and setting up a special chunk for the semantic meaning of each.**

The PDP approach avoids such ontological excess by representing all these shades of meaning with various patterns in a single set of units representing microfeatures. The patterns for sentences (1) and (2) might share, e.g., the microfeature values spherical and game object, while the pattern for sentences (2) and (3) share the values small and hard. One interesting upshot here is the lack of any ultimate distinction between metaphorical and literal uses of language. There may be central uses of a word, and other uses may share less and less of the features of the central use. **But there would be no firm, God-given line between literal and metaphorical meanings;** the metaphorical cases would simply occupy far-flung corners of a semantic-state space. There would remain very real problems concerning how we latch on to just the *relevant* common features in understanding a metaphor. But it begins to look as if we might now avoid the kind of cognitivist model in which understanding metaphor is treated as the computation of a nonliteral meaning from a stored literal meaning according to high-level rules and heuristics. Metaphoric understanding, on the present model, is just a limiting case of the flexible, organic kind of understanding involved in normal sentence comprehension. [my bold]

On Clark's perspective, metaphor might be termed *extreme shading* or *microfeatural sparseness*. The advantage of the microfeatural perspective is that the meanings of words are not wholly predetermined - only certain microfeatures will be activated according to

context. However, in conflict with this connectionist microfeatural approach (see what Clark (1989) says in bold above), Lakoff and Johnson (1980) above predetermine the meanings of the words in italics by ascribing ‘whole’ semantic value to them. Clark (1989) is intellectually honest when he says that there would remain the very real problems for connectionists to explain how we latch on to just the *relevant* common features in understanding a metaphor. And as I have said before, I am not trying to claim that connectionism effortlessly offers explanations for all aspects of cognition. However, as Clark indicates above, connectionism conflicts with ‘the computation of a nonliteral meaning from a stored literal meaning according to high-level rules and heuristics’. Such a perspective relies upon concepts being seen as *atomic* and *enduring* in order that they can be manipulated according to set of rules. The cognitive linguistic programme may have dispensed with the notion of cognition as the rule-governed manipulation of symbols. But the cognitive linguists, Lakoff and Johnson, are still wedded, at least with regard to metaphor, to a classical assumption of symbols being atomic and enduring. Despite what Lakoff (1987a: 338) says, then, he does ‘contradict’ connectionism.

*The Idealised Reader and Lakoff and Johnson (1980)*

In chapter 6, (6.6), I introduced the notion of the idealised reader (IR), a non-analytical reader with little vested interest in a text and thus one who is parsimonious with their processing, generating only automatic inferences. In chapter 6, we saw that the low detection rates of lexical anomalies (e.g. ‘surviving dead’) in the experiments of Reder and Kusbit (1991) and Barton and Sanford (1993) suggest that it is *automatic* for the mental representation of a reader to be non-compositional with regard to lexical items in a text. This is due to top-down processing prevailing over the processing of discrete lexical items. I highlighted also that to *detect* anomalies such as ‘surviving dead’ each word had to be

fully analysed, i.e. compositionally. Compositional analysis, thus, requires effort and so compositional processing is *strategic* and not automatic. Related to this, in 7.4.1, we saw how the non-compositional processing of McClelland, Taraban and St. John's (1989) connectionist network and the capacity of McClelland and Kawamoto's (1986) network to shade meaning lends support for the fact that non-compositional language processing is *automatic* in humans.

Now, since IR's processing is *automatic*, non-compositional mental representation is, then, in line with this reader's minimum investment of cognitive effort. Let me return to Lakoff and Johnson's (1980) approach to metaphor. From the perspective of IR, to view the meanings of 'win', 'wipe out' and 'shoot' *compositionally* in relating them to the semantic field of war, regardless of the accommodative effects of their contexts, is hardly an *automatic* thing to do. It must, then, be that Lakoff and Johnson's (1980) analysis of the above sentences is *strategic* and so is not in line with the *automatic* non-compositional mental representation of a reader who does not invest much cognitive effort.<sup>4</sup>

### *Barsalou vs Lakoff*

In common with connectionism, Barsalou (1987; 1989) points out that how people represent a concept alters according to lexical environment. Earlier Anderson and Ortony (1975) had given an indication of this phenomenon in a series of experiments. They showed that when subjects are asked to memorise sentences like 'the man lifted the piano' and the 'man tuned the piano' that 'something heavy' is a superior recall cue to the former than 'something with a nice sound' and vice versa. Barsalou (1982) refers to this phenomenon as 'context-dependent information'. That is, only a portion of the knowledge associated with a particular category becomes activated according to the particular context. In 5.2.1, I

highlighted Barsalou's notion that many concepts are actually created 'on the fly' - ad-hoc categories - formed according to specific goals, as a function of context dependence, e.g. 'things to take with you in event of a fire'. I also highlighted how Lakoff (1987a: 45-6) regarded his work as being in line with that of Barsalou (1983). For Barsalou (1989), humans possess in long-term memory a large amount of loosely organised knowledge associated with a category. The connectionist philosopher Bechtel (1990: 266) sees Barsalou's perspective as compatible with that of connectionism:

It is somewhat difficult to make sense of this [Barsalou's] view within a rule-based account of cognition, since concepts would seem to be the atoms of such systems, but much easier to make sense of from a connectionist perspective where what exists in long-term memory are only connections. These enable the subject to produce representations that play the role of concepts and may be used in solving problems for which even rules might be invoked, but the concepts need not be fixed, atomic structures as they are in most rule-based accounts. Thus, implementing concepts in a connectionist system might allow us to explain in a straightforward manner some characteristics of concepts that might be otherwise difficult to explain.

Stillings, N.A. et al. (1995: 96-7) also makes the link between Barsalou's dynamic theory of concepts and connectionism. In the last section we saw that Lakoff and Johnson (1980) implicitly regard concepts as being enduringly atomic, seemingly underwritten by the classical assumption of compositionality. Despite Lakoff regarding his work as being in line with that of Barsalou, Barsalou's non-classical approach and compatibility with connectionism pose difficulties for Lakoff's view of metaphor.<sup>5</sup>

In 7.5.3, I tease out the implications of the above for CDA's use of Lakoff and Johnson (1980). While teasing out these implications in that section, I will also draw upon the distinction between *explanatory* and *constitutive* metaphors. This is what I address immediately below in 7.5.2.

### 7.5.2 Explanatory vs Constitutive Metaphors

One of the mistakes of Lakoff and Johnson's (1980) perspective is a conflation of *explanatory* and *constitutive* metaphors. An *explanatory* metaphor is one which assists the understanding without permeating that understanding. A *constitutive* metaphor is one where the metaphor does not explicate the theory but is part of the theory itself. The tendency in Lakoff and Johnson is to regard all metaphors as being constitutive. I derive the distinction between explanatory and constitutive metaphor from Ungerer and Schmid (1996: 147-9). However, the discussion that follows is my elaboration upon this distinction. On hearing 'you could cut the fog with a knife' we act on this material to realise the thickness of the fog but do not imagine that the fog was actually a solid. We have the experiential encyclopaedic knowledge that this is not true. This is, then, purely an *explanatory* metaphor to explicate the thickness of the fog. A *constitutive* metaphor is one where we are unable to readily access experiential encyclopaedic knowledge. Consequently it is unsurprising that many metaphors of science are 'constitutive of the theories they express rather than merely exegetical' (Boyd, 1993: 486; see also Kuhn, 1993; Schön, 1993 for similar perspectives). Kuhn gives the example of 'the atom as miniature solar system', the sun as nucleus and electrons as planets. This, however, conflicts with models of the atom inspired by work in quantum mechanics etc. As a more technological example, consider the example of the computer virus metaphor. Here are Ungerer and Schmid (1996: 147-8):

Though ordinary language users will not know much about the organism called virus, they may have a rich if indirect experience of its unpleasant effects on humans and animals...The metaphorical explanation may ultimately remain vague, but it seems to satisfy the conceptual needs of average computer users, so for them these metaphors do not fulfil an explanatory function but are constitutive for the conceptualization of computer malfunctioning.

Interestingly, what Ungerer and Schmid contend with regard to satisfaction of conceptual

needs ties in neatly with the notion of *shallow* processing and a reader's natural parsimony in processing if they have no specific goal (chapter 6). [For example, consider the generation of partial semantic value (Sanford, 1990), and Vonk and Noordman's (1990) experimental data where processing became shallow because the textual material was difficult to relate to background knowledge.] One more example of where constitution can occur is in the *philosophy of mind*. For similar reasons to the previous discussion, it is difficult to derive direct experience of the mind. Metaphors of mind such as the *mind as a computational device* as we saw in chapter 2 are more likely to lead to constitution for the non-specialist.

Now, consider constitutive metaphors from the perspective of Sanford and Garrod's primary processing principle. In the example of 'the mind as a computational device', if familiar knowledge cannot be engaged, the primary processing principle is *defeated*. Of course, for the specialist, familiar knowledge can be engaged. We can see, then, that there is in fact a *continuum* running from explanatory to constitutive which depends on whether comprehension is in accordance with the primary processing principle or not. A metaphor where there is a high degree of engagement from experiential knowledge will be more *explanatory* while a metaphor with a low degree of engagement of experiential knowledge will be more *constitutive*. When the primary processing principle is defeated, this also means that the accommodative (7.2.2) effects of lexical company are hindered. Consequently, in such cases, processing is likely to be to some extent compositional.<sup>6</sup>

### 7.5.3 Implications for How CDA Treat Metaphor

#### *Orientation*

In what follows, I consider how use of Lakoff and Johnson (1980) has led two CDA authors to regard metaphors much more in terms of *constitution* and not *explanation*. Secondly, I indicate how CDA examination of metaphor in certain texts does not accord with how the ‘non-resistant’ idealised reader (IR), referred to in 1.7 and 6.6 (i.e, one who makes minimum cognitive effort), would read these texts. In other words, I show how the CD analyst and IR would produce different discourses from the same text, and that therefore a CD analyst cannot necessarily be said to be *explaining* (interpretation-2) how a non-resistant reader, such as IR, is positioned by the text.

*Lee (1992: 90-92)*

Consider the following news text and the analysis by Lee (1992: 90-92), which I outlined in chapter 1, in the light of my discussion on metaphor above [Lee (1992) endorses Lakoff and Johnson’s (1980) view of metaphor]:

The black township of Soweto, which has been **simmering** with unrest since the riots on June 16 and the shooting of 174 Africans, **erupted** again today.

At least three Africans were shot dead, according to witnesses, although police deny this. The black hospital of Baragwanath nearby was reported to be ‘overcrowded’ with injured Africans.

The Minister of Justice, Mr Jimmy Kruger, announced in Pretoria this evening that he is reimposing the ban on public gatherings which lapsed last Saturday. The ban will continue until the end of the month. The nightmare of many whites in Johannesburg of a black march on their city almost came true today when between 20, 000 and 25, 000 angry Africans began moving in procession out of Soweto toward John Vorster Square, police headquarters in Johannesburg, where they planned to protest against the detention of black pupils.

Police with automatic rifles and in camouflage uniform headed the marchers off after they **swept through** a roadblock...

[my bold]

... metaphorical process that treats the people of Soweto as some kind of natural force, **specifically here as a volcano** which had been 'simmering' with unrest and then 'erupted'. This is echoed in the later report that the marchers had 'swept through' a roadblock, like a river. Note, too, that the emotions of individuals and the actions that they give rise to are transferred onto the place where they live. It is 'the township' that has been simmering and that now erupts, rather than the Sowetans experiencing feelings of anger and deciding to march. The effect of these processes of metaphor and metonymy is arguably to distance **the reader** from the subjects of the report. In speaking of the Sowetans as a natural force and as a place, the emotions of the people involved and the decisions which they make to engage in particular actions are eliminated from the process of interpretation. The situation is seen as resulting from some kind of inevitable set of natural laws rather than from human feelings and decisions. This tendency to downplay the agentive element in events initiated by relatively powerless groups is a general one. [my bold]

Although it is not mentioned explicitly, I take the 'reader' posited by Lee to be a non-resistant one, given the rationale of CDA to highlight how text can position non-resistant readers.

Consider what I highlighted in bold from the perspective of IR. 'Volcano' is an elaborative inference which could only be generated by the parsimonious processor, IR, if the semantic constraints were sufficiently rich. Recall from 6.2.4 the example of high semantic association from McKoon and Ratcliff (1989a) in the sentence from which 'sew' is inferred:

The housewife was learning to be a seamstress and needed practice so she got out the skirt she was making and threaded her needle.

From a cognitive linguistic perspective, 'sew' was readily generated because of the high degree of cognitive interdependence between 'sew' and 'needle' and the fact that the cognitive interdependence is predicated on motor-activity. On this basis, for *non-resistant* IR, who invests minimum cognitive effort, we can see, then, that 'volcano' would be a weak inference at best. Indeed, IR, possessing normal adult knowledge that people are not volcanos etc, would regard the metaphors, 'simmering', 'erupt' etc as *explanatory* metaphors



rather than being *constitutive*, *pace* Lee. So Lee has not actually indicated how a non-resistant reader, who invests minimum cognitive effort, processes the text. Rather, he has produced a *strategic inference* in order to suit *his* line of interpretation. That is, he rather ‘unshallowly’ imports the meanings of ‘erupt’<sup>7</sup> and ‘simmer’ *compositionally* from the semantic fields for ‘volcano’ and ‘natural forces’. But this is in fact rather arbitrary, since alternative semantic fields could easily be found for ‘simmer’ (cooking), ‘erupt’ (teeth; skin problems) etc. Crucially, though, it is not in line with the more shallow discourse derived by IR where the mental representation of lexical items in a sentence is *automatically* non-compositional.<sup>8</sup>

*Fairclough (1989: 120)*

Consider the following from Fairclough (1989: 120):

*As the cancer spreads [headline]*

As the riots of rampaging youths spread from the south, even the most optimistic have fears for the future, afraid worse is yet to come. How far can the trouble spread? If it comes to Scotland, where will it strike?

The metaphorical representation of social problems as **diseases** illustrated here is extremely common. Notice it incorporates a metaphor for **disease** itself, as a vague, subhuman and unthinking force (*where will it strike*).

The ideological significance of **disease** metaphors is that they tend to take dominant interests to be the interests of society as a whole, and construe expressions of non-dominant interests (strikes, demonstrations, ‘riots’) as undermining (the health of) the society *per se*. [my bold]

The first thing to notice is that Fairclough does not flag explicitly the notion of a non-resistant reader. But we may suppose that, for Fairclough, a non-resistant reader would read the text in the manner of *his explanation*, letting themselves be positioned by the text into drawing upon the macro-context script of ‘social problems as diseases’. Now, in the above, Fairclough performs a logical operation ‘cancer is a disease’ along the lines of  $\forall(x) [H(x)]$

→ S (x)] where H is a hyponym and S a superordinate.<sup>9</sup> As I indicated in 6.4, readers are usually shallow with regard to logical operations (Sanford, 1990; Wason, 1966) when they have little vested interest in a text. Since IR is one who has little vested interest in a text, we can suppose that IR would not generate the logical inference that Fairclough does. Also, since IR possesses normal adult experiential knowledge that countries are not bodies, IR (as well as any reader) is *unlikely* to treat ‘cancer’ as being a *constitutive* metaphor. We saw in 5.3.2 that metaphors can be functional in that they act as cues for us to realise functions. Since ‘cancer’ is unlikely to be constitutive to readers, that is, its *ideational* function does not predominate, its function is instead likely to be *interpersonal*, being eye-catching, having impact because of its *emotive connotations* etc. And this seems borne out by the position of ‘cancer’ in the most interpersonal aspect of news text - the headline. Fairclough, however, fails to see this since he does not distinguish between the headline and the text in terms of function, regarding both in ideational terms. For example, Fairclough regards the ‘it’ of ‘where will it strike?’ in the text body as referring to ‘cancer’ in the headline, when surely it refers to ‘trouble’ in the immediately previous sentence. Similarly in the Lee news text, words like ‘erupt’ etc are also presumably chosen for their ability to have connotative impact, i.e. their interpersonal function, when instead Lee regards them from an ideational perspective.

Fairclough continues thus (1989: 120) [in an analysis which is endorsed by Weber (1996: 7)]:

An alternative metaphor for the ‘riots’ might for instance be that of the argument – ‘riots’ as *vociferous protests* for example. Different metaphors imply different ways of dealing with things: one does not arrive at a negotiated settlement with cancer, though one might with an opponent in an argument. Cancer has to be eliminated, cut out.’

Recall from chapter 6 the psycholinguistic evidence that for a reader with no specific goal, causal consequent inferences are shallowly generated at best. So, while Fairclough generates the causal consequent ‘cancer has be eliminated’, non-analytical IR would not. In fact, Fairclough has produced a *strategic* inference in line with his own vested interest in the text, i.e., to boost the line he has adopted.<sup>10</sup> Such an inference, however, is not in line with the much less vested interest in a text by non-resistant IR. Thus, in this respect, Fairclough’s non-resistant reader and non-resistant IR produce different discourses.

#### *The Similarity between Lakoff and Johnson & Fairclough and Lee*

It should be apparent that Fairclough and Lee have much in common with Lakoff and Johnson (1980) in not emphasising the *micro*-context of interpretation and the reader’s potential for closure. Instead, both Lakoff and Johnson (1980) and the above authors operate at a *macro*-level of analysis where linguistic specificities are seen in terms of their membership of larger sets of phenomena. What I hope I have shown above is that these macro-phenomena are merely macro-*constructs*, constructed in line with the presumption in both CDA authors and Lakoff and Johnson (1980) that *macro*-phenomena allegedly guide our thinking and govern the micro-context of interpretation.<sup>11</sup> Having indicated how metaphors which CDA regard as constitutive are actually explanatory, finally let me use this distinction to highlight a scenario where CDA *is* correct to regard a metaphor as being constitutive.

#### *Prospects for Constitutive Metaphors in CDA*

Consider now the following from Fairclough (1992a: 208):

The vocational preparation product is usually a programme. Its design and implementation are therefore central parts of the marketing process, and should start from the needs of potential customers and clients and the benefits for which they are looking. (Further Education unit 1987: 51)

The message to course designers and teachers is a more elaborate variant of the marketing maxim 'Give the customers what they want'. Such wordings effect a metaphorical transfer of the vocabulary of commodities and markets into the educational order of discourse. But in contemporary Britain the metaphor is more than just a rhetorical flourish: it is a discursive dimension of an attempt to restructure the practices of education on a market model, which may have (as this extract suggests) tangible effects on the design and teaching of courses, the effort and money put into marketing, and so on.

Investigation of the *social world* involves the problematic investigation of the ontologically subjective while *physical world* investigation is an attempt to know the ontologically objective (Searle, 1995). The largely uniform nature of the ontological objective (at the macroscopic level) in the physical world means a reasonable likelihood of generalisation from experimental data. But the fluidity of the ontological subjective in the social world means that generalisation from one's experiential knowledge is not always secure. Because of this, the use of the 'market' metaphor, for example, may then defeat the primary processing principle, leading to constitutive conceptualisation for a non-analytical reader. Another way of seeing this is to say that for a non-analytical reader, constitutive metaphors might be said to give rise to a form of *shallow processing* since experiential knowledge cannot so readily be accessed.

### *Summing-Up*

- i) Fairclough / Lee and IR derive different discourses from the same text.
- ii) Different levels of cognitive investment in a text lead to different discourses being derived from the text. Above, the different discourses were the *non-shallow* discourse

generated by proxy for non-analytical readers by CD analysts and the shallow discourse generated by proxy for IR by myself.

iii) As in 6.6, CD analysts wrongly assume that the high cognitive effort they invest in interpretation by proxy for a non-resistant reader is replicated by such a reader who may have little vested interest in a text and so invest much less cognitive effort. By exploring the micro-context of interpretation (in contrast to CDA), I have shown that a non-analytical reader (i.e. IR) is *not* positioned by the metaphors in the above texts into making a particular interpretation as deemed by the CD analysts cited above. In other words, for the texts which Lee and Fairclough analyse in 7.5.3, (where readers are not coerced into interpretations), the macro-context does not necessarily govern the micro-context of interpretation of a non-analytical reader.

Again, as in 6.6, other things follow from the above:

- since different levels of vested interest and thus different levels of cognitive labour lead to different (by proxy) interpretations, the principle of ‘partiality of interpretation’ from the same text has thus been demonstrated.
- CDA (Fairclough and Lee) ‘over-interpret’ by proxy for the non-resistant reader. Indeed, both CDA authors as well as Lakoff and Johnson (1980) implicitly read macro-concepts into the micro-context of reading, without taking into account the nature of readers and their capacity for linguistic closure.
- since the discourse of *non-analytical* IR (i.e one in line with empirical psycholinguistic evidence) conflicts with the discourse of the CDA *non-analytical* reader (i.e. a reader not supported by empirical psycholinguistic evidence), the notion that there is critical

hermeneutic exegetic privilege in the above CDA explanations is considerably weakened.

In discussions of metaphor, the mistake CDA makes is to ignore the issue of richness of experiential knowledge and thus to see metaphor as always constitutive. But importantly, *pace* CDA, it is *not* metaphors which *construct* conceptualisation. Rather, metaphors are merely *cues* from which the *reader* constructs conceptualisation. If someone says, ‘I’ve been watching soap operas all evening and I’m feeling *mentally flabby*’ - the metaphor is not *mentally flabby*. Rather it is in the processing that the listener has to produce to *inhibit* the usual physical notion of flabbiness and leave other associations of flabbiness such as inactivity etc that could be associated with the mental domain. Conceptualisation, though, may become constitutive because the primary processing principle is defeated in cases where experiential encyclopaedic knowledge cannot be accessed. By using the distinction between explanatory and constitutive metaphors yoked to the primary processing principle, I have been able to separate an example of sound CDA from unsound CDA.

## 7.6 Endpoints

In chapter 6, we saw how psycholinguistic evidence for inference generation etc was incompatible with certain *symbolic / logical empiricist* text processing assumptions in CDA.

In chapters 4 and 5, I used connectionism and cognitive linguistics to *problematise* the symbolic postulates on which CDA was based. But we can now go further than ‘problematising’. With regard to psycholinguistic evidence for *shallow* processing, at least, cognitive linguistics and connectionism are much more compatible than symbolicism (the conflict between cognitive and connectionism on metaphor processing aside). The alternative framework I construct in the next chapter, for the highlighting of mystifying

discourse, from the compatibilities I have demonstrated above will, then, be a *non-symbolic* one. This framework will be a more developed and systematic version of the processing profile of IR. Partly because it is a non-symbolic framework, it will highlight text that leads to mystification in IR's reading, text which *symbolic* CDA would not necessarily detect.

## Notes

1. Another cognitive linguist, Lakoff (1987a: xiv), like Langacker, positions against the notion of cognitive 'building blocks':

'Thought has *gestalt properties* and is thus not atomistic; concepts have an overall structure that goes beyond merely putting together conceptual 'building blocks' by general rules'.

2. With more patent endorsements of connectionism, here is Sanford (1990: 527):

'In general, the ease of recognizing the classes of problems and situations which one knows (cf Sanford, 1987), compared to the slowness of overt logical inference, provides strong grounds for looking toward **parallel pattern matching** as a support for understanding rather than toward a uniform type of inference making.'

[my bold]

'The argument that there are two kinds of inferential mechanism, a fast pattern-matching facility and a facility corresponding to a slower, classical inference engine of understanding, is not new. In essence, the argument is basic to all schema theories of understanding (such as Schank and Abelson's 1977, script theory) and was put forward in a psychological context by Sanford and Garrod (1981). Indeed, the process finds its natural formulation at a **subsymbolic level, which studies of connectionism are beginning to make clear and which discourse psychologists would be unwise to resist.**'

[my bold]

Another link between Sanford and Garrod and connectionism is with regard to the nature of pronoun inferencing. Here are Sanford and Garrod (1994: 701) on Garrod and Sanford (1994):

'In the chapter by Garrod and Sanford (this volume), a discussion is given of the idea that reference resolution is really just part of a more general process of interpretation, and not prior to more general interpretation. This view is quite different from the ones discussed up to now which see coreference (argument repetition) as essentially prior to other sorts of interpretation...'

But this idea that coreference processing is coterminous with that of other sorts of interpretation is actually a facet of St. John and McClelland's (1992: 131) *connectionist* model (see chapter 4). Consider the following

which refers initially to the model's absence of distinction between elaborative and coherence inferences:

'The model, however, makes no distinction between these inferences. One process, activating the story based on constraints from the input, is used to produce a complete interpretation. **Representing the explicit text, drawing prediction inferences, drawing coherence inferences, and resolving pronouns all result from this same process.**' [my bold]

In both Garrod and Sanford (1994) and St. John and McClelland (1992), pronoun resolution is not regarded as separate but as just part of 'a more general process of interpretation'.

3. I confine discussion to Lakoff and Johnson (1980) and not more recent work on metaphor, e.g., Kittay (1987) and Gibbs (1994) since this is not drawn upon in CDA.

4. I should qualify all this by saying that only metaphoric processing for which top-down strategies can be engaged [i.e. for conventional metaphors] is automatically non-compositional. Novel metaphors such as in poems etc do not necessarily lead to immediate engagement with standard top-down strategies. In such cases, the 'literary metaphor' can draw the reader into analysis and thus into investing more cognitive effort. Processing, then, with 'literary metaphors' is more likely to be initially compositional (and thus bottom-up) compared with conventional metaphors. However, this is consonant with *strategic*, effortful processing and not automatic processing.

5. Another advantage of the connectionist approach over an 'atomic' view of concepts is that it explains Rosch's (1978: 40-41) *later* qualifications over prototype theory (see 5.2.1). Since prototypes 'emerge' on the Barsalou / connectionist perspective, this clearly twines with Rosch's admonitions that prototypes are 'fictions' and not 'stored representations'.

6. Consider the following from Kress's (1989a: 1), 'Linguistic Processes In Sociocultural Practice':

'I had tried to think of a title which did not separate 'language' from 'society' or 'culture', or language from its 'context', or talk about language and its social functions, or any of these formulations. The fact that I have not succeeded in my wish will be everywhere apparent in the book. Indeed, the very structure of the English language, with its preference for nominal, object-like forms rather than for the process-oriented forms of verbs, makes it a difficult and perhaps impossible task. So, quite often where I have used ungainly and awkward circumlocutions, that has been my reason: to try and invent ways of talking in which the linguistic and the social appear as one (though the last formulation signals yet another failure!).'

I would argue that Kress's 'failure' lies not with the propensity for object-like forms in English but with the difficulty of conceptualising 'language', 'society', 'culture', given the level of their abstraction from our direct *experiential knowledge*. One reason why these concepts are difficult to conceptualise is their *super*-basic-level nature, and so prototypes (i.e. familiar and 'graspable') are difficult to generate from them. Cognitive



interdependence exists between basic-level nouns and verbs. Thus, from a linguistic compositional input of basic-level nouns, a linguistic output can ensue which *goes beyond* the original compositional input. In contrast, there is little cognitive interdependence at the *super*-basic level. So, while ‘culture’, ‘society’ and ‘language’ may be linked together grammatically or morphologically (sociocultural), conceptualising them in unison, going beyond the original compositional input, is another thing all together. To sum up: we are thus most likely to understand a book title like ‘Linguistic Processes in Sociocultural Practice’ in a *compositional* manner unless we make the *considerable effort* not to.

7. Similar extrapolation from ‘eruptions’ to ‘natural disaster’ are expressed in Fairclough (1995a: 114).

8. The partiality of Lee’s analysis is visible. Lee (1992: 93) argues that ‘this tendency to downplay the agentive element in events initiated by relatively powerless groups is a general one’. But the ability of the ‘angry Africans’ to organise a huge march (20 000 - 25 000) and ‘plan’ a protest is in conflict with Lee’s supposition that agency is mitigated in the metaphorisation.

9. Sontag (1988) is endorsed in Fairclough (1992a: 197-8) and Shephard (1994). Consider the following from Sontag (1988: 28-9):

*‘Full-blown [AIDS] is the form in which the disease is inevitably fatal. As what is immature is destined to become mature, what buds to become full-blown (fledglings to become full-fledged) - the doctors’ botanical or zoological metaphor makes development or evolution into AIDS the norm, the rule.’*

Like Fairclough (1989: 120), Sontag performs a compositional analysis, extrapolating from the particular to more general categories, along the lines of  $\forall(x) [H(x) \rightarrow S(x)]$ , using the general categories of ‘botanical terms’ or ‘zoological terms’. But as I indicated in 6.4, logical relation inferences are *not* normally automatically produced (Sanford, 1990; Wason, 1966). To sum up, what Sontag produces is a strategic, *effortful* inference to suit her own line of interpretation and, thus, an inference which is not consonant with the result of automatic or minimum cognitive investment natural language processing.

10. Consider a perspective where cancer is removed via chemotherapy. In chemotherapy, the ‘whole’ of the patient suffers. If Fairclough had generated ‘chemotherapy’ in his strategic inference rather than the surgeon’s knife, then this would have had to imply the bizarre situation that the whole of the country would ‘suffer’ while the ‘cancer’ of the specific part of the country was being dealt with. Fairclough’s choice of ‘surgeon’s knife’ in his strategic inference, then, is *doubly* strategic since it supports the line he is taking more so than a strategic inference which alluded to ‘chemotherapy’.

11. Lee (1992: 71) is explicit about metaphoric processes working at a macro-level:

‘Some time ago, as I drove into an underground car park, my seven-year-old daughter asked me why the car radio had suddenly stopped working. My spontaneous (and probably somewhat unhelpful) reaction was to say

that it was because of 'radio shadow'...one domain of experience (that of radio transmission) is being structured in terms of a different domain of experience (the perception of light) through the use of language. The metaphorical process is operating at a more general level.

## **CHAPTER 8: AN ALTERNATIVE FRAMEWORK FOR THE ANALYSIS OF MYSTIFYING DISCOURSE PRODUCED BY A NON-ANALYTICAL READER**

### **8.1 Introduction**

In previous chapters, I have indicated many problems with how CDA highlights mystifying discourse. In this chapter, I construct an *alternative* framework from the compatibilities I showed between psycholinguistic evidence for shallow processing, connectionism and cognitive linguistics. The rest of this chapter will be given over to use of the framework to highlight how certain text can lead to mystifying discourse, though, on the whole, text which CDA would not detect as such. Firstly, to assist the reader, let me condense some of the arguments in previous chapters as to why CDA fails as a framework for highlighting mystifying discourse and thus why an alternative framework is necessary.

### **8.2 Why an Alternative Framework for the Analysis of Mystifying Discourse is Necessary**

#### *8.2.1 An Alternative Non-Analytical Reader*

In 1.4 and 1.5, I showed how there were several conflicting notions of a non-analytical reader in CDA and several conflicting assumptions of what constitutes the processing make-up of a non-analytical reader. In chapter 6, I showed how many of these assumptions were inconsistent with more recent psycholinguistic evidence, this evidence problematising how CDA highlights mystifying text. I also showed that because CDA has not given enough attention to the nature of the micro-context of interpretation, they regard the inferences they

make, as readers with a vested interest in a text, as being equivalent to those made by a non-analytical reader who may have little vested interest in a text. CDA's non-analytical reader, then, has not been adequately developed. There needs, then, to be an alternative framework for the highlighting of how certain text can lead to mystification in reading (lead to mystifying discourse) based on a *non-analytical reader who has a developed and consistent processing profile and in line with the psycholinguistic evidence I highlighted in chapter 6.*

Widdowson (1996) contains a critical examination of Fairclough's analysis of a text, which supplies information to pregnant women. He criticises Fairclough for 'interpretation by proxy' arguing that Fairclough should enquire as to the perlocutionary effect of the text on mothers rather than estimating what it may be. Analysts cannot be expected to be reliable in interpreting by proxy for discourse communities they do not belong to. This is why, for Widdowson, ethnographic studies should be performed on how texts are interpreted by members of discourse communities not inhabited by analysts. Naturally, this is the best procedure. It would of course be arduous, though, to conduct such ethnographic studies following every text analysis. Indeed, ethnographic study would not be so necessary if the analyst was actually a potential addressee of the text under scrutiny. I explained above why there needs to be an alternative framework for the analysis of mystifying discourse and that in contrast to CDA it should be based on a non-analytical reader who has a developed and consistent processing profile. What needs to be added, then, is that this is only for *text for which an analyst could be considered a potential addressee.*

### *8.2.2 An Alternative Vocabulary for Inference Generation*

We have seen how attitudes to inference generation in CDA are inconsistent with one another and that there is little awareness of inference typology in psycholinguistics. The

absence of any systematic vocabulary for inference generation in CDA based on psycholinguistic evidence means that their text analysis: i) neglects the standard inferences readers make, such as instantiations; ii) makes the mistaken assumption that causal antecedent inference constructed over adjacent clauses or sentences are necessarily weak inferences and thus mystifying of causality; iii) assumes that non-analytical readers invest the same cognitive effort as analysts, both generating strategic inferences. Furthermore, when inference theory *is* drawn upon in CDA (e.g. Fairclough, 1989), the age of the sources in discourse analysis (e.g. Brown and Yule, 1983) means important work on inference generation which has transpired in the late eighties and early nineties is absent. If inference generation is to be linked to mystification in reading, then, a greater level of consistency and awareness of different types of inferences are needed as well as a vocabulary for inference generation based on more recent psycholinguistic evidence.

### *8.2.3 The Incompatibility of Symbolic Postulates on Which CDA is Based with Psycholinguistic Evidence of Chapter 6*

When the psycholinguistic evidence of chapter 6 conflicts with the notion in CDA that inferences generated across clauses or sentences are weak representations, it also conflicts with the *symbolic / logical empiricist* basis of this that syntactic structure has priority in mental representation. We also saw in chapter 6 how for a reader investing minimum cognitive effort, mental representation of a sentence is automatically non-compositional. But this conflicts with the *symbolic* notion common in CDA that mental representation of sentences *is* compositional. In operating on symbolic postulates which *conflict* with psycholinguistic evidence, CDA has been using *non-performance* approaches to language processing. This is partly because such postulates have been absorbed from *non-performance* based theories of language in logical empiricism and early Chomsky.

Now, with regard to psycholinguistic evidence for *shallow* processing, we saw in chapter 7 that cognitive linguistics and connectionism are much more compatible than symbolism. For example, connectionist networks *automatically* shade meaning non-compositionally. For both connectionism and cognitive linguistics, linguistic input is not the same as process output, thus also dovetailing with the psycholinguistic evidence for automatic top-down processing. While CDA often draws upon *non-performance* approaches to language processing, the aspects of connectionism and cognitive linguistics I highlighted in chapter 4 and 5 are concerned with *performance*. In chapter 6, I outlined psycholinguistic evidence that inferences in reading are ubiquitous and varied, leading Sanford (1990: 515) to go as far as saying that studying text comprehension was like ‘a subset of the study of inference making’. Connectionism is in line with the prominence of inference generation in psycholinguistics since in connectionist networks inferences are not ancillary to processing (*pace* symbolism) but are an *inherent* aspect of language processing coterminous with the processing of syntax and semantics.

To sum up: a framework for the highlighting of mystifying discourse which is inspired by the psycholinguistic evidence of chapter 6, and one based on a non-analytical reader who makes minimum cognitive effort, is incompatible with the *symbolic* notion of compositional mental representation and also the downplaying of inference generation in favour of *symbolicism* inspired syntax-first strategies. However, such a framework *is* compatible with connectionism and cognitive linguistics.

### 8.3 An Alternative Framework for the Analysis of Mystifying Discourse

#### 8.3.1 Orientation

The alternative framework I offer for the analysis of mystifying discourse takes into account the recommendations of 8.2. The basis of the alternative framework is the *non-analytical* idealised reader (IR) I introduced in 6.6 and 7.5, a reader who has little vested interest in a text and is largely unfamiliar with its subject matter. In this chapter, I will build upon the processing profile of IR and also delineate this profile in a more concise and systematic manner than hitherto. Because IR is largely unfamiliar with the subject matter of a text, I concentrate on *news text* and in particular ‘hard-news’ stories<sup>1</sup>, which I assume IR would be unfamiliar with. I also choose news texts since they have no real specific addressee and so it would not be incongruous of me, the analyst, to interpret news texts ‘by proxy’ for IR. In other words, I too could be a possible addressee of ‘hard-news’ stories in newspaper texts.

Consider now the following from Brown and Yule (1983: 266):

While...it is, in principle, impossible to predict the *actual* inferences a reader will make in arriving at an interpretation of a text, we may be able to make predictions regarding particular aspects of individual texts which readers will generally have to interpret on the basis of inference. Such predictions will be closely related to some concept of ‘depth of processing’. Clearly, the reader who casually skims across the news article...while sitting in the dentist’s waiting room, is likely to be ‘reading’ the text in a qualitatively different way from the reader who is anticipating being asked comprehension questions after he has finished the text.

I concur with the argument of Brown and Yule that predictions about the actual inferences a reader makes ‘will be closely related to some concept of ‘depth of processing’’. I show that the cast of *inferences* IR generates from certain news text, in line with this non-analytical reader’s propensity for *shallow* processing, means subject matter is mystified for

IR. As I have said, my focus is on a reader with little vested interest in a text and one who is largely unfamiliar with subject matter. But this does not preclude the fact that different readers with different motivations will produce different discourses from the texts I analyse below. And indeed, at times I will indicate that the texts I examine below may not lead to mystification for readers who are *more* familiar with the subject matter and *have* a vested interest in the text. In other words, just like my analyses of CDA commentaries in chapters 6 and 7, the analyses below are in line with the ‘partiality of interpretation’ principle.

In the next section, I detail a processing profile for IR by listing compatible processing ‘principles’ from cognitive linguistics, connectionism and psycholinguistic evidence for shallow processing. These principles are the framework that guides my discourse analysis of IR. The psycholinguistic evidence I marshal for this framework has validity in the sense that it is largely *consensus* evidence in current psycholinguistics and is unobscure, being readily available in standard psycholinguistic and cognitive psychology reference books (Gernsbacher, 1994; Eysenck and Keane, 1995; Harley, 1995). Because it is, as I say, largely consensus evidence, it has a consistency that compares favourably with the rather inconsistent psycholinguistic assumptions of processing in CDA which were outlined in chapter 1.

Since IR is based on the compatibility between psycholinguistic evidence for *shallow* processing and by definition *non-symbolic* elements from cognitive linguistics and connectionism, use of IR necessarily *prohibits* modes of analysis which are predicated on *symbolicism*. Thus, I do not follow a CDA approach to analysis, e.g., I make no separation between syntactic meaning and semantic meaning. Consequently, my highlighting of text from which IR derives a mystifying discourse does not necessarily coincide with what CDA



highlights as being mystifying text, given their *symbolic* processing assumptions. This lack of coincidence also transpires because my framework deals more systematically than CDA on the issue of inference generation in text processing and its relation to mystifying discourse. All the same, my use of ‘principles’ of analysis mirrors that of CDA since (with the CDA authors I have discussed in this thesis, at least) a specific text comprehension model is not used. Rather CD analysis has transpired through a set of principles about how reading takes place, e.g. syntax-first strategies.

### *8.3.2 The Processing Profile for the Idealised Reader: The Framework Principles*

The following principles were demonstrated to be compatible in chapter 7. I separate them into their respective provenances below to ease the reader’s comprehension:

#### *Psycholinguistic Principles*

1) IR is a relatively parsimonious processor. Since IR has no specific reading goal, IR has little vested interest in a text. IR thus reads in line with the minimalist hypothesis since IR’s reading only involves *automatic* inferences. These include coherence inferences as well as elaborative inferences generated where information is quickly and easily available.

2) IR’s reading does not involve generation of *strategic* inferences (unlike CD analysts) since these involve time and work.

3) Instantiations of superordinate / general / abstract categories are a type of elaborative inference which are usually automatic for IR if relevant encyclopaedic knowledge is readily available.

4) Causal antecedent inferences are usually automatic. They can be divided as follows:

*connecting* - the reader establishes a link between the current focal event and prior information thus instituting a causal antecedent for the new event.

*reinstatement* - information from prior text which is currently not activated can be reactivated by the reader in order to establish a causal antecedent.

*elaborative* - the reader utilises background knowledge to establish a likely but unmentioned causal antecedent.

*Connecting* causal antecedent inferences are readily instituted by IR. But the strength of the *elaborative* causal antecedent inference is dependent on the strength of causal relation of the particular scenario in memory, e.g, *A punches B. The next day B is covered in bruises.* [inference - *the punching caused the bruises* is strong]; *A goes to a neighbour's house to play. The next day A is covered in bruises* [inference - *playing caused the bruises* is less strong].

5) Causal consequent inferences are usually *not* generated by IR or generated only weakly by IR when the situation is *not* highly constrained and very familiar, given the processing effort incurred in forecasting potential consequents.

6) For IR, causal *antecedent* inferences are usually much more probably generated than causal *consequent* inferences.

7) Topicalisation in a sentence will affect the generation likelihood of causal consequent inferences.

8) IR tries to relate textual material as soon as possible to background knowledge (primary processing principle).

9) IR *automatically* makes richer conceptualisation around main characters of narratives at the expense of secondary characters.

10) IR is a shallow processor of logical form if the context is unfamiliar. The logical offence of *affirming the consequent* can be expected in unfamiliar contexts, or in contexts in which the reader has little vested interest.

11) Lexical items from a sentence do not figure compositionally in IR's mental representation of a sentence.

12) Instrument inferences are likely to be generated by IR when knowledge is quickly and readily available such as when the instrument is part of a well-known *semantic field*, e.g., unlock (with a key), stab (with a knife), shoot (with a gun).

13) IR is largely unfamiliar with the subject matter of the textual material under analysis, but possesses otherwise 'normal' adult encyclopaedic knowledge.

### *Connectionist Principles*

1) linguistic input (sentential structure for example) is not equivalent to process output in IR's discourse. With this as a guiding line, I do not read off meaning from the application of a metalanguage. I also do not seek to find significance in sentential *form* (*pace* CDA) separate from other factors.

2) inferences in IR's discourse are inherent to processing, so syntax, semantic and inferential processing are interactional.

3) Unlike for some CDA authors, inferences are not weak representations *per se*. They can, however, be *weakly* generated, e.g. causal *consequent* inferences. Connectionist models capture the strength / weakness aspect of inference generation because of their capacity to capture probabilities.

4) Connectionist network output of a sentence is *automatically* non-compositional in line with IR's non-strategic processing. This does not preclude an analyst's compositional *examination* of a sentence, though. For instance, I analyse compositionally basic-level and superordinate categories in the texts that follow. But I do not make an *equation* between my compositional analysis of a category and compositional mental representation in the discourse IR produces.

### *Cognitive Linguistic Principles*

1) the basic level is the category level where the largest amount of information about an item is understood with the least cognitive labour - i.e., basic-level categories are characterised by *cognitive economy*.

2) basic-level categories are more directly understood with minimum cognitive effort since they lead to generation of prototypes, and thus the familiar, through being associated with our neurophysiological capacity for motor-interaction, gestalt perception and image generation. The converse of this is that a context where motor-interactivity etc is emphasised (though perhaps without use of basic-level categories) will also be 'directly understood'.

3) superordinate / general / abstract categories are *not* characterised by cognitive economy. To yield more information, they require *more* cognitive effort than with basic-level categories. On their own, or in context where there is an absence of sufficient basic-level categories or sufficient constraints, and when minimum cognitive effort is invested, they are unlikely to lead to *elaborative* inference generation.

4) prototypical causation involves 'direct manipulation' by an agent.

5) The framework prohibits the notion that use of a basic-level nominal produces *objectifying* effects since basic-level object categories and basic-level *action* categories are cognitively interdependent.

6) metaphoric constitution is more likely to transpire when experiential background knowledge cannot be accessed, i.e. when the primary processing principle is *defeated*.

7) linguistic input (surface structure) is not equivalent to process output. With this as a guiding line, I do not

read off meaning from application of a metalanguage.

8) syntax and semantics processing is interactional, i.e. 'syntax-first' approach prohibited.

I recognise that the vocabulary of the framework mediates between the text and my analyses. For instance, I use the terms 'coherence' and 'elaborative' even though there is no absolute distinction between these inferences. However, I retain the use of these terms for heuristic reasons, making it easier to articulate the probability of inference generation. I also continue to use the term 'inference' for heuristic reasons even though on a connectionist perspective, inferences are *inherent* to processing, being high-level descriptions, approximating underlying processing. The connectionist background to the framework suffers from being selective. It draws principally on work by McClelland and his collaborators. This selectivity is unavoidable given the constraints of what I am able to accomplish in the thesis. However, my choice of McClelland is significant in that he is one of the fathers of connectionist approaches to language processing. In previous chapters I have criticised CDA for their sentential focus. This was in relation to CDA asserting that a certain kind of semantic-syntactic form was the best representation of events, i.e., AGENT-PROCESS-AFFECTED. Like CDA, I too focus on sentences in my analysis. However, sentences are merely the *starting point* of my analysis. Where I differ from CDA is that I am interested in the *holistic* processing *output* of a sentence for IR (i.e. including inference generation). I follow the connectionist and cognitive linguistic principle that treats syntactic and semantic information as well as inferences interactively and so do not make an equation between just sentential form and mental representation (*pace* CDA).

One of the aims of this thesis has been to counter the rather impoverished appreciation of

cognitive matters and the inconsistent processing assumptions in CDA as well as to reveal underlying, unrecognised, notions of mental representation in CDA. Given this aim, and the constraints of the thesis, I have not been able to devote attention to the *socio-cultural* situation of readers. Also because I limit IR to having little vested interest in a text, my framework is limited in trying to reflect the range of interpersonal and intrapersonal variation in reading. IR is, all the same, *a part of* intrapersonal reading variation; I assume most readers, at some point, must have read news text non-analytically. To substantiate this point, here are McKoon and Ratcliff (1992: 440) whose minimalist hypothesis informs much of my framework:

For different readers, minimalist processing with little strategic processing will occur in different situations. For some readers, it might be a rare occurrence; for others, it might happen in such situations as reading a magazine on an airplane, reading the newspaper though the morning fog over breakfast, or reading texts in a psychological experiment.

Recall also from chapter 1 how Hodge and Kress (1993: 22) indicate the possibility of such ‘morning fog’ minimalist processing, in the case of those who would demur at the labour involved in de-transformation:

As readers of this editorial we should have to be alert and willing to engage in mental exercise to get beyond the seductive simplicity of the final form...where everything seems to be there on the surface...  
**...few commuters on the 8.05 from Brighton would have the energy to perform the mental gymnastics required.** [my bold]

Of course it could be argued that even for ‘morning fog’ readers some of the inferences generated are likely to be strategic inferences. That is, in reality it is *more likely* that a non-analytical reader produces fewer strategic inferences than automatic ones, and is *more likely* to have non-compositional than compositional mental representation. In this respect, then,

IR is indeed *idealised*. Nevertheless, basing such a reader on *only* automatic processing in 6.6 and 7.5 enabled us to see the tendency in CDA to regard interpretation by non-analytical readers as being equal to the higher level of cognitive work of the analyst. That is, the device of IR is useful in shedding light on the micro-context of interpretation of a non-analytical reader, something neglected in CDA. It is on this basis, then, that I use the device of IR in what follows in this chapter. One final point in this paragraph is that I prefer to use the phrase ‘non-analytical reader’ to describe IR rather than ‘lay-reader’. ‘Lay-reader’ is simply too large a category. Lay-readers can take up a variety of positions on a text; they can be non-analytical to varying degrees but they can also read analytically also to varying degrees (see: Wallace, 1992: 45-7; Widdowson, 1984a). Furthermore, an ‘off-duty’ CD analyst could read non-analytically.

As I have said, because of the constraints of my thesis, my focus on inferencing related to a *particular reader* and my aim to counter the paucity of appreciation of cognitive matters in CDA, my framework does not pay enough attention to socio-cultural considerations in reading. However, since IR is a reader with little vested interest in a text, their reading is not then *engaged* in the *social* or *political* sense, mitigating to some extent the absence of attention to socio-cultural considerations in this thesis. It could be argued, indeed, that my eschewing of a socio-cultural approach to reading, in favour of a cognitive one, has been justified. This is because I was able to show that the macro- or socio-cultural context does *not* in fact weigh heavily on the micro-context of interpretation, by a non-analytical reader of news text, as CDA supposes! So, CDA is an inappropriate mode of discourse analysis for the reading of news text. What I have said about the absence of socio-cultural consideration in IR applies also to my analyses in chapters 6 and 7 of CDA commentaries and their text data, via the perspective of IR.

Because my framework is overt about its cognitive and philosophical make-up, it is transparent why I think IR would process in one way as opposed to another and why I prohibit modes of processing such as an all-or-nothing approach to inference, compositional processing, whole semantic value etc.<sup>2</sup> That is, any disagreement with my analysis of the texts in this section is, I hope, much more likely to be with the framework rather than with any of my own personal attitudes which may inadvertently slip in. Of course, it is quite possible that I put ‘pressure’ on my ‘principles’ to fit my intuitively formed ideas about the text data below. There is *always* this danger. To try to overcome this, I try to link my analysis as explicitly as possible to the framework principles which I have laid bare, unlike CDA, which is permeated with unrecognised symbolic assumptions. In this way, in contrast to CDA, it should be more apparent if inadvertently I have smuggled in other assumptions ‘to suit a particular line’. I acknowledge, however, that my selection of texts for analysis may be governed by own predilections.

### *8.3.3 Reinforcing the Distinction between Discourse Analysis and Text Linguistics*

Consider the following from Widdowson (1997: 153):

Text does not signal its own meaning, so linguistic analysis, no matter how detailed, cannot result in understanding of how and why a text means what it does.

Indeed I would argue that the more detailed the linguistic analysis, the further one is likely to get from the significance of the text. And this follows because only some of the semantic meaning encoded in linguistic form is activated as contextually appropriate on a particular occasion.

IR (albeit with a focus on cognition rather than socio-cultural cognition) can be taken as a set of contextual conditions. Widdowson’s points above also concur with my highlighting of the different discourse derived by the detailed analysis, higher cognitive effort and thus strategic inferencing of the CD analyst in contrast to the ‘significance of the text’ for the

lower cognitive effort making, non-analytical IR. Furthermore, Widdowson's point that 'only some of the semantic meaning encoded in linguistic form is activated as contextually appropriate on a particular occasion' ties in with the psycholinguistic evidence we saw in chapter 6 for the partial activation of semantic value. Since only *some* semantic meaning is activated in the shallow reading of IR, reading off significance from a *text metalanguage* by proxy for IR is likely to misrepresent IR's discourse. This is because text metalanguage does not take account of how automatic processing is non-compositional. In relation to this point, we saw in 4.3.3 how Trew applied Hallidayan metalanguage to a text and proceeded to *read off* significance from this metalanguage without taking into account the micro-context of interpretation of a non-analytical reader. Trew, then, effectively *over-interpreted* the text for a non-analytical reader. A corollary of all this is that although some linguists (e.g. Stubbs, 1983) make no distinction between text linguistics and written discourse analysis, the distinction is actually a real and valid one. Text linguistics is concerned more with the formal properties of a text, with the whole semantic potential of text. It is analyst centred. (Micro-) discourse analysis is concerned with how different readers on different occasions read texts. It is (particular) reader centred. Since my focus below is on the relationship between a particular reader's mystifying discourse and inference generation, this explains why I do not devote much attention to a more text linguistic focus.

The rest of this chapter is devoted to showing how the IR framework is able to highlight text which leads to mystifying discourse for a non-analytical reader with little vested interest in a text and largely unfamiliar with its subject matter.



## 8.4 'Protest Mob Storm Tube HQ' News Text: Evening Standard 7 August 1996

### 8.4.1 *How the Text Leads to Mystification for IR as to the Nature of Causal Antecedence and Causal Consequence*

Consider the following text which began on page 1 of the Evening Standard 7 August 1996 and ran onto page 2. Page 1 was dominated by the report and the headline and sub-headline were in very large type.

#### *Evening Standard Text 1*

#### **PROTEST MOB STORM TUBE HQ**

Demonstrators fight past security men to occupy chief's office

by Luke Blair, Nick Pryer and Allan Ramsay

#### **PAGE ONE**

1. Protestors today smashed their way into London Underground headquarters **as a mass demonstration on the streets and the Tube strike brought central London to a standstill.**

*[CONFLATED CAUSAL CONSEQUENTS + FRONTING OF 'MASS DEMONSTRATION']*

2. An angry mob overpowered security guards at the LU head offices above St James's station and rushed up seven floors to march into the chairman's office.

3. There they threw papers around and unfurled banners from the windows declaring 'Don't squeeze the tube.'

*[ABSENCE OF PROTESTORS' PERCEPTION THAT LU MANAGEMENT'S RENEGING ON A PAY DEAL (CAUSAL ANTECEDENT) leading to an inevitable strike PRODUCES EXTRA TRAFFIC POLLUTION (CAUSAL CONSEQUENT), i.e. CARS LEAD TO POLLUTION, TUBES DO NOT]*

4. LU chairman Peter Ford said: 'I was astonished when they burst into my office.

5. I asked them what they wanted and they didn't seem too sure.

6. 'They started chanting slogans, throwing paperwork around and opening the windows and unfurling banners outside'.

7. 'I was worried that they would damage my family photographs so I asked them to respect them and they did.

8. I explained that I was a keen cyclist and asked them if there were any points I could clarify about the dispute.

9. 'I even told them how much Underground drivers earn.

10. Then they said they wanted to lock the doors but I told them if this was a non-violent protest they would have to play by the rules, which they did.'

11. Police were later questioning a number of the protestors.

12. At the same time, during the height of the rush hour, several hundred cyclists

**PAGE TWO**

from the same group of protestors - Reclaim the Streets - converged on Trafalgar Square.

13. The protest, ironically aimed at combating car use and supporting public transport,

*[ABSENCE OF PROTESTORS' PERCEPTION THAT LU MANAGEMENT'S RENEGING ON A PAY DEAL (CAUSAL ANTECEDENT) leading to an inevitable strike PRODUCES EXTRA TRAFFIC POLLUTION (CAUSAL CONSEQUENT), i.e. CARS LEAD TO POLLUTION, TUBES DO NOT] caused massive knock-on effects on routes already overloaded with the Tube strike traffic.*

*[PRESENCE OF AN ACTUAL CAUSAL CONSEQUENT]*

14. Scuffles broke out as anarchist cyclists confronted police and angry motorists.

15. They were greeted by shouts of abuse, fist waving, and a cacophony of tooting car horns from cabbies and drivers as they deliberately dismounted and remounted and pedalled slowly around the Square.

16. Edmund King, head of campaigns for the RAC, said: 'I'm appalled by the action of a bunch of anarchists stopping the traffic.

17. It's ironic that they are campaigning against pollution when they've caused 10 times more in Trafalgar Square by blocking the traffic.'

*[PRESENCE OF A PERCEIVED CAUSAL CONSEQUENT BY NON-SYMPATHIZER]*

18. After an hour the cyclists peeled off down Whitehall to Parliament Square where they again slowed traffic.

*[PRESENCE OF AN ACTUAL CAUSAL CONSEQUENT]*

19. Tempers rose to boiling point when several vanloads of police trapped the demonstrators in a bus lane.

20. A stand-off followed, with lines of officers blocking both ends of the street.

[I telephoned 'Reclaim the Streets' (RTS) and a spokesperson informed me that the purpose of the protest was three-pronged: i) to show solidarity with the tube strikers who were campaigning for a better pay deal and decided to strike after the London Underground (LU) management withdrew a pay offer; ii) to highlight how the extra pollution caused by more cars being on the road because of the tube strike was ultimately a result of LU management renegeing on a pay deal; iii) to support the tube as a form of public transport over polluting car culture and its others negative effects e.g. noise, danger for pedestrians].

*Establishing the Nature of the Protest*

From the processing profile of IR, we know that:

i) instantiation of general categories / superordinates is a fairly automatic process if it is dependent upon knowledge which is readily available.

ii) IR automatically searches for a causal antecedent (as in fact would any reader). Here is Graesser et al. (1994: 379):

Comprehenders attempt to explain *why* episodes in the text occur and *why* the author explicitly mentions particular information in the message. Thus, comprehension is typically guided by why-questions rather than other types of questions (e.g. what-happens-next, how, where, or when). There is extensive evidence that causal explanations of actions, events, and states play a central role in our understanding of narrative (Black and Bower, 1980; Bloom et al., 1990; Bower et al., 1979; Fletcher, 1986; Graesser, 1981; Rumelhart, 1975; Schank, 1986; Singer, 1990; Trabasso and Sperry, 1985; Trabasso et al., 1989; van den Broek, 1990)...

iii) IR tries to relate textual material to background knowledge as soon as possible (Sanford and Garrod's primary processing principle).

Now consider the first three sentences and my bolding in the following:

1. Protestors today smashed their way into London Underground headquarters as a mass demonstration on the streets and the **Tube strike** brought central London to a standstill.
2. An angry mob overpowered security guards at the LU head offices above St James's station and rushed up seven floors to march into the chairman's office.
3. There they threw papers around and unfurled banners from the windows declaring '**Don't squeeze the tube**'.

Initially there is no mention of the nature of the protestors. So, rather anomalously, we have mention of a general category before the instance. The text, it could be argued, is inconsiderate in that it places general categories before instances which as we saw in Garnham's (1981) experiments leads to a processing jolt and extra processing time (see 6.2.5). By sentence 3, there is *still* no mention of the nature of the protestors. There is though mention of 'tube strike' (1) and 'don't squeeze the tube' (3). As I have said, IR will

seek the nature of the causal antecedent, posing *why*-questions (Graesser et al., 1994: 379 above) in the discourse they derive from the text [cf the notion of ‘discourse as dialogue’ in Hoey (1983: 168-188); Widdowson (1984b: 39-53), and Cook (1994: 48-51)], constructing an *elaborative* causal antecedent inference in line with familiar encyclopaedic knowledge.

Now, in the absence of information as to the true nature of the protest on page one, and the combination of i), ii), iii) above, IR will quickly try to relate the text material to the familiar.

It is likely that IR will see the protestors as strikers, i.e., will instantiate ‘protestors’ and the other general categories as *striking London Underground employees*. In other words, IR generates an *elaborative* backward causal antecedent inference from background knowledge that strikers are usually angry, disaffected etc. Given that causal antecedent inferences can be generated readily, the exploitation of what is an automatic inference (i.e. IR initially setting up an incorrect causal antecedent) can be seen in a sense to be insidious.

There *is* however a picture of some cyclists in traffic with one car foregrounded and with the caption: ‘Demonstrating cyclists surround a motorist at Trafalgar Square today’. But there is no strong evidence on page one for IR to think that the ‘demonstrating cyclists’ in ‘Trafalgar Square’ are the same people as the ‘protestors’ in the LU management’s office other than the briefest hint when Peter Ford explained that he was ‘a keen cyclist’. But even this comes as late as line 8. Moreover, for IR to link these demonstrators with the protestors who invaded LU underground, and create a rationale for why ‘demonstrating cyclists’ would invade LU underground headquarters, is not in line with their processing labour profile. Indeed, as we saw with the ‘survivors’ text in chapter 6, global expectation can frustrate anomaly detection (i.e. mention of cyclists) at a local semantic level. As Sanford and Garrod (1994: 717) aver:

...processing at a local semantic level, where case assignment and other attachments take place, can be dominated by more global aspects of coherence establishment. In particular, if a text statement fits well with

a piece of pre-established knowledge or can be understood easily on the basis of pragmatics, then that link seems to be made, even if there are details of a local nature which are inconsistent with that interpretation.

We shall see in a later section on the processing of ‘characters’ in the text that ‘cyclists’ are backgrounded in another way in IR’s discourse.

For the nature of the protest to become clearer, IR has to turn to page two. But when IR turns the page, the reasons behind the protest are complex. The protestors are from ‘Reclaim the Streets’ and the protest is ‘aimed at combating car use and supporting public transport’ (13) but there is no *explicit* mention that the protestors are also supporting the tube strike. On page one in (3), there is of course the declaration on a banner ‘don’t squeeze the tube’. But previously for IR this was from the perspective of striking LU employees and so IR would have to incur processing labour to reactivate this information in the context of environmental protestors. This would most likely also involve the extra physical effort of turning back to page one.

#### *Revision of Interpretation through Text Bias*

I do not deny that IR will revise interpretation; (they may of course simply abandon the text). But the revision only takes place once IR moves from the front page (the text takes up almost all of it) and then onto the second page where the text layout is smaller. Moreover, when there is finally enough information for revision of the causal antecedents, the text immediately *bias*es against the ‘Reclaim The Streets’ (RTS) protestors. For example (12), there is topicalisation of ‘At the same time, during the height of the rush hour, several hundred cyclists...’ etc making the height of the rush-hour the focus rather than what RTS were trying to achieve. Other biases include: the protestors ‘caused massive knock-on

effects on routes already overloaded with the Tube strike traffic’ and descriptions of the cyclists as ‘anarchists’, the abuse they receive from drivers, cyclists going on to slow traffic etc. Edmund King signals his distaste where his comments reiterate the reporting: ‘anarchists’, ‘caused more pollution’, ‘irony’. So to sum up, IR can revise their interpretation but this revision occurs in the context of a great deal of text bias against the ‘protestors’. Since also it has taken until the second page to find out the nature of the ‘protestors’, the significance and reasons for the protest are considerably mystified and downplayed in the discourse IR derives from the text.

#### *The RTS (Perceived) Cause-Consequent*

Let us examine the issue of causal consequence in the text. The issue, it will become apparent, is a complex one. From the phone call I had with an RTS spokesperson, the perspective of RTS is the following: it is *London Underground (LU) management* who *caused* the extra traffic pollution since *they caused* the tube strike by renegeing on a pay deal, leading to a great increase in the number of commuters going to work by car; (see the spokesperson for RTS in ‘The Guardian’ article below for substantiation of this). However, this *perceived cause-consequent* is not explicitly mentioned in the text.

Consider now, sentences 13 and 17:

13. The protest, ironically aimed at combatting car use and supporting public transport, **caused massive knock-on effects on routes already overloaded with the Tube strike traffic.**

*[PRESENCE OF AN ACTUAL CAUSAL CONSEQUENT]*

17. It’s ironic that they are campaigning against pollution **when they’ve caused 10 times more in Trafalgar Square by blocking the traffic.**

*[PRESENCE OF A PERCEIVED CAUSAL CONSEQUENT BY NON-SYMPATHIZER]*

Here are clearly described causal consequents of the ‘Reclaim the Streets’ action. The first is not in dispute since it is indicated elsewhere, e.g. in a Guardian article, which we shall later examine. The second is a *perceived* cause-consequent by Edmund King, someone clearly out of sympathy with the protest. It could be argued, then, that this text is *biased against* the protestors since their perceived cause-consequent is *absent* while a perceived cause-consequent of a non-sympathizer is *present*.

*Text Presence Bias → Discourse Bias*

It should be clear now that there are different types of bias with regard to the Evening Standard text. On the one hand, there is bias which is *present in* the text. For example, the connotations of ‘mob’ and the commonplace derogatory sense of ‘anarchist’ [as opposed to the advocate of a particular political system ironically not based on chaos] are both negative ones and thus biased against the RTS protestors. These particular connotations are also consistent ones in news texts. Given the definition of *text* in 1.7.2, ‘linguistic forms in a stretch of language, and those interpretations of them which do not vary with context’ (Cook, 1994: 24), we can call the connotative bias in ‘anarchist’ and ‘mob’ examples of *text-presence* bias. Of course, if a particular reader’s processing concurs with the text presence bias against the protestors, then that reader’s *discourse* is *biased* also against the protestors.

*Text Absence Bias → Mystifying Discourse as Bias*

But there is another form of text bias that leads to a different form of discourse bias - one which is a by-product of *mystification*, which itself stems this time from *text absence*. While the perceived cause-consequent of a non-sympathizer (Edmund King) is *present* in the text,

*absent* in the text is the RTS protestors' *perceived cause-consequent* that it was LU management's renegeing on a pay deal that caused the extra traffic pollution. This *text absence*, as I indicated earlier, is a form of bias.

But what happens when we relate this text absence to a particular reader? As we saw in 6.2.4, causal consequents are not constructed when the reader has little vested interest in a text or is unfamiliar with subject matter. But they can be constructed when the reader *is* familiar with subject matter. So a reader *familiar* with the eco-protest (e.g., heard a radio report earlier in the day; has eco-protest sympathies etc) could produce in their discourse the necessary identification with RTS to generate the *RTS perceived cause-consequent*. And if the subject matter is familiar enough, then the causal consequent might be treated as being *automatic* (see 6.2.4). A second type of reader, one not as familiar with the subject matter *but* with a vested interest in the text nonetheless, could also eventually generate the RTS *perceived* cause-consequent if the trouble is taken to piece together certain aspects of the text, e.g., 'Tube strike brought central London to a standstill' (1); protestors breaking into LU headquarters (1); declaring 'don't squeeze the tube' (3); 'combatting car use and supporting public transport' (13); 'campaigning against pollution' (17) etc). In such a circumstance, the perceived causal consequent would be *strategically* generated rather than automatically. But despite the text absence bias, in both cases above the reader's *discourse* would not be *biased* against the protestors. Finally in this paragraph, related to the above, if a reader puts in the required cognitive work, it is possible to generate inferentially that RTS slowing the traffic (sentence 18) would lead to greater pollution, if it had not been mentioned in sentence 17 already (i.e. Edmund King's non-sympathetic perceived cause-consequent). But again this would be a *strategic* generation.



I have mentioned two different types of reader who are able to generate the RTS perceived cause-consequent. But what about IR? As I indicated in 8.3.2, given that IR has little vested interest in the text, and has no prior knowledge of the protest, they will not generate this perceived cause-consequent. It is an *absence* in IR's discourse. Moreover, not only would IR *not* generate this perceived cause-consequent in their *discourse*, but being a non-analytical reader, IR would *not notice* its absence either. Consequently, IR's discourse is not only *mystifying* as to the full rationale of RTS but also inadvertently *biased* against the protestors.<sup>3</sup>

*Why the Text-Based Principle of 'Reporting both Sides' needs to be Discourse-Based*

In disputes or protests where there are opposing sides such as in the above, it *seems reasonable* that a news text should report 'both sides of the story' so as to avoid bias. But this is only a *text-based* principle. Much of this thesis has tried to show that highlighting a text as being mystifying needs to take account of the *reader's contribution* to the text. We saw in 6.6, for example, that because CDA does not properly take account of a reader's contribution, they wrongly assume that an absence of *agent-process-patient* structures means causal antecedent inferences were only weak representations, and thus mystifying of causal relations. Furthermore, we saw above that despite its textual absence some readers (either with vested interest or familiar with subject matter) *could* generate the RTS perceived cause-consequent. In other words, while there is text absence bias, the discourse of such readers is *not* biased against the protestors. In the context of these readers, then, the principle of 'reporting both sides' so as to avoid bias does not seem so relevant.

But how does the text-based principle of 'reporting both sides' relate to IR? In contrast to

the above, we saw that IR, a reader with little vested interest in the text and one largely unfamiliar with its subject matter, would not generate the RTS perceived cause-consequent.

For IR, it is both a text absence *and* a discourse absence. And now to the crux of this section. On the basis of the above, I want to argue that:

*since:*

a) Edmund King's perceived cause-consequent is present in the text, but would not be generated as an inference by IR if it were absent, as this would involve strategic generation

*then:*

b) the RTS perceived cause-consequent should *also* be present since again this also would not be generated by IR, as again it would involve strategic generation.

In other words, when the particular reader is *IR*, the text really *should* report both sides.

To sum up: in the case of (*perceived*) *cause-consequents* absent from a text, the principle of 'reporting both sides' has only real significance in the context of a *particular reader*.

This reader should not be one who has a vested interest in the text and is then willing to invest a high degree of cognitive labour. Neither should this reader be one who is familiar with the subject matter of the text. Rather this reader should be one who makes minimum cognitive effort and one largely unfamiliar with a text's subject matter (i.e. IR). *What I have done, then, is to adjust the text-based principle of 'reporting both sides' into a discourse-based principle. The value-judgement that texts should 'report both sides' to avoid bias has thus been grounded in a processing criterion, i.e. one which relates to minimum processing effort.*

*Using I.R as a Yardstick for the Success of a Text*

From what I have argued above, the IR framework can be used as a yardstick to gauge the *success* of news texts in discourse with regard to the issue of cause-consequents, and can also be used to improve the text. Below I offer an improved alternative fragment which removes the prospect of IR mystifying discourse bias related to text absences:

**ECO-PROTEST MOB STORM TUBE HQ**

*Cycle* Demonstrators fight past security men to occupy chief's office

1. *Eco*-protestors today smashed their way into London Underground headquarters as the tube strike and a mass *cycling* demonstration brought central London traffic to a standstill.
2. An angry mob of *cyclists* *dismounted* and overpowered security guards at the LU head offices above St James's station and rushed up seven floors to march into the chairman's office.
3. There they threw papers around and unfurled banners from the windows declaring 'Don't squeeze the tube'.

Removing the potential for text absences to lead to IR mystifying discourse bias does not mean that text presence biases ('mob', 'anarchists' below etc) are also removed. The IR framework is based on a minimum processing effort criterion that aims to remove the prospect of generating information strategically, i.e. information that takes a good deal of cognitive effort. So since 'mob', 'anarchists' etc are biases *present in* the text, they take *much less* cognitive effort to detect than to locate text absence bias; this is why I do not remove *text presence biases* in my alternative text fragment. Furthermore, I have not removed what CDA would regard as 'nominalisations' - 'protestors' and 'demonstration' either. But the addition of 'eco-' and 'cycle' etc, which do not take up that much space, mean there is enough information this time for IR to *instantiate* 'protestors' and 'demonstration' in discourse. In other words (*pace* a CDA type analysis), in the original Evening Standard text, it was not the syntactic form of 'protestors' and 'demonstration'

which was the obscuring factor. Rather mystification transpired because of their abstract / general nature, reducing the accessing of appropriate encyclopaedic knowledge, as well as the absence of appropriate information to instantiate these categories. With the alternative text, then, mystification as to causal antecedence is reduced in the discourse of IR. But what about causal consequence? The causal consequent of the strike that the protestors are trying to highlight (see Guardian text below for corroboration) might go something like the following:

4. A spokesperson said the protest was in support of the tube drivers' strike. It aimed to make LU bosses aware that reneging on a pay deal was causing greater traffic pollution with thousands of tube-commuters taking their cars into central London today.

I include 4 since i) explicit mention of supporting the tube strike means this avoids the cognitive labour involved in generating this rationale from 'don't squeeze the tube'; ii) IR would not produce the RTS perceived cause-consequent as it would involve identification and thus more than minimum cognitive labour. I keep the causal consequent in 5 since it is part of the event.

5. The mass-cycling protest throughout central London, however, caused massive knock-on effects on routes already overloaded with the Tube strike traffic.

6. Edmund King, head of campaigns for the RAC, said: 'I'm appalled by the action of a bunch of anarchists stopping the traffic.'

7. It's ironic that they are campaigning against pollution when they've caused 10 times more in Trafalgar Square by blocking the traffic.'

'Anarchists' in 6 is an example of text presence bias and I keep this for the same reasons as above. I keep 7 since it is also a *perceived cause-consequent* which, again, IR would not generate, given the effort involved with identifying with the perception. [We can imagine

that in an eco-protest-friendly newspaper that such a casual consequence might be omitted.

In these circumstances, again it would not be generated by IR.] Finally, consider the following from Sanford and Garrod (1994: 704):

The primary processing account assumes that a text is easy to read if it can be mapped onto familiar background knowledge structures and that the writer should try to make such mappings possible.

Since the nature of the protest is not familiar enough for IR to generate the RTS perceived cause-consequent, The Evening Standard text should have made available the rationale of the protest in the first place so IR would have this as ‘familiar background knowledge’.

#### *8.4.2 Contrast with Other ‘Reclaim the Streets’ Texts*

Firstly, consider the following from the same edition of the ‘Evening Standard’ (Wednesday 7 August 1996) but this time from page 2 and by one of the same authors of the first piece. Compared to the first text, text 2 is in smaller type, is only one of five stories on the page and so does not take up the whole page:

##### *Evening Standard Text 2*

CRAWL TO WORK IS THE WORST YET

by Luke Blair

Industrial Correspondent

1. LONDON today suffered its worst traffic congestion of the summer as the Tube strike and the mass protest by cyclists caused widespread chaos.
2. At the height of the morning rush hour, parts of central London came to a complete standstill as the protest, ironically aimed at supporting public transport, caused massive knock-on effects on routes already overloaded with Tube strike traffic.
3. Roads around Westminster, including some of the busiest in London, were blocked solid, according to motoring organisations.

4. Emergency services said the public were doing their best to let them through, but a spokesman for the ambulance service added: 'Obviously it gets difficult in a gridlock situation.'

*[In 1 to 4 PRESENCE OF NEGATIVE CAUSAL CONSEQUENTS OF THE PROTEST]*

5. A spokesman for the protest group, known only as Chris, said: 'The idea was not to block traffic but to get motorists to leave their cars at home.'

6. He said the group supported both the striking Tube drivers and the idea of public transport.

7. A London Underground spokesman, however rubbished the protest as 'totally self-defeating'.

8. He said: 'they are not helping the cause of public transport at all. We need to keep the roads clear on today of all days, if only so that our buses can get through...' [my bold]

Sentence 1 indicates that the protest is by cyclists and so allows some instantiation of 'protest' as to the nature of the causal antecedent fairly early on in the text's reading, unlike Evening Standard Text 1. Sentence 2 indicates that the protest 'aimed at supporting public transport', again allowing more instantiation. However, the impact of this is reduced since this text is much less salient than Evening Standard text 1. Moreover, there is *still* no explicitly mentioned *perceived cause-consequent* by the protestors that because LU management reneged on a pay deal, there is more traffic and thus more pollution on the roads. Admittedly there is a hint of all this in 5 and 6 but IR would not invest the necessary cognitive effort to 'expand' these hints. Besides, this hint does not compare favourably to 1 - 4 and 8 where there is explicit specifying of *negative* causal consequents of the cycle protest. In this respect, the text still leads to mystifying discourse for IR about the rationale of the protest.

*The Guardian August 8 1996*

**Motorists fume over show of support for strikers**

**CYCLE PROTEST ADDS TO TUBE DISRUPTION**

**Alex Bellos**

1. Bicycle campaigners ended a rally that brought chaos to London's rush hour traffic yesterday by invading

the office of London Transport's chairman and using their bicycle locks to chain themselves to his window.

2. Forty protestors entered LT's headquarters in Westminster after 500 cyclists had brought traffic around Trafalgar and Parliament squares to a standstill.

**3. The action was in support of yesterday's strike by tube drivers.**

*[ABSENCE OF PROTESTORS' PERCEPTION THAT LU MANAGEMENT'S RENEGING ON A PAY DEAL (CAUSAL ANTECEDENT) leading to an inevitable strike PRODUCES EXTRA TRAFFIC POLLUTION (CAUSAL CONSEQUENT), i.e. CARS LEAD TO POLLUTION, TUBES DO NOT]*

4. An LT spokeswoman said that the chairman, Peter Ford, was in his seventh floor office when about 10 cyclists came in, started throwing papers around, and hung a banner out the window.

5. They spent 10 minutes having a 'fairly good-humoured chat', in which they all agreed public transport was under-funded, until the police arrived with bolt cutters.

6. Nine protestors were arrested and charged with offences including assault, breach of the peace, criminal damage and theft.

7. Scotland Yard said that protestors punched and kicked a police sergeant.

8. They also grabbed his video camera, police radio and helmet.

9. The event was organised by the anti-car pressure group Reclaim the Streets, which has close links to London's monthly Critical Mass bicycle demonstrations.

10. While participants said the aim was not to infuriate car drivers, an inevitable consequence was that traffic - already slower than normal because of the tube strike - was brought to a standstill in several places.

*[SEPARATION OF CAUSAL CONSEQUENTS OF TUBE STRIKE AND PROTEST]*

11. Many drivers honked their horns in anger and threatened violence at the cheering protestors.

12. Groups of cyclists converged on Trafalgar Square at 9am, where a man in his 40s was arrested for breach of the peace.

13. Reg Wagland, aged 62, a retired electrician, said the protest was to highlight transport problems and show solidarity. 'We are here to support the tube drivers. Their employers reneged on a deal.'

*[ABSENCE OF PROTESTORS' PERCEPTION THAT LU MANAGEMENT'S RENEGING ON A PAY DEAL (CAUSAL ANTECEDENT) leading to an inevitable strike PRODUCES EXTRA TRAFFIC POLLUTION (CAUSAL CONSEQUENT), i.e. CARS LEAD TO POLLUTION, TUBES DO NOT]*

14. Another cyclist said: 'By cycling instead of driving Londoners can show their concern about the chronic underfunding of public transport, which results from placing profits before the environment, the health and safety of workers and quality of life.'

15. Bearing banners saying Squeeze Cars Not Tubes, the demonstration moved on to Parliament Square.

*[ABSENCE OF PROTESTORS' PERCEPTION THAT LU MANAGEMENT'S RENEGING ON A PAY DEAL (CAUSAL ANTECEDENT) leading to an inevitable strike PRODUCES EXTRA TRAFFIC POLLUTION (CAUSAL CONSEQUENT), i.e. CARS LEAD TO POLLUTION, TUBES DO NOT]*

16. In front of the House of Commons the protestors got out of their saddles and held their bikes in the air.

17. The demonstration was criticised by motorists' organisations.

18. The RAC's head of campaigns, Edmund King, said: 'To stage this protests on a strike day when many

people have no alternative but to use their cars is selfish and counter-productive.’

19. An AA spokesman, Paul Watters, said: ‘Direct action protests really do cause chaos, and some of the more illegal activities can cause disruption to the emergency services and people having to make urgent journeys.’

[*PRESENCE OF A PERCEIVED CAUSAL CONSEQUENT OF PROTEST BY NON-SYMPATHIZER*] [my bold]

Like the previous texts, The Guardian text does not attempt to conceal the fact that the protest caused huge disruption. However, the crucial point is that IR’s discourse is less mystifying, with regard to causal antecedence in the early part of the text, than the one derived from Evening Standard Text 1. This highlights how the nature of the ‘protestors’ is significantly downplayed in ‘Evening Standard’ text 1. IR knows it is a ‘cycle protest’ from the start as well as knowing explicitly that the cyclists are supporting the tube strike. Since posing *why-questions* as to causal antecedence is a natural part of discourse derivation even for a reader with little vested interest in a text, a rationale can begin to be formulated that this is some kind of ecological protest. However, again, the Guardian text does not specify the *RTS perceived causal consequent* at sentence 3, i.e., that LU management from the perspective of the bicycle campaigners have caused the extra traffic pollution and nor does it seek to flesh this out at sentence 13. The mention of the banner ‘squeeze cars not tubes’ (sentence 15) again hints at this consequence. But a fuller generation of the consequence would not be forthcoming from IR since again it is a consequent based on the perception of the campaigners, requiring more processing effort in order to produce the necessary identification. Compare also sentence 10 in The Guardian with sentence 1 from the Evening Standard text 1:

*The Guardian*

10. While participants said the aim was not to infuriate car drivers, an inevitable consequence was that traffic - already slower than normal because of the tube strike - was brought to a standstill in several places.

[*SEPARATION OF CAUSAL CONSEQUENTS OF TUBE STRIKE AND PROTEST*]



*Evening Standard Text 1*

1. Protestors today smashed their way into London Underground headquarters **as a mass demonstration on the streets and the Tube strike brought central London to a standstill.**

*[CONFLATED CAUSAL CONSEQUENTS + FRONTING OF 'MASS DEMONSTRATION']*

In the Guardian, there is a separation of the causal consequents of the tube strike and the protest. That is, the tube strike meant that the traffic was slower than normal and the protest meant that the traffic was brought to standstill in places. In Evening Standard text 1, these consequents are not separated. Indeed, with the fronting of 'mass demonstration on the streets', the impression created for me is that the demonstration is an equal contributor to the central London standstill if not the major contributor.

*8.4.3 Mystification through Automatic Selective Process around Main Characters*

Recall the information as to selective processing of narratives we looked at in 6.5. We saw how a number of studies indicated that *main* and *subsidiary* characters are processed in different ways. We saw how Morrow (1985) showed that when a character (B) is perceived from the point of view of character (A) that the reader's perception of states is bound up with the perspective of character (A). In effect, (A) is the main character and (B) is the subsidiary character. We also saw how Garrod and Sanford (1988) examined the differences in reference resolution with both proper names and role descriptions and that sentences with references to proper names were read more rapidly than those with references to role-defined characters. Both these experiments demonstrated that causal explanation of main characters is *automatically* sought. IR only generates automatic and not strategic inferences and so will only *automatically* seek causal explanation of main characters and not seek *strategic* explanation of secondary characters which requires more effort. With these points in mind, consider the 'character' of Peter Ford in the following seven sentences from Evening

Standard text 1 and the discourse IR derives<sup>4</sup>:

4. LU chairman Peter Ford said: 'I was astonished when they burst into my office.
5. I asked them what they wanted and they didn't seem too sure.
6. 'They started chanting slogans, throwing paperwork around and opening the windows and unfurling banners outside'
7. 'I was worried that they would damage my family photographs so I asked them to respect them and they did.
8. I explained that I was a keen cyclist and asked them if there were any points I could clarify about the dispute.
9. 'I even told them how much Underground drivers earn.
10. Then they said they wanted to lock the doors but I told them if this was a non-violent protest they would have to play by the rules, which they did.'

The fact that sentence 6 is in fact a repetition of sentence 3 is an example of *text bias* furthering the case that the real nature of the protest is being delayed. And now to the discourse bias. In *seven* sentences, the protestors are viewed from Peter Ford's perspective thus imparting to them subsidiary status as 'characters'. Furthermore, 'protestors' are 'role-defined characters' whereas Peter Ford is a proper name. Given that IR would seek automatic causal explanation of the main character, IR's discourse would inadvertently downplay the perspective of the 'protestors', leading to even more mystification as to the purpose of the protestors. In other words, again, IR's *discourse* is inadvertently *biased* against the protestors. But what about a reader more familiar with the nature of the protest and so one who more easily relates the picture of 'demonstrators' to 'protestors' in the text? For such a reader, the 'secondary' billing of the protestors will *not* lead to such a mystifying discourse of the purpose of the protest and so this reader's discourse would not be biased in the same way as IR's discourse. By the same token, a reader with a more vested interest in the text can seek causal explanation for the actions of the secondary characters. However, this would be a strategic procedure rather than being an automatic one.

A word or two about apparent anomalies. Consider sentence 9 'I even told them how much

LU drivers earn'. On page one where sentence 9 is found, IR would still not know that the protestors are eco-protestors and is highly likely to assume that the protestors are LU employees, perhaps tube drivers. On this assumption, Peter Ford's utterance would seem odd. Peter Ford also explains that he was a keen cyclist (8), which also seems anomalous. However, both these anomalies are from Peter Ford's perspective, the main character perspective, and so the oddity of 'cyclist' is somewhat diminished in IR's discourse. In the same way, in sentences 16 and 17, Edmund King is given more primary character status in being allowed to voice his opinion of the protestors who in consequence receive the 'billing' of secondary characters. Again, as secondary characters, the nature of the protest is diminished in IR's discourse.

And a few final points concerning the Guardian text:

4. An LT spokeswoman said that the chairman, Peter Ford, was in his seventh floor office when about 10 cyclists came in, started throwing papers around, and hung a banner out the window.
5. They spent 10 minutes having a 'fairly good-humoured chat', in which they all agreed public transport was under-funded, until the police arrived with bolt cutters.

With regard to Peter Ford, the perspective has become 3rd person and not 1st person. Thus the protestors are not perceived from the viewpoint of a main character. The configuration of the Guardian text means that the nature of the protest is not as downplayed in IR's discourse as compared to IR's discourse from the 'Evening Standard' (text 1).

## **8.5 News in Brief Text - 'Sex Assault Sailors Jailed': The Mirror October 8 1997**

### *8.5.1 Instantiation of the Text Body by the Headline*

Compare the following two ‘News In Brief’ texts:

i) **INDECENT ASSAULT SAILORS JAILED**

SIX Navy ratings were kicked out of the service and jailed for up to 10 months yesterday after two recruits were *indecently assaulted* in an initiation ceremony. [*CAUSAL CONSEQUENT PRESENT*]

Their young victims were attacked on board HMS Southampton. Two other ratings were given three-month suspended sentences by a Portsmouth court martial.

ii) THE MIRROR October 8, 1997:

**SEX ASSAULT SAILORS JAILED**

SIX Navy ratings were kicked out of the service and jailed for up to 10 months yesterday after two recruits were *indecently assaulted* in an initiation ceremony. [*CAUSAL CONSEQUENT PRESENT*]

Their young victims were attacked on board HMS Southampton. Two other ratings were given three-month suspended sentences by a Portsmouth court martial. [*my italics*]

Only the headlines in the texts differ, with the second headline being authentic. I will comment initially on text i). While ‘indecent assault’<sup>5</sup> may suffice in the headline, we would expect to be supplied with details in the body of the text. But this is not the case. Notice how the superordinate verb ‘assault’ is used in both the headline and the text body. It is difficult to form prototypes around superordinates and thus difficult to access the familiar. Superordinates, unlike basic-level categories, are not characterised by cognitive economy – a high amount of information yielded from a minimum cognitive effort. IR, of course, does not invest much cognitive effort and so processing of the text is shallow with regard to the possibilities for what type of assault took place. Text i) is mystifying in reading for IR.

Now consider text ii), which is the actual text. The headline is SEX ASSAULT SAILORS JAILED. Before I continue with the analysis here are McKoon and Ratcliff (1989b: 1143):

...when a [general] category is contextually defined, then information about a most typical exemplar of the category is encoded to a high degree of inference processes...these processes can be claimed to apply only for contextually defined categories that have a most typical exemplar...Other categories may have so little definition that there is no most typical exemplar (e.g. *Dorothy thought about the animal*, from Roth and Shoben, 1983).

‘Indecent assault’ on McKoon and Ratcliff’s line would not lead to a ‘typical exemplar’ for IR, in the same way that ‘Dorothy thought about the animal’ would not. But what of ‘sexual assault’? I would not like to claim that ‘sexual assault’ would lead to a sharply delineated typical exemplar such as penile penetration. But all the same, it is more *constraining* than ‘indecent assault’. That is, the modifier ‘sex’ (which could be construed as a basic-level event, given that ‘rape’ is a basic-level event (see, Rifkin, 1985)) leads to some degree of instantiation of ‘assault’ (or to use Rosch’s (1978) phrase, leads to some ‘translation’ of the more general category) such that the assault much more obviously included penile sexual contact, which characterises ‘rape’.<sup>6</sup> Because ‘sex assault’ is more constraining than ‘indecent assault’, and a prototype is partially emergent at least, ‘sex assault’ has some association with ‘direct manipulation’. Crucially though, this ‘translation’ of ‘assault’ in the headline does not require much cognitive effort since basic-level categories are characterised by cognitive economy. Thus, the above ‘translation’ is one readily produced by readers and not just by IR.

Now, as the interpersonal function presides in a headline, we would not expect the ideational function of this headline to be the ultimate guide for how the text is processed. However, ‘indecently assaulted’ describes the nature of the events in the *body* of text ii), whereas ‘sex assault’ is in the headline. Furthermore, there is no mention of the actual details of the assault in the text body. This is to be expected since it is a ‘news-in-brief’ text. Consequently, it is likely for IR that the general category in the body of the text – ‘indecently

assaulted' [i.e. without a typical exemplar] - has *some* instantiation by the more constraining 'sex assault' of the headline towards the typical exemplar of 'rape'. Given also IR's propensity (or any reader's) to try to 'ask why?', to generate a causal antecedent from a specified causal consequent (being 'kicked out of the service and jailed...'), we would expect such instantiation in IR's discourse.

If the events being described were of a 'typical sexual assault', then this instantiation of the main body of the text by the headline, even though the function of a headline is usually more interpersonal than ideational, would not matter. However, what actually took place was not a typical exemplar of 'sexual assault'. Consider the following from the Guardian as an indication of the real nature of events:

*The Guardian: 8th October 1997*

COURT MARTIAL CONDEMNS GANG'S 'DISGUSTING CRUELITIES'  
NAVY WARNS BULLIES AS 8 RATINGS JAILED (by David Pallister)

Royal Navy sailors were given a stern warning yesterday that bullying will not be tolerated, as a court martial jailed eight naval ratings who carried out a cruel initiation ceremony on two recruits.

They were sentenced to between three and 10 months, and six of them were also dismissed from the service.

Captain John Wright, the president of the court martial panel of five officers sitting at HMS Nelson in Portsmouth, told them: 'This is a case in which you all as a gang used violence against two young men in their first sea draft.

'You humiliated and caused great distress to your victims, such that in one case the young, promising sailor is likely to be unable to pursue his career in the Navy because of the mental anguish he still suffers as a result of your actions.

'The Navy will not tolerate such behaviour. Your actions were not horseplay, they went well beyond that and were despicable and in some cases disgusting acts of cruel ill-treatment.'

The two-day hearing had been told that the recruits, aged 20 and 21, had been subjected to **sexual abuse** and beatings on board the destroyer HMS Southampton, while it was anchored off the coast of Oman in April this year.

One had been taped to a table and the other had been trussed up with rope and sodomised with the metal handle

of [a] mop, known as a 'doodlebug'. One of the ratings' defence lawyers told the hearing that humiliating initiation rites were a routine part of Navy life.

Kieran Fuller, aged 18, and 25-year-old Antony Wilkinson, who both admitted one charge of assault, were sentenced to three months suspended detention. They were left with their Navy careers intact for playing lesser parts in the assaults.

James Rowland, aged 20, who admitted two charges of **indecent assault**, was dismissed from the Navy and jailed for five months. Gary Weston, aged 21, who admitted a charge of indecent assault and another of assault, faced the same jail term and dismissal.

Alasdair Whall, aged 22, who admitted two charges of assault and one charge of conduct prejudicial to good order and Naval discipline, and Lee Phillips, aged 23, who pleaded guilty to two assault charges and one charge of indecent assault, were also dismissed and jailed for five months.

Former able seaman Michael Thomson, aged 27, who had already left the Navy, had admitted four charges of assault and one **indecent assault**. He was dismissed and jailed for eight months.

The final defendant, Gareth Jones, aged 25, had pleaded guilty to two charges of **indecent assault** and three charges of assault. Jones, who was married three months ago, was jailed for 10 months and dismissed from the service. [my bold]

In the circumstances as described, the inclusion of 'sex assault' in the Mirror text is misleading since it leads to some instantiation at least of prototypical sexual contact when in fact what happened was not prototypical sexual contact. There is mention that the court heard of 'sexual abuse' and this is presumably the argument being put forward by the prosecuting lawyers. Now, the length of 'The Guardian' article means more information can be provided than in a 'News In Brief' text. This enables the NP 'sexual abuse' to be given sufficient contextualisation such that the reader will know by that point that *non-typical* 'sexual abuse' has occurred. Recall from chapter 5 the following from Lakoff (1987a: 452) on the symbiosis between prototypicality and basic-level categorisation:

Rosch has observed that simple basic-level expressions are used to refer to a prototypical instance of a basic-level category, but that it is misleading to use such an expression to refer to a nonprototypical instance. For example, if a sparrow lands on the front porch, it is not misleading to report this by *There's a bird on the porch*.

But it would be quite misleading to use such a sentence to report that an eagle had landed on the porch or that a penguin had waddled up the front steps. Similarly, if John hit a baseball with a bat in the usual way by swinging the bat at the ball, we could straightforwardly report that *John hit a ball*. But if he hit a beachball with a pizza platter, or if he hit a ball by throwing a rock at it, it would be misleading to describe such an event

to someone who didn't see it as *John hit a ball*, even though such a description, strictly speaking, would be true. *Hit a ball* has an associated conventional image that characterizes the normal case, and with no further modification we assume that the normal case holds. Thus, conventional images are used to understand even the simplest, most straightforward sentences with no idioms in them.

By the same token, use of 'sex assault' is misleading since what took place was a non-typical instance.<sup>7</sup>

The crucial point about 'News in Brief' texts are their size. Since sufficient details cannot be supplied, on reflection text i) shows what might be called *responsible mystification*. Text ii) on the other hand suffers from *headline instantiation of the text body*. In other words, the pragmatic use of 'SEX ASSAULT' to capture a reader's attention exceeds its interpersonal function and becomes embroiled in the ideational function of the text as a whole.

### 8.5.2 Contrast with Other News Texts

*News in Brief Text - 'Sailors Jailed': The Times, 8th October 1997*

Now, as an actual example of 'responsible mystification', consider the 'News In Brief' report of the same events by 'The Times':

The Times: 8th October 1997

#### **SAILORS JAILED**

Eight sailors from the destroyer *HMS Southampton* were jailed for subjecting two recruits to a degrading initiation ceremony. The eight mechanics had forced the two 20-year-olds to undergo a series of disgusting acts when they joined for their first sea posting. [my bold]

'Acts' here can be treated as an ad-hoc superordinate category. Superordinates are not



characterised by cognitive economy. Since IR does not make much cognitive effort and the fact that context is not so constraining, a *typical* exemplar is not readily generated in IR's discourse when what in fact happened was *non-typical*. In other words, the text is being *responsibly mystifying* since 'acts' is not readily instantiated in the discourse of IR.

*News In Brief - Sex Attack Arrest: The Times October 22nd 1997*

#### SEX ATTACK ARREST

1. A 23-year-old man has been arrested on suspicion of a **sexual attack** on a woman who was dumped in the boot of her car after being abducted in Loughborough town centre and **raped** in fields last Saturday.
2. The **assault** lasted seven hours.
3. The car was later found by a farmer near Bolton. [my bold]

In the headline, again 'sex attack' is more constraining, for IR, towards the typicality of penile penetration associated with 'rape' than 'indecent assault'. This is confirmed when IR reaches the body of the text and encounters 'raped' i.e., 'direct manipulation', a high degree of interaction between agent and patient, which in turn instantiates 'assault' in sentence 2. As a result, the headline, which in general usually has an interpersonal rather than ideational emphasis, does not have to provide the ideational component and thus instantiate the superordinate category 'assault' in the text body.

*Chaplain Wins Sex Case Fight: Evening Standard 2 June 1998*

#### CHAPLAIN WINS SEX CASE FIGHT

#### CAPTAIN WEEPS AT COURT MARTIAL VERDICT

By Ed Harris

AN ARMY chaplain who was accused of indecently assaulting a soldier's wife broke down in tears this afternoon as he was cleared by a military court.

Captain Richard Landall, 41, was found not guilty of four charges of indecent assault. All the offences were

alleged to have taken place in Germany in November, after the woman had sought pastoral help from him. While chaplain to the Second Battalion Royal Regiment of Fusiliers and stationed in Celle, near Hanover, he was alleged to have grabbed the woman's breast, rubbed his groin against her, put his hand down her knickers and kissed her mouth.

The father-of-four was cleared at a court martial at Aldershot barracks...'

The headline includes the modifier 'SEX CASE'. But because of the length of the article, the actual charges brought against the chaplain are detailed and so, in this text, we actually find out the nature of the indecent assault. Consequently the function of the headline which is usually in the favour of the interpersonal function is not tipped in favour of the ideational function.

#### **8.6. 'Interpreter 'Was Prepared to Prostitute Herself' News Text: Evening Standard 21 November 1997**

We saw in 8.4.1 how incorrect causal antecedents can be generated because of the propensity of readers (in 8.4.1, IR) to want to instantiate general categories by relating them to familiar knowledge. In the following, I show how a misleading causal antecedent can be constructed because of irresponsible text configuration, rather than instantiation. Consider the following:

*Evening Standard News Text 1: 21 November 1997*

#### **INTERPRETER 'WAS PREPARED TO PROSTITUTE HERSELF'**

'Squadron leader, his mistress and the other man'

1. THE YOUNG lover of a squadron leader accused of murdering his wife was 'tricky and manipulative' and prepared to prostitute herself with an older man, a court heard today.
  2. Nicholas Tucker, 46, is said to have murdered his wife over his love for Dijana Dudokovic, whom
-

### The 'schoolboy' love letters: Page 5

---

he met in Bosnia.

3. However, when he telephoned her about his wife's death, she 'realised her future had broken down and she was ready to do anything.
4. She said that she was even ready to prostitute herself and even go with an old man if she could find one.'
5. She went off to live with a Swiss restaurateur in Verbier.

[Photo of attractive blonde woman - caption - Serbian interpreter Dijana Dudukovic, 23, was described in court today as 'tricky and manipulative'.]

We have seen (6.2) amongst the minimalists and later constructionists that there is a consensus that the reader searches for the causal antecedent. And so even IR, a reader with little vested interest in a text, seeks to generate causal antecedence in discourse derivation.

Consider now the complement 'tricky and manipulative' and its position. Dudokovic is the *global theme* of the article being the subject of the headline which is reinforced by her picture. She is also the *sentence theme* of sentence 1 which establishes a relational function with 'tricky and manipulative'. The effect of both the global and the local topicalisation is that IR would make a *connecting* causal antecedent inference between Dudokovic's 'tricky and manipulative' personality and the death of Tucker's wife. This is reinforced by 'tricky and manipulative' in the photo caption. The most straightforward *elaborative* causal antecedent inference here is that *Dudokovic manipulated Tucker to commit murder*. Indeed, the seeming immorality of Dudokovic is initially abetted by the information that she 'was prepared to prostitute herself'. The 'prostituting' is with an **older** man and so IR may assume that this is Nicholas Tucker, aged 46, as he is mentioned in the next sentence (2).

But Nicholas Tucker is not the 'older man' and indeed 'older man' becomes '**old** man' in sentence 4. It is difficult also to know what Dudokovic means by 'her future' without a great deal of speculative forward elaborative inferencing as to why. But this would be a *strategic* inference which would not be forthcoming from IR. Furthermore, the fact that her

future had broken down does create some kind of narrative conflict. If she had egged Tucker on to murder his wife, then, she would presumably be more prepared for the outcome. All the same, sentence 4 *conflicts with* but does not *cancel* the previous IR causal antecedent inference that she manipulated Tucker into the murder. Now consider the follow-up text to this story on page 5:

*Evening Standard News Text 2: 21 November 1997*

**'The Squadron leader's letters of love, love, love'**

by PAUL CHESTON

**Courts Correspondent**

1. The RAF officer accused of murdering his wife sent schoolboy-style love letters to his girlfriend, constantly repeating the words 'I love you' for five pages, the court heard today.
2. Squadron Leader Nicholas Tucker, 46, is said to have murdered his wife Carol over his love for Serbian interpreter Dijana Dudukovic, now 23, whom he met while acting as a United Nations military observer with the peacekeeping force in Bosnia.
3. **However, the jury was told that Miss Dudukovic, was 'tricky and manipulative' and wanted to make a secure future for herself in western Europe.**
4. She is now married and living with another man in Switzerland, the court was told.
5. Father-of-two Tucker is accused of staging a fake road accident on a lonely Suffolk country road to cover his attempt to strangle and then drown his wife in July, 1995.
6. The court has been told that he had an affair in Bosnia with Miss Dudukovic and also brought her to London to stay in a double room in the RAF Club in Piccadilly.
7. Bertrand Du Pasquier, who met Miss Dudukovic while acting as a UN refugee official in the former Yugoslavia, told Norwich Crown Court how she came to stay with him at his home in Geneva at the request of her parents, who feared for her safety in former Yugoslavia.
8. 'Tucker rang her at least once a day, sometimes twice, sometimes lasting five minutes, sometimes 50 minutes.
9. He started ringing her from the first day she arrived here,' he said.
10. 'I don't believe during that time she wrote to him but she received letters.
11. One had five or six pages with the same words repeated all the time like a schoolboy tradition - but the sentence was not what a schoolboy would say.
12. He was repeating 'I love you' for line after line on every page. It was signed 'Nick'. She burned the letters.'
13. Two days after Mrs Tucker's death the Squadron Leader rang to tell Miss Dudukovic what had happened,

the court heard.

14. 'After that call she realised it was not possible to go to him in England as was planned,' said Mr Du Pasquier.

15. 'She realised her future had broken down and she said she was ready to do anything, even prostitute herself and go with an old man if she could find one.'

16. A few days later she went with Mr Du Pasquier and one of his children to stay in his chalet in Verbier.

17. She met a restaurant owner there and stayed on with him.

18. 'I never heard anything from her again,' said Mr Du Pasquier.

19. Tucker of Honington, Suffolk denies murder.

20. The case continues. [my bold]

IR now knows that 'future' from page 1 refers to a desire to secure a future in Western Europe. Sentence 3 does not now have the topicalisation effects that the complement, 'tricky and manipulative', had on page 1. However, despite the length of the article, the elaborative causal antecedent inference from page 1 that Dudokovic manipulated Tucker to commit murder is **not** cancelled. Consider 13-17:

#### *CAUSE*

13. Two days after Mrs Tucker's death the Squadron Leader rang to tell Miss Dudukovic what had **happened**, the court heard.

#### *CONSEQUENCES*

14. 'After that call she realised it was not possible to go to him in England as was planned,' said Mr Du Pasquier.

15. 'She realised her future had broken down and she said she was ready to do anything, even prostitute herself and go with an old man if she could find one.'

16. A few days later she went with Mr Du Pasquier and one of his children to stay in his chalet in Verbier.

17. She met a restaurant owner there and stayed on with him. [my bold]

Recall from chapter 5 the following from Ungerer and Schmid (1996: 102) on superordinate verbs:

...their main function is to highlight one very general attribute which is part of a whole range of basic-level action categories...Other candidates for superordinate action categories with a salient general attribute are HAPPEN, BECOME, BEGIN and STOP.

The verb 'happen' (see sentence 13) highlights a general attribute and so can be taken as a superordinate verb. Now let us return to the text. The consequences are well specified and certainly we would expect IR to generate a *connecting* causal antecedent inference between sentence 13, Tucker 'rang to tell Dudokovic what had happened' and sentence 15 '...she realised her future had broken down'. In other words, IR knows *generally what* [i.e. something in the phone call] caused Dudokovic to realise her future had broken down but does not know *why*, something any reader is generally interested in; see quote from Graesser et al. (1994: 379) in 8.2.1 above. For IR, encyclopaedic knowledge of the causal relations between sentence 13 and 15 could not be that strong in memory since they do not know the characters of the story. Thus, IR's discourse at this part of the text is mystifying as to these causal relations. In other words, while IR makes a *connecting* causal antecedent inference, the *elaborative* causal antecedent inference, an inference based on background knowledge, is necessarily shallow. This can all be tied to the results of Vonk and Noordman (1990). In other words, in the absence of sufficient background knowledge, even coherence inferences can be shallow.

As before, any speculation as to Dudokovic's 'role' in the murder, so as to provide a stronger elaborative causal antecedent inference, is *strategic* inferencing and not the automatic variety of IR. Besides, that is all it would be - speculation. The fact is there is no evidence in text 2 that she was or was not an accessory to murder. The upshot of all this is that the automatic causal antecedent inference derived in IR's discourse from text 1 that Dudokovic manipulated Tucker to murder his wife is still *not* cancelled when IR finishes text 2. Because this is a report of what happened in court and there has been no verdict passed as yet, there are no *facts* yet pertaining to whether Dudokovic was party to the

murder. Consequently, in such a circumstance, newspapers should show prudence in topicalisation and condensation of information. Inferences generated because of the natural propensity of readers to try to ascertain causal antecedence, prompted by such topicalisation and condensation, may not be cancelled because evidence may not yet be available to assist abrogation. To illustrate the point, it might be fruitful to compare the above with the following text:

*The Sun News Text: September 5th 1995*

## HYPNO-FIST

### Bruno: Telly Paul gave me better trance

PHOTO OF  
PAUL MCKENNA

PHOTO OF  
FRANK BRUNO

caption: Paul...hotel sessions  
WORLD EXCLUSIVE

caption: Frank...no-nonsense mood

WORLD boxing champ Frank Bruno last night revealed how he won his title with help from telly hypnotist Paul McKenna.

The mesmerising star put Big Frank in a tough, no-nonsense mood for his fight with American Oliver McCall. The pair, who are long-time pals, chatted like 'best-mates' to make the boxer think positively during two sessions days before Saturday's bout.

New WBC heavyweight champion Frank 33, told The Sun exclusively: 'Paul came round to my hotel and talked me through a load of things.

'He didn't hypnotise me - but he helped me relax and it was really constructive. He built up my confidence through calm, friendly chat.

'He was a very nice, cool bloke who talked common sense and helped me get in the right frame of mind.'

In 6.2.4, from McKoon and Ratcliff (1989a), we saw that for an inference to be generated automatically, a fair degree of semantic associative richness must be present. In the above

text, the semantic field ‘mesmerising’, ‘sessions’, ‘hypno’, ‘trance’ etc create a strong semantic associative richness and thus IR is likely to construct the elaborative inference that Bruno was hypnotised. However, later in the text we have a quote from Bruno (which I assume to be true) that McKenna did not in fact hypnotise him. The semantic associative inference that Bruno was hypnotised is thus *cancelled*.

### **8.7 ‘The Thin End of the Veg’ Text: Today October 20 1995**

The text I analyse below depends on an understanding of the complex category ‘vegetarianism’. I shall indicate why I think it is complex below and how the discourse of IR would be mystifying. Consider the following extract:

#### **The Thin End of the Veg<sup>8</sup>**

**As Linda McCartney withdraws her ‘too fatty’ burgers, is life without meat really so healthy?**

by DOMINIC MIDGELEY

TODAY Friday October 20 1995

1. ITS growing number of disciples would have us believe that it is the wonder diet for the next millennium.
2. Turn vegetarian, they say, and your chances of dying from cancer could plummet by 40 per cent, and from heart disease, by 28 per cent.
3. What’s more you would suffer 80 per cent less food poisoning and reduce your chances of contracting diabetes or arthritis.
4. But on Wednesday, the vegetarian passport to long life and happiness was left looking a bit tattered around the edges.
5. Celebrity veggie Linda McCartney withdrew from sale thousands of packs of her Beefless Burgers after a TV programme revealed they contained twice as much vegetable fat as advertised.
6. While the list of ingredients on the packets gave a fat content of 11.2 per cent, laboratory tests by ITV’s The Big Story showed levels of up to 22.7 per cent.
7. And the revelation coincided with a report from the Food Commission which found that 17 out of 21 popular



brands of veggie sausages and burgers were high in fat.

8. The worst offender was a Granovita sausage which contained more fat than a pork sausage.

9. [ital] *So are Britain's three million vegetarians being taken for a an unhealthy ride by veggie food producers?* [ital]

10. Sue Dibb of the Food Commission thinks so. 'The high price of these products and the relatively cheap ingredients suggest manufacturers are trying to cash in on the growing interest in vegetarian food', she says.

11. As if this wasn't enough, the Meat and Livestock Commission is weighing in with a campaign that claims the fat content of pork can be as low as four per cent, roughly the same as cheese.

12. 'Vegetarians claim they are healthier,' says Commission spokesman Phil Saunders. 'They claim they live longer, suffer less from every disease known to man and probably at least a third will live forever.'

13. 'But you can get too much fat from a vegetarian diet just as can from a meat diet'...

[picture of vegetables - caption 'Burgers are off: Linda and the veggies are under fire']

As an indication of the complexity of the issue of vegetarianism, I reproduce an extract from the British Nutrition Foundation web site [www.nutrition.org.uk](http://www.nutrition.org.uk):

Variations in strictness of vegetarianism are great and are largely dependent on the person's beliefs and reasons for adopting vegetarianism. This may be for a variety of personal, philosophical, ecological and economical reasons.

The issue is a complex one and so vegetarianism may not be appreciated by everyone to the same degree. The web-site defines different types of vegetarianism in terms of what is excluded.<sup>9</sup> For instance:

*Lacto-ovo-vegetarian*

Exclusion of all meat, fish and poultry; milk, milk products and eggs are still consumed

In other words, the vegetarian types are defined *negatively*, hence use of the term *exclusion*.

The *positive* definition, what is eaten by the vegetarian type, is secondary. However, while we have allusion to the negative definition in the headline:

As Linda McCartney withdraws her 'too fatty' burgers, is life without meat really so healthy?

**negative definition**

the actual thrust of the article rests predominantly and thus anomalously on a positive definition, i.e., that vegetarians eat manufactured ‘veggie-food’:

4. But on Wednesday, the vegetarian passport to long life and happiness was left looking a bit tattered around the edges.

9. So are Britain’s three million vegetarians being taken for an unhealthy ride by veggie food producers?

**positive definition**

Consider all this broken down logically. Use of the term ‘vegetarian’ below does not include fruitarians and macrobiotics:

- i) If X is a vegetarian, then X eats vegetables            **valid**
  - ii) If X eats vegetables, then X is a vegetarian        **invalid** [*Affirming the Consequent*]
- i) does not imply ii).*

On the basis of this, one could argue that the picture and the caption are misleading since it is wrong to ‘define’ a vegetarian by what they eat at the expense of what they *do not* eat. In the same way:

- iii) If X eats manufactured ‘veggie-food’, then X is a vegetarian        **valid**  
[This of course is not true *absolutely* - carnivores can, should they choose, eat manufactured ‘veggie-food’ but eaters of such food are at least *likely* to be vegetarians.]
  - iv) If X is a vegetarian, then X eats manufactured ‘veggie-food’        **invalid** [*Affirming the Consequent*]
- iii) does not imply iv).*

The text, then, commits the logical offence of affirming the consequent. Recall the difficulty subjects had with the Wason test (6.4.2) and how in most cases subjects affirmed the consequent. However, when a logical problem was couched in a familiar context then processing became easier. IR, by definition, has little vested interest in the text: IR has, then,

little vested interest in vegetarianism. Firstly, this would then diminish the likelihood that IR would notice the affirming of the consequent. Secondly, the processing labour necessary to notice that the consequent is affirmed is increased with the problems of defining vegetarianism. So, since IR only makes minimum cognitive effort, the processing labour needed to notice that while the positive definition is present, the negative definition is absent, would not be invested in IR's discourse. The overall effect is that IR is not likely to notice how the concept of vegetarianism is being mystified in *discourse*. Or to put it another way, similar to the 'Reclaim the Streets Evening Standard Text 1', IR's *discourse* is inadvertently *biased* against 'vegetarianism'. As a coda to the above, I am not suggesting that people have difficulty in processing logically. Rather, this kind of processing will incur a higher degree of cognitive labour if the reader cannot draw upon familiar encyclopaedic knowledge in discourse. And for a reader with little vested interest in the text, this cognitive labour is unlikely. Conversely, a reader with a vested interest in vegetarianism will no doubt notice the logical offence in deriving a different discourse from that of IR. That is, the discourse of such a reader is *not biased* against vegetarianism.

## 8.8 Fairclough's (1989) 'Quarry Load-Shedding Problem' News Text

### 8.8.1 *Highlighting how the Text Can Lead to a Different Mystification in Reading Compared to that Found by Fairclough*

Recall from chapter 1, the use of the following text from Fairclough (1989: 51):

#### Quarry load-shedding problem

UNSHEETED lorries from Middlebarrow Quarry were still causing problems by shedding stones on their journey through Warton village, members of the parish council heard at their September meeting.

The council's observations have been sent to the quarry management and members are hoping to see an improvement.

In his commentary, Fairclough (1989: 51) argues that causality in the above is not clear because of the surface syntax - 'it is difficult to take seriously the notion that the *lorries* are the cause of the problem'. He contends that 'unsheeted lorries' contains 'unspecified causality' 'for *unsheeted* implies the failure of a process to happen - someone did not put sheets over the loads' and that the cause of the problem in a different representation could be this 'someone'. However, from the psycholinguistic evidence I highlighted in chapter 6, causal antecedence can be inferred *across* sentences, this evidence conflicting with Fairclough's syntax-first symbolism. Since IR readily tries to generate causal antecedence, there would be little difficulty in inferring that the quarry management is at the head of the causal chain since they own the lorries etc. To gloss this slightly, IR should have little problem establishing a *connecting* causal antecedent inference between 'quarry management' and 'shedding stones'. The richness of the *elaborative* backwards causal antecedent inference is of course very much dependent upon the richness of the background knowledge of a reader as to how a quarry management might manage its lorry drivers etc. To sum up: the text is actually not mystifying of causality in the way in which Fairclough contends.

Now let me show how from the perspective of the IR framework the text can lead to a *different* mystification from that identified by Fairclough. The noun 'problems' is here an ad-hoc superordinate. Superordinates are less 'directly understood' than basic-level categories, and are not characterised by cognitive economy. So, since IR is a parsimonious reader, little effort will be invested to ascertain the possible nature of these 'problems'. As I have indicated previously, causal *consequent* inferences are not generated in the discourse

of IR or only shallowly at best. They are more likely when the situation is *highly constrained* by being familiar, having a high degree of semantic associative richness and where there is only one obvious outcome. However, a causal consequent inference here would be weak at best since the causal consequent of stones being shed does not fit these criteria. Recall also (5.4.2) that the word ‘cause’ characterises non-central causation (Lakoff, 1987a: 55). Moreover, recall (5.6.2) how Ungerer and Schmid (1996: 102) suggest that verbs like ‘cause’ and ‘become’ can be seen as superordinate verbs because:

...their main function is to highlight one very general attribute which is part of a whole range of basic-level action categories.

Now, taking into account that:

- i) since both ‘problems’ and ‘cause’ are superordinates, they are *not* characterised by cognitive economy
- ii) ‘cause’ characterises *non-central* causation
- iii) IR has little vested interest in the text

any causal consequent inference for IR, a reader who invests minimum cognitive effort, as to how the stones caused problems would be *very weak* indeed. What Sanford (1990) (6.2.4) says about topicalisation is also pertinent to this text. ‘UNSHEETED lorries from Middlebarrow Quarry’ is *theme* and ‘were still causing problems by shedding stones on their journey through Warton village’ the *rheme*. Since ‘shedding-stones’ is included in the rheme, again, it makes it less likely that an elaborative inference in IR’s discourse, as to what the ‘problems’ are, will be produced.

Unlike Fairclough's analysis, the IR framework does not just focus on the text itself at the expense of a particular kind of reader. [However, in the absence of an explicit notion of a reader here, but with Fairclough's comments on the syntax and absence of agency, the implication seems to be that, for Fairclough, readers *consume* the text (see 9.2).] An analysis of the text *only* would reveal that a *causal consequent* has not been specified. The IR framework, however, attends to the *discourse* perspective of a particular reader and so is both text *and* discourse based. On this basis then it is not just that the text does not stipulate the causal consequent, but IR, being non-analytical, *would not notice this anyway*. That is, IR's discourse is *inadvertently* mystifying as to what the 'problems' were.

Now consider the following variations:

UNSHEETED lorries from Middlebarrow Quarry were still causing problems by shedding stones on their journey through Warton village, members of the parish council heard at their September meeting.

a) Particularly concerned were village mothers.

OR

b) Particularly concerned were village mothers afraid of injury to their babies.

The council's observations have been sent to the quarry management and members are hoping to see an improvement.

With a), the context is beginning to become constrained and so we might expect the causal consequent to have *some* degree of generation, i.e., more than in the original. With b) a (possible) causal consequent has been specified and processing of the event for IR is less *shallow*. I should stress here that I am not suggesting that the 'Quarry' text is deliberately mystifying the nature of the event in a sinister way because it does not specify the causal consequent. I just use this text to show how the IR framework isolates text which can lead to mystification in reading from a perspective which conflicts with the mystification Fairclough perceives.

### 8.8.2 *The Interpersonal Function of the Headline*

Consider again the headline that Fairclough (1989: 50-51) refers to and his commentary:

Quarry load-shedding problem

‘the grammatical form in which the headline is cast is that of a *nominalization*: a process is expressed as a *noun*, as if it were an entity’

Fairclough (1989), in the above, is highlighting how the nominalisation has, in the removal of agent and patient, mystified the causality of the event. Similar to Fairclough’s analysis of the ‘cancer’ text in 7.5.3, this is an example of treating a headline from an ideational perspective when its function is interpersonal. Like Fairclough, I too want to criticise the headline - but from an interpersonal perspective rather than an ideational one. The sentence ‘the lorry shed its load’, because of the use of the basic-level ‘lorry’, and thus more directly understood category, can lead to the generation of a gestalt. However, ‘Quarry load-shedding problem’ is difficult to generate a gestalt from because of the lack of directly understood basic-level categories. There are two, loosely speaking, superordinate categories [‘load’ and ‘problem’] which are necessarily only *indirectly meaningful* for IR. It is usual for a headline’s interpersonal function to be more prominent than its ideational function. One effective way of achieving this is to use directly meaningful concepts, i.e., basic-level categories, rather than the superordinate categories used. To sum up: construing the ideational function of the headline as being mystifying is inappropriate. Rather, it is its interpersonal function which is *ineffective*.

As an example of a headline which *is* successful from an interpersonal perspective, consider the following newspaper headline and extract from The Express (Saturday January 23 1999):

### **Klinsmann lifts English 2006 Cup bid**

JURGEN KLINSMANN gave England's 2006 World cup campaign a huge boost last night - by walking out on his role promoting Germany.

Former Tottenham star Klinsmann had been acting as an ambassador in the Germans' battle with England and South Africa to stage the finals.

However, his plea for a testimonial match was turned down by the German Football Federation – 'We will not make any exceptions,' said vice-president Franz Beckenbauer - and he quit his post.

'They made me sick with this refusal and I'm not going to continue with the bid,' said Klinsmann. Fedor Radmann, head of the German bid committee, said he deplored the withdrawal but added: 'We have to accept it unfortunately.'

The 'Klinsmann lifts English 2006 Cup bid' headline is characterised by punning expected in newspaper headlines. But from the perspective of cognitive linguistics, this punning is salient. When football teams win competitions, they lift trophies. Until IR encounters the text body, naturally this is all they have to go on. Klinsmann, a German and former Spurs footballing hero, is subject, the 'cup bid' is object. It is likely that the motor-interactive sense of 'lift' will feature in the initial processing, particularly given that prototypical causation involves direct manipulation of a patient as a real object with use of hands, body etc (Lakoff, 1987a: 54-55) (5.4.2). The headline is actually an abstract 'lifting'. However, before the text body is processed, IR interprets Klinsmann as being involved in direct causation as AGENT in the reviving of the World Cup in 2006, instead of the rather indirect causation that is actually the case from the text body. Again, it is inappropriate to say that the headline 'mystifies' from an ideational point of view. Rather, it is successful from an interpersonal point of view because of what is associated with prototypical transitivity.



## 8.9 Providing a More Comprehensive Explanation of Trew (1979: 100-1) / Montgomery (1995)

While I have in previous chapters criticised CL as well as CDA, I endorse the following analysis of Trew (1979) (and Montgomery's (1995) follow-up) but show how my framework provides a fuller explanation of the mystification which Trew and Montgomery highlight.

Consider first the text and then the examination by Trew (1979: 100-1), an analysis subsequently endorsed by Toolan (1988: 231); Lee (1992: 101-2); Montgomery (1995; 240):

DAY 2 *The Times*

### **Split threatens ANC after Salisbury's riots**

After Sunday's riots in which 13 Africans were killed and 28 injured, a serious rift in the ranks of the African National Council became apparent today.

In *The Times* report itself there is a reference to the killing, but in a way that is significant. It is in the clause: After Sunday's riots in which 13 Africans were killed.

...it is in the passive form, and the agent is deleted. But more than this, the description is changed from 'shot dead' to 'killed' so that any reference to the manner of death is deleted. The new description gives no hint of the agent or the manner of death - there is only a suggestion of a cause resulting from the way the 'riots' are made focal and made the context of the deaths.

Trew (1979: 100-1) does not regard as particularly strong the causal antecedence inference that the riots killed the 13 Africans. In other words, for Trew, this is mystifying of the actual causal relations. Before I return to this analysis by Trew, let me cite again another of Trew's analyses. Previously (1.4.1) we saw another instance of Trew (1979: 98-9) regarding a causal antecedence inference as a *weak* representation:

*The Times*

Eleven Africans were shot dead and 15 wounded when Rhodesian Police opened fire on a rioting crowd of about 2,000.

...the syntactic agent is deleted ('11 Africans were shot dead by...') **and is identified only weakly by implication** through the temporal conjunction with the police opening fire ('when police opened fire on a rioting crowd of about 2,000'). [my bold]

In 6.6.2, I outlined how Keenan et al. (1984) showed that the strength of a causal antecedence inference was proportional to the degree of strength of causal relation in memory so that causal strength decreased from 1 to 4:

level

- 1      Joey's big brother punched him again and again.  
The next day his body was covered with bruises.
- 2      Racing down the hill, Joey fell off his bike.  
The next day his body was covered with bruises.
- 3      Joey's crazy mother became furiously angry with him.  
The next day his body was covered with bruises.
- 4      Joey went to a neighbor's house to play.  
The next day his body was covered with bruises.

The inference then that police were responsible for the deaths, then, would be fairly strong for most people since it is safe to say that *people being shot dead after police open fire* is a fairly strong causal relation in the memories of most people, similar to 1 above. But while Trew was wrong about the strength of the inference based on '11 Africans were shot and 15 wounded...etc', he is actually *correct* about the strength of the inference in 'After Sunday's riots in which 13 Africans were killed'. Why then is Trew correct when he says 'there is only a suggestion of a cause resulting from the way the 'riots' are made focal and made the context of the deaths'? When Trew says this, it is because he is only able to make a *connecting* causal antecedent inference. However, the *elaborative* causal antecedent inference he makes is not so strong since the sentence resembles 4 above. In other words, the causal relation in memory for Trew (and most people) that riots lead to deaths is not as strong as people being shot dead when police open fire on them.

Finally, consider the following from Montgomery's (1995: 240) commentary on Trew's text:

...the agent remains completely unspecified. This vagueness is reinforced if anything by the selection of 'were killed' rather than 'were shot' **which would have at least implied someone to do the shooting.** [my bold]

Again I provide a fuller explanation from the perspective of my framework on why Montgomery's point is essentially correct. If 'were shot' had been used, 'with guns' would be a likely accompanying instrument inference. It is a likely inference given the cognitive interdependence between 'gun' and 'shoot', the high degree of motor-interaction etc. It thus facilitates IR's natural ability to seek causal antecedence and thus seek a different causal agent other than the 'riot'.<sup>10</sup>

### **8.10 Endpoints**

I hope I have shown that the framework:

- i) reveals aspects of text which could lead to mystification in non-analytical reading, given the conditions of idealised reading I have set out.
- ii) reveals that IR would not notice this mystification and that their discourse is inadvertently biased.
- iii) has demonstrated enough systematic breadth to be able to highlight evidence of sound CDA with regard to the relationship between inference generation and mystification.

In the final chapter (chapter 9), I provide an overarching contrast between CDA's and my framework's assumptions of text processing as well as a summary of what I have achieved in this thesis.

## Notes

1. To make clear what constitutes 'hard news', here is Bell (1991: 14):

'Newswriters' basic distinction is between hard news and features. Hard news is their staple product: reports of accidents, conflicts, crimes, announcements, discoveries and other events which have occurred or come to light since the previous issue of their paper or programme. The one-off, unscheduled events such as fires and disasters are sometimes called 'spot news'. The opposite to hard news is 'soft' news, which is not time-bound to immediacy. Features are the most obvious case of soft news. These are longer 'articles' rather than 'stories' covering immediate events. They provide background, sometimes 'editorialize' (carry the writer's personal opinions), and are usually bylined with the writer's name...For both newswriters and researchers, the boundaries between hard and soft news are unclear...journalists spend much of their energy trying to find an angle which will present what is essentially soft news in hard news terms. Journalists and media researchers both recognize hard news as the core news product, the typical against which other copy will be measured. Hard news is also the place where a distinctive news style will be found if anywhere.'

2. The framework would also disallow a post-structuralist approach to (non-literary) language. Post-structuralism's focus on 'full' signifiers makes it a bottom-up based approach to meaning production, neglecting the expense of top-down expectation processing and how the semantic value of signifiers in context is partial in *automatic* natural language processing. With literary texts, it is a different matter. Literary texts, particularly poems, often do not present obvious beginnings for top-down processing. They often initially draw the reader into bottom-up tactics, make the reader incur more cognitive effort and thus *strategic* processing, and in so doing, they can lead to the refreshment of conventional top-down strategies (cf Cook, 1994). That is, with literary text, readers may begin with 'full' signifiers in the absence of obvious top-down beginnings.

3. Another *absence* in the first page of the *text* is the nature of the protestors. I.R.'s mystifying discourse as to the nature of the protestors, imagining them to be striking L.U employees, is another example of discourse which is (initially at least) *biased* against the protestors. Recall from chapter 3 that when I looked at CDA's approach to highlighting mystifying text, we saw that the *absence* of agent-process-patient structures to describe events, and the need to generate agency inferentially, was seen as leading to mystification. But we saw in chapter six that even though *causal antecedence* may be 'absent' in a text because it is *indirectly present*, it can, all the same, be readily generated even if a reader has little vested interest in a text. Thus, not all text absences can be necessarily related to mystifying discourse.

4. To show the Evening Standard's (text 1) narrow selection of information, and downplaying of the aims of the protest, compare with The Daily Telegraph's report 'Cyclists' Demo Adds to Tube Strike Chaos' (Thursday August 8<sup>th</sup> 1996 p.6):

'One of them started throwing my papers around,' Mr Ford said. 'I asked them to leave my family photographs alone which they did.'

'I asked them what the problem was and one said they were in favour of bicycles, so I told them I was a keen cyclist. Another said they were supporting the Tube Drivers. A third said they were campaigning for more investment in public transport. I told them that on that issue we were on the same side.'

One more point. Compare the headline of the Telegraph text ('Cyclists' Demo Adds to Tube Strike Chaos') and the headline from the Guardian text ('Cycle Protest Adds to Tube Disruption'). Both mention the *protest* in the context of 'cycles' and so the pictures are easy to relate to the text. In effect, this highlights the absence of a link between 'cycles' and 'protestors' in 'Evening Standard' text 1, how the picture is difficult for IR to relate to the text, and how, as a result, the nature of the protest is downplayed.

5. The category of 'indecent assault' corresponds to a legal definition. The legal definition of 'indecent assault', from Ashworth (1991) is as follows:

*'In English law, the offence of indecent assault is charged for all forms of sexual assault other than rape, buggery, and attempts to commit those crimes'*

There is no typical indecent assault, then, since its range is so wide and disparate, from bottom-pinching to forced oral sex. While the phrase 'indecent assault' is commonly encountered in newspapers, I assume most people are unaware of the legal definition.

6. Data from the British National Corpus would seem to support a reading of 'sex assault' as being close to, if not, equivalent to 'rape'. Here are a few contexts for 'sexual assault':

*'In our recent study of papers in the mid-1980s we broadened the focus and examined not only rape but also other forms of sexual assault and not only reporting of trials but also other stages, including the search and post-conviction.'*

*'The narrow definition of rape and its failure to increase the categories of sexual assault may be attributed to the early origins of the offence.'*

*'In common parlance, however, rape is used to describe the sexual assault of men as well as women, as, for example, in the prison context.'*

*'On the other hand, if a mature woman is sexually assaulted by a stranger in daylight or at home, then she has not been raped by a man but by a monster.'*

As further support that 'sex assault' is closer to rape than 'indecent assault', consider another 'News in Brief' text and what is bolded:

THE TIMES Wednesday October 22nd 1997 - NEWS IN BRIEF COLUMN p.2

## SEX ATTACK ARREST

1. A 23-year-old man has been arrested on suspicion of a **sexual attack** on a woman who was dumped in the boot of her car after being abducted in Loughborough town centre and **raped** in fields last Saturday.
2. The **assault** lasted seven hours.
3. The car was later found by a farmer near Bolton. [my bold]

The text sets up equivalences, I would argue, between 'sex attack' and 'rape'. 'Assault' in sentence 2 receives instantiation by 'sex attack' and 'rape' from sentence 1. There would be little, if any, change in meaning if 'attack' in 'sexual attack' in sentence 1 were interchanged with 'assault' in sentence 2. In other words, there is a large amount of similarity between 'sex assault' and 'rape' as much as between 'sex attack' and 'rape'.

7. The relationship between prototypicality and instantiation is often exploited in 'trick questions'. Consider the 'Trivial Pursuit' question - *What mammals can jump the highest?* Instantiations are automatic when they rely on familiar knowledge. So automatically, 'mammal' would be instantiated as one with legs since 'jump' is prototypically associated with 'legs'. But it is the automatic relating to familiar knowledge in instantiations that can lead the player astray. 'Whales', the answer, are neither prototypical 'jumpers', having no legs, and are prototypical mammals neither.
8. This is an example of 'soft news'. See note 1 above.
9. As an indication of the complexity of 'vegetarianism', below is a list of types of vegetarian diets found on the British Nutrition Foundation web site [www.nutrition.org.uk](http://www.nutrition.org.uk) web-site:

Term	Description
<i>'Semi' or 'demi' vegetarian (pescatarian)</i>	Exclusion of red meat or all meat, but fish and other animal products are still consumed; some people also exclude poultry
<i>Lacto-ovo-vegetarian</i>	Exclusion of all meat, fish and poultry; milk, milk products and eggs are still consumed
<i>Lacto-vegetarian</i>	Exclusion of all meat, fish and poultry and eggs; milk and milk products are still consumed
<i>Vegan</i>	Exclusion of all foods of animal origin; diets comprise vegetables, vegetable oils, cereals, pulses such as beans and lentils, nuts, fruit and seeds.

*Fruitarian*

Exclusion of all foods of animal origin as well as pulses and cereals. Diets mainly comprise raw and dried fruits, nuts, honey and olive oil.

*Macrobiotic*

Sometimes referred to as Zen Macrobiotic diet. The diet progresses through a series of levels, gradually eliminating all animal produce, fruit and vegetables and leading to a restricted diet of cereal (brown rice) only. Fluids are also severely restricted.

10. In chapter 5 (5.4.3), we saw that Martin (1989: 43) regarded 'killing techniques' as helping to 'immobilise the most unsavoury part of the seal hunt'. Since the 'action' 'killing' is 'the modifier of an abstract noun', Martin argues this 'refers to the 'killing indirectly'. In 5.4.2, I flagged how Lakoff (1987a: 55) regarded 'kill' as expressing 'direct causation, with cause and result expressed in a single morpheme - the closest possible connection'. On this basis, I criticised Martin since if 'kill' is the closest possible connection, the syntactic position of 'killing' in 'killing techniques' diminishes in significance. But Lakoff's (1987a) cognitive linguistic analysis of 'kill' and my use of it in chapter 5 are in tension with my analysis of 'kill' here. This is because 'killed' in 'After Sunday's riots in which 13 Africans were killed...' is *not* the 'closest possible connection' between the 'cause and effect' in this context. 'Shot dead' is a much better candidate. One corollary that flows from this is that cognitive linguistics needs to deal more with examples from *discourse* and not with isolated lexical examples. Indeed, it is not just discourse context which would have to be taken into account but *genre* as well. For example, killing in a Bond movie is different to killings as reported in news text.

But what about the tension I have drawn attention to? The tension centres around the issue of agency absence. However, the nature of agency is only relevant in the 'After Sunday's riots...' sentence and not in the 'hunting seals' text. So, if the 'cause' of death in a particular context is straightforwardly understood, then, 'kill' adequately expresses the connection between cause and effect.

## CHAPTER 9: CONCLUSION

### 9.1 Introduction

To begin with, I provide an overarching comparison of my framework and CDA by comparing ‘master-metaphors’ so that the reader can see what I have broadly tried to achieve in the thesis. Then from such a broad perspective, I move on to what are more specific demonstrations of the thesis.

### 9.2 Shallow Reading vs Consumption Reading

#### *9.2.1 The Consumption Metaphor as Constitutive Metaphor*

The metaphor I have used for a type of processing based on the scenario of the idealised reader has been that of *shallow* reading. This is very much in contradistinction to the metaphor of text *consumption*, which is often used in CDA. Their notion that texts are *consumed* by readers is particularly prevalent in the work of Fairclough, for example, (1992a: 71-73; 78-79; 85), (1995a: 49-50; 58-9), (1995b: 9) but also in the work of others associated with CL / CDA, for example in Chilton (1988: 79):

...most people don't or won't think round and beyond the language they consume.... Chilton (1988: 79)

and also in the following from Simpson (1993 : 182):

All this brings us on to the question of the position of the analyst relative to the interaction between producers and consumers of a text.



In 7.5.2, I indicated how ‘philosophy of mind’ metaphors are likely to defeat the primary processing principle and constitute processing, particularly for a reader who has little expertise or vested interest in this field. Consequently for such a reader, the primary processing principle is likely to be defeated when coming across the notion of ‘consuming texts’. In other words, the ‘consumption’ metaphor is likely to be a *constitutive* metaphor for such a reader.

### *9.2.2 The Consumption Metaphor as Misleading Discourse*

Certain things follow from the consumption metaphor being constitutive. The consumption metaphor can lead to mystification in the reading process about the *actual nature* of the reading process. In the light of the evidence for shallow / selective processing I outlined in chapter 6, it can hardly be said that everything in a text is *consumed* in the reading process. So much of text processing does not involve ‘consumption’ of the text but generation of *inferences*. We saw also that for IR, a reader with no specific goal, reading is more likely to involve the inadvertent downplaying of phenomena such as secondary characters. But the consumption metaphor would imply that main and secondary characters are processed to the same ‘degree’. The consumption metaphor would also imply that the mental representation of words in a sentence mirrors their compositionality in the sentence. But IR’s processing is also in line with the concept of *accommodation* which is made use of in cognitive linguistics and connectionism. This refers to the phenomenon of automatic meaning reduction or shading along a sentence. That is, while a sentence in a text may be compositional, the mental representation of it is not compositional since readers make *closure* along the length of a sentence. The ‘consumption’ metaphor also reinforces the notion that sentences can be taken as a *representational* medium rather than serving as ‘evokers’ of encyclopaedic knowledge. Moreover, the ‘consumption’ metaphor detracts

from viewing text as an *external device* to be *interacted* with. To sum up, the notion that non-analytical readers ‘consume’ text misleads as to the nature of the *micro-context* of interpretation of a non-analytical reader, as well as going some way to explaining why the notion of a non-resistant reader in CDA is so undeveloped.

### *9.2.3 The Consumption Metaphor as Legitimation of CDA*

So, the consumption metaphor, then, bears little relationship to actual processing. But its use is more insidious than that. The consumption metaphor can actually be regarded as a legitimising metaphor for the practice of CDA. It legitimises a detailed dissection of the text by the CD analyst by proxy for the non-resistant reader on the assumption that the non-resistant reader ‘consumes’ the whole text, making a mental copy of it, and so ingesting ‘encoded ideologies’ etc. The ‘job description’ of the CD analyst that naturally flows from this is to show readers what encoded phenomena have been ingested or would be ingested if these had not been pointed out. Since all of the text is consumed, it allows the analyst to dissect the text without any conceptual limitations, picking out anything they choose since *anything* could be relevant for the *consumptive* reader. And the analyst, free to dissect any part of the text they choose, usually picks out aspects of the text which suit *their own* predilections. The consumption metaphor thus erases not only inter-personal variation amongst readers but the intra-personal variation based on degree of vested interest and familiarity with the text’s subject matter that I have tried to highlight in this thesis.

### *9.2.4 Saying One Thing: Doing Another*

A CD analyst may feel that my points traduce CDA. Fairclough (1996), for example, is explicit about the fact that ‘diversity of interpretation of texts is a central assumption’ in

Fairclough (1992a). Here, for instance, is Fairclough (1996: 50):

...social meanings (including discourse) cannot simply be read off from the text without considering patterns and variations in the social distribution, consumption and interpretation of the text.

But with regards to variation in ‘consumption’ this is merely lip-service. As Widdowson (1995a) points out there is rarely any attempt to *illustrate* ‘diversity of interpretation’.

Furthermore, there are criticisms in CDA of how critical linguistics:

...construe texts as merely producing ideological effects upon a passive recipient...’

Fairclough (1992a: 29)

But there is irony here since what Fairclough criticises CL for is actually in line with his consumption metaphor perspective. Indeed, regardless of their criticisms of CL, CDA, particularly Fairclough, often proceed in something akin to CL manner, supported by the legitimising consumption metaphor (e.g. Fairclough’s (1989) analysis of the ‘Quarry load-shedding problem text’ in 1.4.1 and 8.8.1).

### 9.2.5 Consumption Constitutive Metaphor vs the Shallow Explanatory Metaphor

Of course, ‘shallow’ processing is another metaphor. In contrast to CDA, I have supplied a good deal of psycholinguistic detail as to why ‘shallow’ processing is a more plausible metaphor than ‘consumption’ for the text comprehension of a reader with no specific goal and unfamiliarity of subject matter (see chapter 6). Since I have provided psycholinguistic evidence for ‘shallow’ processing, ‘shallow’ is, then, much more of an *explanatory* rather than *constitutive* metaphor.

To sum up section 9.2: the approach I have taken to text processing highlights how the degree of a reader's cognitive investment in a text affects the type of inferences generated and thus the discourse derived from the text. My approach has been, then, explicitly *interactionist*. While CDA *accords* in theory with an *interactionist* principle, in practice the tendency is to assume that the text is merely mentally facsimiled by a non-resistant reader. In other words, there is little or no attempt to show how the static symbols on a page are actualised *dynamised* by a *particular* reader.

I move on now to what are more specific demonstrations of the thesis.

### **9.3 Specific Demonstrations of the Thesis**

#### *9.3.1 The Problematic Processing Assumptions in CDA*

Recall from chapter 1 where I listed the main processing assumptions of CDA and the ones which guided how CDA highlighted (predominantly) mystifying text. In this thesis I have shown these assumptions to be problematic. I now contrast my framework with these assumptions.

#### *Inferences as Weaker Representations*

In CDA, inferences across clauses or adjacent sentences are often regarded as weaker representations either explicitly or implicitly because more weight is given to the surface structure of a sentence and its ability or inability to reflect reality. On the basis of my framework, informed by recent psycholinguistic evidence, inferences across clauses or adjacent sentences are not weak representations per se. A causal antecedent inference across

adjacent clauses or sentences, where the causal relation is strong in memory, is not a weak inference. The same applies to instantiations, where the background knowledge is readily available, and instrument inferences, which are highly constrained by context. Causal consequent inferences are likely to be *weakly* generated, though, for a reader with no specific reading goal, especially if there are many alternative possible consequents.

### *Inference as Work vs Automatic Gap-Filling*

Fairclough (1989) outlines two types of inferencing: ‘automatic gap-filling’ which requires minimum cognitive labour while ‘inferences’ are those which require higher than minimum cognitive labour. His assumption, emanating from Brown and Yule (1983) and adopted by Gough and Talbot (1996), was that readers are *willing to work* to produce *inferences* necessary for what he regards as the coherence of a particular text, even if the subject matter is unfamiliar. However, we have seen that if the subject matter is unfamiliar for a reader who makes minimum cognitive effort: i) causal consequents are not produced; ii) even coherence inferences, those which are normally instituted, can be shallowly generated, readers satisfying themselves that a text is coherent because it is cohesive. In other words, on the basis of the psycholinguistic evidence I discussed in chapter 6, if readers are unfamiliar with the subject matter of the text and have no specific goal, it is likely that they will *not* work to create inferences as Fairclough supposes. While Fairclough makes a distinction between automatic gap-filling and inferences, using McKoon and Ratcliff (1992) my distinction has been between *automatic* and *strategic* inferences. The latter require work and so depend on the inclination and vested interest of the reader. In terms of this vocabulary, Fairclough supposes that a reader will naturally seek to make the *strategic* inferences he makes in line with his vested interest in a text.

*A High Degree of Nominalisation in a Text Requires Extra Processing Effort*

We saw in 3.2.3 that Hodge and Kress (1993) regard transformations as being psychologically real and in 1.4.2 that a non-energetic reader would not process deeply enough to recover the 'original form' of a nominalisation. Hodge and Kress (1993) in effect were detailing a form of shallow processing but it is a variety of shallow processing that is disallowed by my framework. The derivational theory of complexity has been discredited and so looking at nominalisation in terms of psychologically real transformations is incorrect.

*Excessive Nominalisation Make a Text Abstract and Distant From Concrete Events*

CDA makes the point that since nominalisations remove participants, events are made more *distant*. CDA tend, however, to conflate *ideational distance* with *interpersonal distance*.

I have no problem with nominalisations producing interpersonal distance since formal and so 'distant' letters are characterised by a higher degree of nominalisation, and thus deletion of 1st person participants, than informal letters. In other words, I have no problem with the fact that *interpersonal distance* is created though the absence of participants. But this does not mean, as I showed in chapter 4 and 5, that *ideational distance* is necessarily created by nominalisations through the absence of participants and the presence of 'object-like' nominals. Basic-level noun categories are cognitively interdependent with basic-level action categories. So just because a basic-level noun is used to describe an *action*, it does not follow that its nominal form leads the reader to see the action in an abstract, distant 'objectified' way. Ideational distance can be created, though, when superordinate / abstract categories, *nominal or verbal*, lack sufficient information in the text to lead to their instantiation (e.g. sufficient basic-level categories).

In further contrast to CDA, what I have tried to show is that ideational distance *can* be related to a particular discourse derivation *even if* participants and processes are being used to describe events. Just because participants are present in a text does not necessarily mean that there are no differences in the ‘depth’ to which different participants are processed in *discourse*. Compared to primary characters, secondary characters are *automatically* read in a more shallow way for a reader who has little vested interest in a text. Or put another way, in the discourse of such a non-analytical reader, secondary characters are *ideationally distant*.

#### *Confusing Semantic Transitivity with Syntactic Transitivity*

Following the experimental evidence of Taraban and McClelland (1988), psycholinguistic evidence of top-down expectation overriding bottom-up processing etc, as well as the principles of connectionism and cognitive linguistics where syntax and semantics are interactive, my framework prohibits the imparting of semantic transitivity - the structure *agent-process-patient* - to just *any subject-verb-object* structure, *pace* CDA.

#### *Metaphors are Conceptually Active*

Metaphors *per se* do not *constitute* processing. If they do, it is because they cannot be related to a high degree of experiential encyclopaedic knowledge (predominantly based on sight), leading to the defeat of Sanford and Garrod’s primary processing principle. When they do not constitute processing, when they are *explanatory*, it is because they can be related to a high degree of experiential encyclopaedic knowledge, i.e., in accordance with the primary processing principle. The degree of absence of experiential knowledge is proportional to the constitutive power of the metaphor. A corollary of the connectionist and

cognitive linguistic principle that linguistic input is not equivalent to process output is that metaphor cannot be associated with any one particular lexical item. Rather metaphor is then best seen as the *processing output* which has involved a high degree of microfeatural closure via the constraints of lexical environment.

### *9.3.2 Situating CDA Processing Assumptions in an Historical Context*

In chapter 3, I showed how symbolic assumptions of processing inadvertently underpin much of CDA, assumptions in line with much Anglo-American-Austrian philosophy in the twentieth-century and symbolic cognitive science. This was brought into relief by setting them alongside assumptions of mental representation in connectionism and cognitive linguistics. The inconsistencies in CDA with regard to psycholinguistic assumptions (see: 1.4) no doubt arise, in part, because CDA is largely nescient of its cognitive foundations. In contrast to CDA, my framework has been *overt* about its philosophical and cognitive foundations and has tried to be systematic and consistent.

### *9.3.3 The Partiality of Interpretation Principle*

I have, to some extent, reflected *intra-personal* variation based on degree of vested interest in the micro-context of interpretation. For example, the medical textbook a person read reluctantly as part of a compulsory course leads to a different discourse for the same person when they are reading it again for information about their own illness. In drawing attention to variation of vested interest in a text, (little in the case of IR but high for a CD analyst), I have demonstrated the principle of partiality of interpretation for the same text. Of course ethnographic studies are ultimately better in trying to ascertain the diversity of interpretation of texts. But as I said in 8.2.1, always seeking ethnographic support for a discourse analysis



is arduous. Furthermore, ethnographic research faces the problem, similar to the memory-recall experiments of early constructionists, that inferences might be generated in response to questioning, which might not have been generated in ordinary discourse. This problem is not faced by my framework since it is informed by psycholinguistic experimentation where this has been dealt with.

#### *9.3.4 Exploring the Micro-context of Interpretation to Show CDA 'Over-Interpretation'*

By using the distinction between automatic inferences and strategic inferences, I have shown that what a CD analyst *explains* ('interpretation-2') as being the news text interpretation of a non-resistant reader does not coincide with the discourse of non-resistant IR. It most likely accords with their own vested interest in the text. In other words, CDA over-interprets by proxy for a reader with little vested interest in a news text. [Of course, I recognise that I have also had a vested interest in a text in showing how CDA is misguided. Unlike CDA, my use of IR framework, where only automatic inferences are generated, has prevented the production of strategic inferences in my analyses.]

#### *9.3.5 Where CDA is an Inappropriate Mode of Discourse Analysis*

Since the discourses (derived predominantly from news text) of the non-resistant IR and the non-resistant reader in CDA do not coincide, I have in effect shown that the assumption in CDA that the macro-context *always* weighs heavily on the micro-context (Fairclough, 1992a: 85-86) of interpretation, by a non-resistant reader of news text, is not the case. This is hardly surprising when one considers that usually readers of news text are at liberty to vary the level of interest in the text. It is reasonable to conclude that in a reading situation, where readers are not coerced into a particular interpretation, that CDA is an inappropriate mode

of discourse analysis. A corollary of all this is that discourse analysis needs to be differentiated into (at least) two spheres:

a) Analysis where the macro-context is more likely to impinge upon and restrict interpretation in the micro-context, e.g. in regimented 'institutions' such as religious cults, the armed forces etc. These seem to me more natural constituencies for the kind of analysis performed in CDA.

b) Analysis where the micro-context of interpretation is much less restricted by the macro-context, e.g. as in the reading of news text, as I have tried to show in this thesis.

### *9.3.6 Highlighting Recent Work on Inference Generation in Psycholinguistics*

I have flagged recent important work in cognitive psychology and have attempted to plug some of the gaps in discourse analysis that deals with inference generation (e.g. Brown and Yule, 1983). The psycholinguistic research I have drawn on also lends support for the 'discourse as dialogue' principle of written discourse (Hoey, 1983: 168-188; Widdowson, 1984b: 39-53; and Cook, 1994: 48-51). For example, with regard to causal antecedence, readers are geared to process texts in something resembling a 'dialogical' manner in their posing of 'why' questions (Graesser et al., 1994). However, I have shown that the value of the '(inferential) response' depends upon familiarity with subject matter.

### *9.3.7 An Alternative Framework for the Highlighting of Mystifying Discourse*

#### *Orientation*

I have indicated a host of problems with how CDA highlights mystification in news text and

how the notion of a non-analytical reader in CDA is undeveloped. My framework predicts how certain text can lead to mystification for a *non-analytical* reader who has little vested interest in a text and is largely unfamiliar with its subject matter. I have shown how mystification for this non-analytical reader is connected with inference generation but, in contrast to CDA, I have provided a detailed processing profile for such a reader. Symbolic attitudes in CDA towards inference generation are often in conflict with recent psycholinguistic research. My framework, rooted in empirical psycholinguistic study, enables a more plausible, comprehensive and thus consistent perspective on inference generation in reading and how this relates to mystification. Owing to their different assumptions about mental representation and cognition, where my non-symbolic ‘framework’ detects text that can lead to mystification in reading is not necessarily coincident with where symbolic CDA detects mystification.

#### *Use of I.R as a Yardstick for the Success of a News Text*

Discourse analysis of text from the perspective of IR is useful as a yardstick as to the success of a text for a reader who invests minimum cognitive effort and is largely unfamiliar with the text’s subject matter. With regard to cause-consequents, the IR framework can reveal *discourse bias* as a by-product of mystification related to *text absence* in a news text as opposed to just simply *text bias* which is *present* in the text. Moreover, the IR framework can indicate how news text can be improved so as to remove the necessity for strategic processing<sup>1</sup>, and thus reduce the potential for mystification for a reader who invests only minimum cognitive labour. Via the IR framework, the text-based principle of ‘reporting both sides’ was adjusted into a discourse-based principle, i.e., news texts should report both sides *for IR*. Thus, the value-judgement that texts should ‘report both sides’ was, in effect, grounded in a processing criterion.

### 9.3.8 *The Distinction between Text Linguistics and Discourse Analysis*

While certain linguists make little distinction between text linguistics and written discourse analysis (e.g. Stubbs, 1983), by focusing on inference generation my thesis has (inadvertently) affirmed the validity of such a distinction. Inferences such as anaphora etc, which are more *text*-based, are more reliant on *text knowledge* which is *common* amongst readers. They are in line with the definition of text in chapter 1 - 'linguistic forms in a stretch of language, and those interpretations of them which do not vary with context' (Cook, 1994: 24). However, I have concentrated on inferences whose efficacy is related to the degree of vested interest in a text and a reader's familiarity of subject matter. These inferences have been, then, dependent for their generation on a particular reader and thus more *discourse*-based.

My focus on discourse-based inferences means, then, I have not made much use of a text-linguistic metalanguage, unlike CDA. One advantage of this has been that, also unlike CDA, I have not been tempted to read off meaning from the application of such a metalanguage to text. Since this approach does not take into account the perspective of a particular reader, it gives a misleading impression of the 'ideological potential' of the text.

We have seen, though, how text-linguistic phenomena such as topicalisation and information condensation can lead to the misleading generation of causal antecedent inferences. I also highlighted how Sanford (1990) indicates the prospect of causal consequent inference generation can be dependent on topicalisation. The task ahead remains, then, to link text linguistic analysis in a systematically descriptive way to the discourse analytical perspective of IR, and in doing so, to establish a more detailed and systematic metalanguage which *bridges* text and reader. In turn, such a *bridging metalanguage* would allow a greater precision in the prediction of how the discourse derived

by IR can be said to be mystifying.

#### **9.4 Endpoint**

And to my final point, here is Widdowson (1995b: 515) in a critique of Fairclough (1992a):

On a number of occasions in the book, Fairclough makes the point that significance, ideological implication, and so on cannot be just read out of the text. That is to say, interpretation is not a semantic but a pragmatic matter: you do not read meanings out of a text, but into a text. But then we need a theory of some kind which seeks to account for these pragmatic conditions, for the textual constraints on how interpreters reduce and opt for meanings...But we get no such thing.

This thesis has been concerned with ‘how (non-analytical) interpreters reduce and opt for meanings’. I hope that it has gone some way to indicating how this happens.

#### **Notes**

1. While I have concentrated on how text can lead to mystifying discourse for a reader with little vested interest in it, one task ahead would be to make systematic predictions for the kinds of inference likely for readers with a *greater* interest in the text, e.g. language-learners taking part in a reading comprehension activity where task-based *strategic* inferences are more likely to be generated.

## BIBLIOGRAPHY

Altmann, G. and Steedman, M. (1988) 'Interaction with Context During Human Sentence Processing', *Cognition* 30: 191-238.

Althusser, L. (1971) 'Ideology and Ideological State Apparatuses', in L. Althusser (1971) *Lenin and Philosophy*, New Left Books.

Anderson, J.A. and Hinton, G.E. (1981) 'Models of Information Processing in the Brain', in G.E. Hinton and J.A. Anderson (eds) *Parallel Models of Associative Memory*, Hillsdale, NJ: Erlbaum.

Anderson, R.C., and Ortony, A. (1975) 'On Putting Apples in Bottles: A Problem of Polysemy', *Cognitive Psychology* 7: 167-180.

Anderson, A., Garrod, S.C., and Sanford, A.J. (1993) 'The Accessibility of Pronominal Antecedents as a Function of Episode Shifts in Narrative Discourse', *Quarterly Journal Of Experimental Psychology*, 35a: 427-440.

Aristotle (1933) *Metaphysics* (Trans) H. Tredennick, London: Heinemann.

Armstrong, S.L., Gleitman, L.R. and Gleitman, H. (1983) 'What Some Concepts Might Not Be', *Cognition* 13: 263-308.

Ashworth, A. (1991) *Principles of Criminal Law*, Oxford: OUP.

Ayer, A. (1982) *Philosophy in the Twentieth Century*, London: George Wiedenfield and Nicolson Ltd.

Bakhtin, M. (1981) *The Dialogical Imagination*, M. Holquist (ed), trans. C.Emerson and M.Holquist. Austin: University of Texas Press.

Bakhtin, M. (1986) *Speech Genres and Other Late Essays*, C.Emerson and M.Holquist (eds), trans.. V.W.McGee. Austin: University of Texas Press.

Barsalou, L.W. (1982) 'Context-Independent and Context-Dependent Information in Concepts' in *Memory and Cognition* 10: 82-93.

Barsalou, L.W. (1983) 'Ad-Hoc Categories', *Memory and Cognition* 11: 211-27.

Barsalou, L.W. (1987) 'The Instability of Graded Structure: Implications for the Nature of Concepts', in U. Neisser (ed) *Concepts and Conceptual Development: Ecological and Intellectual Factors in Categorization*, pp.101-140. Cambridge: CUP.

Barsalou, L.W. (1989) 'Intra-Concept Similarity and its Implications for Inter-Concept Similarity', in S. Vosniadou and A. Ortony (eds) *Similarity and Analogical Reasoning*, Cambridge: CUP.

Barton, S.B. and Sanford, A.J. (1993) 'A Case-Study of Pragmatic Anomaly Detection:

Relevance-Driven Cohesion Patterns', *Memory and Cognition* 21: 477-487.

Bates, E. and MacWhinney, B. (1982) 'Functionalist Approaches to Grammar', in E. Wanner and L.R. Gleitman (eds) *Language Acquisition: The State of the Art*, pp.173-218. New York: CUP.

Bechtel, W. (1990) 'Connectionism and the Philosophy of Mind: an Overview', in W.G. Lycan (ed) *Mind and Cognition: A Reader*, pp.252-273. Oxford: Blackwell.

Bechtel, W. (1996a) 'What Knowledge Must Be in the Head in Order to Acquire Language?', in B.M. Velichkovsky and D.M. Rumbaugh (eds) *Communicating Meaning: The Evolution And Development Of Language*, pp.45-78. Mahwah, NJ: Lawrence Erlbaum Associates.

Bechtel, W. (1996b) 'What Should A Connectionist Philosophy Of Science Look Like?', in R.N. McCauley (1996) *The Churchlands and their Critics*, pp.121-144. Oxford: Blackwell.

Bechtel, W. and Abrahamsen, A. (1991) *Connectionism and the Mind: An Introduction To Parallel Processing In Networks*, Oxford: Blackwell.

Bell, A. (1991) *The Language of News Media*, Oxford: Blackwell.

Berlin, B., Breedlove, D. and Raven, P. (1974) *Principles of Tzeltal Plant Classification*, New York: Academic Press.

Bierwisch, M. (1967) 'Some Semantic Universals of German Adjectivals', *Foundations of Language* 3: 1-36.

Bierwisch, M. (1970) 'Semantics', in J. Lyons (ed) *New Horizons in Linguistics, vol. 1*, pp.166-84. Harmondsworth: Penguin.

Black, J.B. and Bower, G.H. (1980) 'Story Understanding and Problem Solving', *Poetics* 9: 223-250.

Bloom, C.P., Fletcher, C.R., van den Broek, P., Reitz, L., and Shapiro, B.P. (1990) 'An On-line Assessment of Causal Reasoning During Comprehension', *Memory and Cognition* 18: 65-71.

Bower, G.H., Black, J.B., and Turner, T.J. (1979) 'Scripts in Memory for Text', *Cognitive Psychology* 11: 177-220.

Boyd, R. (1993) 'Metaphor and Theory Change: What Is 'Metaphor' A Metaphor For?', in A. Ortony (ed) *Metaphor and Thought*, 2nd ed., pp.481-532. Cambridge: CUP.

Bransford, J.D., Barclay, J.R., and Franks, J.J. (1972) 'Sentence Memory: A Constructive Versus Interpretive Approach', *Cognitive Psychology* 3: 193-209.

Bransford, J.D. and Johnson, M.K. (1973) 'Considerations of Some Problems of Comprehension', in W.Chase (ed) *Visual Information Processing*, pp.383-438. New York:

Academic Press.

Bredart, S. and Modolo, K. (1988) 'Moses Strikes Again: Focalization Effect on a Semantic Illusion', *Acta Psychologica* 67: 135-144.

Brown, E.K. (1991) 'Transformational-Generative Grammar', in K. Malmkjaer (ed) *The Linguistics Encyclopedia*, pp.482-497. London: Routledge.

Brown, R. (1958) 'How Shall A Thing Be Called?' in *Psychological Review* 65: 14-21.

Brown, R. (1965) *Social Psychology*, New York: Free Press.

Brown, G. and Yule, G. (1983) *Discourse Analysis*, Cambridge: CUP.

Bruner, J. (1973) 'Beyond the Information Given' in J. Anglin (ed) *Studies in the Psychology of Knowing*, New York: W.W. Norton.

Caldas-Coulthard, C. and Coulthard, M. (1996) *Texts and Practices: Readings in Critical Discourse Analysis*, London: Routledge.

Cameron, D. (ed) (1998) *The Feminist Critique Of Language: A Reader*, 2nd ed. London: Routledge.

Carnap, R. (1928 / 67) [Der logische Aufbau der Welt] *The Logical Structure of The World: Pseudo-Problems In Philosophy*, trans. by A. Rolf. George, London: Routledge and Kegan.

Carnap, R. (1935) *Philosophy and Logical Syntax*, London: Kegan Paul, Trench, Trubner and Co. Ltd.

Carnap, R. (1937) *The Logical Syntax of Language*, London: Kegan Paul, Trench, Trubner and Co.

Carter, R. (1997) *Investigating English Discourse: Language, Literacy and Literature*, London: Routledge.

Charniak, E. (1983) 'Passing Markers: A Theory of Contextual Influence in Language Comprehension', *Cognitive Science* 7: 171-190.

Chilton, P. (1988) *Orwellian Language and the Media*, London, Pluto Press.

Chomsky, N. (1957) *Syntactic Structures*, The Hague: Mouton.

Chomsky, N. (1959) 'A Review of B.F. Skinner's Verbal Behaviour', *Language* 35: 26-58.

Chomsky, N. (1965) *Aspects of A Theory of Syntax*, Cambridge, Mass: The MIT Press.

Chomsky, N. and Halle, M. (1968) *The Sound Pattern of English*, New York: Harper and Row.

Chouliaraki, L. and Fairclough, N. (1999) *Discourse in Late Modernity: Rethinking Critical*



*Discourse Analysis*, Edinburgh: Edinburgh University Press.

Churchland, P.M. (1988) *Matter and Consciousness* (revised edition), Cambridge, Mass.: MIT Press.

Churchland, P.S. (1986) *Neurophilosophy*, Cambridge, Mass: MIT Press.

Churchland, P.M. and Churchland, P.S. (1996) 'Connectionism as Psychology', in R.N. McCauley (ed) *The Churchlands and Their Critics*, pp.232-238. Oxford: Blackwell.

Clark, A. (1989) *Microcognition: Philosophy, Cognitive Science and Parallel Distributed Processing*, Cambridge, Mass: MIT Press.

Clark, A. (1993) *Associative Engines: Connectionism, Concepts, and Representational Change*, Cambridge, Mass.: MIT Press.

Clark, A. (1996) 'Introduction', in A. Clark and P.J.R. Millican (eds) *Connectionism, Concepts, and Folk Psychology: The Legacy Of Alan Turing*, vol. 2, pp.1-6. Oxford: Clarendon Press.

Clark, K. (1992) 'The Linguistics of Blame', in M.J. Toolan (ed) *Language, Text and Context*, pp.208-224. London: Routledge,.

Cook, G. (1994) *Discourse and Literature: The Interplay of Form and Mind*, Oxford: OUP.

Cooper, R. (1996) 'Explanation and Simulation in Cognitive Science', in D.W. Green (ed) *Cognitive Science: An Introduction*, pp.23-52. Oxford: Blackwell.

Cotter, C.A. (1984) 'Inferring Indirect Objects in Sentences: Some Implications For the Semantics of Verbs', *Language and Speech* 27: 25-45.

Cottrell, G.W and Small, S.L (1983) 'A Connectionist Scheme for Modelling Word Sense Disambiguation', *Cognition And Brain Theory* 6: 89-120.

Crain, S. and Steedman, M. (1985) 'On not Being Led up the Garden Path: the Use of Context by the Psychological Syntax Processor', in D. Dowty, L., Karttunen and A. Zwicky (eds) (1985) *Natural Language Parsing*, pp.320-354. Cambridge: CUP.

Cruse, D.A. (1977) 'The Pragmatics of Lexical Specificity', *Journal Of Linguistics* 13: 153-64.

Damasio, A. (1994) *Descartes' Error: Emotion, Reason, and the Human Brain*, New York: Grosset-Putnam.

Dell, G.S., McKoon, G., and Ratcliff, R. (1983) 'The Activation of Antecedent Information During the Processing of Anaphoric Reference in Reading', *Journal of Verbal Learning and Verbal Behavior* 22: 121-132.

DeJong, G. (1979) 'Prediction and Substantiation: A New Approach to Natural Language Processing', *Cognitive Science* 3: 251-273.

- Dennett, D.C. (1978) *Brainstorms: Philosophical Essays On Mind And Psychology*, Cambridge, Mass.: MIT Press.
- Dyer, M.G. (1983) *In-Depth Understanding: A Computer Model of Integrated Processing for Narrative Comprehension*, Cambridge, MA: MIT Press.
- Eccles, J. C. (1977) Part II of K.Popper, *The Self And Its Brain*, Berlin: Springer-Verlag.
- Eco, U. (1995) *The Search for the Perfect Language*, London: Fontana.
- Edelman, G. (1992) *Bright Air, Brilliant Fire*, London: Penguin.
- Ellis, D.G. (1992) *From Language to Communication*, Hillsdale, NJ: Lawrence Erlbaum Associates.
- Ellis, J.M. (1993) *Language, Thought and Logic*, Evanston, Illinois: Northwestern University Press.
- Elman, J.L. (1990) 'Representation and Structure in Connectionist Models' in G.T.M. Altmann (ed) *Cognitive Models Of Speech Processing: Psycholinguistic And Computational Perspectives*, pp.345-382. Cambridge, Mass.: The MIT Press.
- Elman, J.L. (1992) 'Grammatical Structure and Distributed Representations', in S. Davis (ed) *Connectionism: Theory And Practice*, pp.138-178. Oxford: OUP.
- Erickson, T.A. and Mattson, M.E. (1981) 'From Words to Meaning: A Semantic Illusion', *Journal of Verbal Learning and Verbal Behaviour* 20: 540-552.
- Eysenck, M. W. and Keane, M.T. (1995) *Cognitive Psychology: A Student's Handbook*, 3rd ed. Hove: Psychology Press.
- Fairclough, N. (1989) *Language and Power*, London: Longman.
- Fairclough, N. (1992a) *Discourse and Social Change*, Oxford: Polity
- Fairclough, N. (1992b) *Critical Language Awareness*, London: Longman.
- Fairclough, N. (1995a) *Media Discourse*, London: Edward Arnold.
- Fairclough, N. (1995b) *Critical Discourse Analysis: The Critical Study Of Language*, London: Longman.
- Fairclough, N. (1996) 'A Reply to Henry Widdowson's Discourse Analysis: A Critical View' in *Language and Literature* 5 (1): 49-56.
- Feldman, J.A. and Ballard, D.H. (1982) 'Connectionist Models and Their Properties', *Cognitive Science* 6: 205-54.
- Fletcher, C.R. (1986) 'Strategies for the Allocation of Short-term Memory During

Comprehension', *Journal of Memory and Language* 27: 43-58.

Flores d'Arcais, G.B. (1987) 'Syntactic Processing During Reading for Comprehension', in M. Coltheart (1987) *Attention And Performance XII: The Psychology of Reading*, pp.619-633. Hillsdale, NJ: Lawrence Erlbaum Associates,.

Fodor, J. (1975) *The Language of Thought*, New York: Crowell.

Fodor, J. (1983) *Modularity of Mind*, Cambridge, Mass.: The MIT Press.

Fodor, J.A. and Garrett, M. (1966) 'Some Reflections on Competence and Performance', in J. Lyons and R.J. Wales (1966) *Psycholinguistic Papers*, Edinburgh: Edinburgh University Press.

Fodor, J.A. and Garrett, M. (1967) 'Some Syntactic Determinants of Sentential Complexity', *Perception and Psychophysics* 2: 289-96.

Fodor, J. and Pylyshyn, Z.W. (1988) 'Connectionism and Cognitive Architecture: A Critical Analysis', *Cognition* 28: 3-71.

Fodor, J.A., Bever, T.G., and Garrett, M.F. (1974) *The Psychology Of Language: An Introduction To Psycholinguistics and Generative Grammar*, New York, McGraw-Hill.

Fodor, J.D., Fodor, J.A., and Garrett, M.F. (1975) 'The Psychological Unreality of Semantic Representations', *Linguistic Inquiry* 6 (4): 515-531.

Foucault, M. (1972) *The Archaeology of Knowledge*, trans. A.M. Sheridan-Smith. London: Tavistock.

Fowler, R. (1986) *Linguistic Criticism*, Oxford: OUP.

Fowler, R. (1988) 'Notes On Critical Linguistics' in C.R. Caldas-Coulthard and M. Coulthard (eds) (1996) *Texts and Practices: Readings in Critical Discourse Analysis*, pp.3-14. London: Routledge.

Fowler, R. (1991) *Language in the News: Discourse and Ideology in the Press*, London: Routledge.

Fowler, R. and Kress, G. (1979a) 'Rules and Regulations' in R. Fowler et al.: 26-45.

Fowler, R. and Kress, G. (1979b) 'Critical Linguistics' in R. Fowler et al.: 185-213.

Fowler, R., Hodge, R., Kress, G. and Trew, T. (1979) *Language and Control*, London: Routledge and Kegan Paul.

Frazier, L. (1987) 'Sentence Processing: A tutorial Review', in M. Coltheart (ed) *Attention and Performance XII: The Psychology of Reading*, pp.559-586. Hillsdale, NJ: Lawrence Erlbaum Associates.

Frazier, L. and Fodor, J.D (1978) 'The Sausage Machine: A New Two-Stage Parsing

Model', *Cognition* 6: 291-325.

Further Education Unit (1987) *Relevance, Flexibility and Competence*, London: Further Education Unit.

Gardner, H. (1987) *The Mind's New Science: A History of the Cognitive Revolution*, New York: Basic Books.

Garnham, A. (1981) 'Anaphoric Reference to Instances, Instantiated and Non-Instantiated Categories: A Reading-Time Study', *British Journal Of Psychology* 72: 377-84.

Garnham, A. (1985) *Psycholinguistics: Central Topics*, London: Methuen.

Garrod, S., O'Brien, E.J., Morris, R.K., and Rayner, K. (1990) 'Elaborative Inferences as an Active or Passive Process', *Journal of Experimental Psychology: Learning, Memory, and Cognition* 16: 250-257.

Garrod, S.C. and Sanford, A.J. (1982) 'Bridging Inferences in the Extended Domain of Reference', in J.Long and A. Baddeley (eds) *Attention and Performance IX*, pp.331-346. Hillsdale, NJ: Erlbaum.

Garrod, S.C. and Sanford, A.J. (1988) 'Thematic Subjecthood and Cognitive Constraints on Discourse Structure', *Journal Of Pragmatics* 12: 519-534.

Garrod, S.C. and Sanford, A.J. (1994) 'Resolving Sentences in a Discourse Context: How Discourse Representation Affects Language Understanding', in M.A. Gernsbacher: 675-698.

Gernsbacher, M.A. (1994) *Handbook of Psycholinguistics*, San Diego, CA: Academic Press.

Gibbs, R.W. (1994) *The Poetics of Mind: Figurative Thought, Language and Understanding*, Cambridge: CUP.

Gough, V. and Talbot, M. (1996) "'Guilt Over Games Boys Play': Coherence as a Focus for Examining the Constitution of Heterosexual Subjectivity on a Problem Page', in C. Caldas-Coulthard and M. Coulthard (eds) *Texts and Practices: Readings in Critical Discourse Analysis*, pp. 214-230. London: Routledge.

Graddol, D. (1996) 'The Development of Scientific English' in D. Graddol, D. Leith, J. Swann (eds) *English, Diversity and Change*, pp. 171-179. London: Routledge.

Graesser, A.C. (1981) *Prose Comprehension beyond the Word*, New York: Springer-Verlag.

Graesser, A.C. and Clark, L.F. (1985) *Structures and Procedures of Implicit Knowledge*, Norwood, N.J.: Ablex.

Graesser, A.C., Singer, M., and Trabasso, T. (1994) 'Constructing Inferences During Narrative Text Comprehension', *Psychological Review* 101 (3): 371-395.

Grosjean, F. (1980) 'Spoken Word Recognition Processes and the Gating Paradigm', *Perception and Psychophysics* 28: 267-283.

Gumenik, W.E. (1979) 'The Advantage of Specific Terms over General Terms as Cues for Sentence Recall: Instantiation or Retrieval?' *Memory And Cognition* 7: 240-4

Hacking, I. (1975) *Why Does Language Matter To Philosophy?*, Cambridge: CUP.

Halliday, M.A.K. (1973) *Explorations in The Functions of Language*, London: Edward Arnold.

Halliday, M.A.K. (1993) 'On the language of physical science', in M.A.K. Halliday and J.R. Martin (eds) *Writing Science: Literacy and Discursive Power*, Basingstoke, Falmer.

Hammersley, M. (1996) 'On The Foundations Of Critical Discourse Analysis', Occasional Paper 42. Centre for Language in Education, University of Southampton.

Hampton, J.A. (1981) 'An Investigation Of The Nature Of Abstract Concepts', *Memory and Cognition* 9: 149-156.

Harder, P. (1997) *Functional Semantics: A Theory of Meaning, Structure and Tense in English* Berlin: Mouton de Gruyter.

Harley, T. (1995) *The Psychology of Language*, Hillsdale, NJ: Lawrence Erlbaum Associates.

Harman, G. (1986) *Change in View*, Cambridge, MA: MIT Press.

Harris, C.L. (1990) 'Connectionism and Cognitive Linguistics', *Connection Science* 2 (1).

Harris, Z.S. (1952) 'Discourse Analysis', *Language* 28: 18-30.

Hobbes, T. ([1651] 1962) *Leviathan: Or the Matter, Forme and Power of a Commonwealth Ecclesiastical and Civil*, London: Collier Books.

Hodge, R. and Kress, G. (1974) 'Transformations, Models and Processes: Towards a Useable Linguistics', *Journal of Literary Semantics* 4 (1): 4-18.

Hodge, R. and Kress, G. (1993) *Language as Ideology*, 2nd ed. London: Routledge.

Hoey, M. (1983) *On the Surface of Discourse*, London: George Allen and Unwin.

Hopper, P.J. and Thompson, S.A. (1985) 'The Iconicity of the Universal Categories 'Noun' and 'Verb'', in J. Haiman (ed) (1985) *Iconicity in Syntax*, Amsterdam: John Benjamins.

Johnson, M. (1987) *The Body in the Mind*, Chicago: University of Chicago Press.

Johnson-Laird, P.N. (1980) 'Mental Models in Cognitive Science', *Cognitive Science* 4: 71-115.

Johnson-Laird, P.N. (1983) *Mental Models*, Cambridge, MA: Harvard University Press.

Just, M.A., and Carpenter, P.A. (1980) 'A Theory Of Reading: From Eye Fixations To Comprehension', *Psychological Review* 87: 329-354.

Katz, J.J. and Fodor, J.A. (1963) 'The Structure of A Semantic Theory', *Language* 39: 170-210

Katz, J.J. and Postal, P.M. (1964) *An Integrated Theory of Linguistic Descriptions*, Cambridge, Mass.: MIT Press.

Kay, P. (1971) 'Taxonomy And Semantic Contrast', *Language* 47: 866-87.

Keefe, D.E. and McDaniel, M.A. (1993) 'The Time Course and Durability of Predictive Inferences', *Journal of Memory and Language* 32: 446-463.

Keenan, J. M., MacWhinney, B. and Mayhew, D. (1977) 'Pragmatics in Memory: A Study of Natural Conversation', *Journal of Verbal Learning and Verbal Behaviour* 16: 549-60.

Keenan, J.M., Baillet, S.D. and Brown, P. (1984) 'The Effects of Causal Cohesion on Comprehension and Memory', *Journal of Verbal Learning and Verbal Behavior* 23: 115-126.

Kimball, J. (1973) 'Seven Principles of Surface Structure Parsing in Natural Language', *Cognition* 2: 15-47.

Kintsch, W. (1988) 'The Role of Knowledge in Discourse Comprehension: A Constructive-Integration Model', *Psychological Review* 95: 163-182.

Kittay, Eva (1987) *Metaphor: its Cognitive Force and Linguistic Structure*, Oxford: Clarendon Press.

Kress, G. (1989a) *Linguistic Processes in Sociocultural Practice*, Victoria: Deakin University Press.

Kress, G. (1989b) 'History and Language: Towards a Social Account of Language Change', *Journal Of Pragmatics* 13: 445-66

Kress, G. (1993) 'Against Arbitrariness: The Social Production of the Sign as a Foundational Issue in Critical Discourse Analysis', *Discourse And Society*, vol 4(2): 169-191.

Kress, G. and Hodge, R. (1979) *Language as Ideology*, 1st ed. London: Routledge and Kegan Paul.

Kuhn, T. (1970) *The Structure of Scientific Revolutions*, 2nd ed. Chicago: University Of Chicago Press.

Kuhn, T.S. (1993) 'Metaphor in Science', in A. Ortony (ed) *Metaphor and Thought*, 2nd ed. pp. 533-42. Cambridge: CUP.

Lakoff, G. (1987a) *Women, Fire and Dangerous Things*, Chicago: University of Chicago Press.

Lakoff, G. (1987b) 'Cognitive Models and Prototype Theory', in U. Neisser (ed) (1987) *Concepts and Conceptual Development: Ecological and Intellectual Factors in Categorization*, pp. 63-100. Cambridge: CUP.

Lakoff, G. and Johnson, M. (1980) *Metaphors We Live By*, Chicago: University of Chicago Press.

Langacker, R.W. (1987a) *Foundations Of Cognitive Grammar: Volume I*, Stanford, CA, Stanford University Press.

Langacker, R.W. (1987b) 'The Cognitive Perspective', *Centre for Research in Language Newsletter, University of California, San Diego* 1 (3): 3-15.

Langacker, R.W. (1997) 'The Contextual Basis of Cognitive Semantics', in J. Nuyts and E. Pederson (eds) (1997) *Language and Conceptualisation*, pp. 229- 252. Cambridge: CUP.

Lee, D. (1992) *Competing Discourses: Perspective and Ideology in Language*, London: Longman.

Leech, G. (1981) *Semantics*, 2nd ed. Harmondsworth: Penguin.

Leech, G.N. and Short, M.H. (1981) *Style in Fiction: A Linguistic Introduction to English Fictional Prose*, London: Longman.

Levinson, S.C (1983) *Pragmatics*, Cambridge: CUP.

Lucas, M.M. Tanenhaus, M.K., and Carlson, G.N. (1990) 'Levels of Representation in the Interpretation of Anaphoric Reference and Instrument Inference', *Memory and Cognition* 18 (6): 611-631.

Lucy, J.A. (1992) *Language, Diversity and Thought: A Reformulation of the Linguistic Relativity Hypothesis*, Cambridge: CUP.

Lyons, J. (1991) *Chomsky*, 3rd ed. London: Fontana.

MacDonald, C. and MacDonald, G. (1995) *Connectionism: Debates On Psychological Explanation, Volume 2*, Oxford: Blackwell.

McClelland, J.L. and Kawamoto, A.H. (1986) 'Mechanisms of Sentence Processing: Assigning Roles to Constituents of Sentences', in J.L. McClelland, D.E. Rumelhart and the PDP Research Group (eds): 272-325.

McClelland, J.L, Rumelhart, D.E. and the PDP Research Group (eds) (1986) *Parallel Distributed Processing: Explorations In the Microstructure Of Cognition, Vol.2: Psychological and Biological Models*, Cambridge, MA, MIT Press.

McClelland, J.L., St. John, M.F. and Taraban, R. (1989) 'Sentence Comprehension: A Parallel Distributed Processing Approach', *Language and Cognitive Processes* 4: 287-335.

- McKoon, G. and Ratcliff, R. (1981) 'The Comprehension Processes and Memory Structures Involved in Instrumental Inference', *Journal of Verbal Learning and Verbal Behaviour* 20: 671-682.
- McKoon, G. and Ratcliff, R. (1986) 'Inferences about Predictable Events', *Journal of Experimental Psychology: Learning, Memory and Cognition* 12: 82-91.
- McKoon, G. and Ratcliff, R. (1989a) 'Semantic Associations and Elaborative Inference', *Journal of Experimental Psychology: Learning, Memory and Cognition* 15 (2): 326-338.
- McKoon, G. and Ratcliff, R. (1989b) 'Inferences about Contextually Defined Categories', *Journal of Experimental Psychology: Learning, Memory and Cognition* 15 (6): 1134-1146.
- McKoon, G. and Ratcliff, R. (1992) 'Inference During Reading', *Psychological Review* 99: 440-66.
- Marslen-Wilson, W. and Tyler, L.K. (1980) 'The Temporal Structure of Spoken Language Understanding', *Cognition* 8: 1-71.
- Martin, J.R. (1989) *Factual Writing: Exploring and Challenging Social Reality*, 2nd ed. Oxford: OUP.
- Miikkulainen, R. (1993) *Subsymbolic Natural Language Processing*, Cambridge, Mass: MIT Press.
- Miller, G.A. (1956) 'The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information', *Psychological Review* 63: 81-97.
- Miller, G.A. (1962) 'Some Psychological Studies of Grammar', *American Psychologist* 1: 748-62.
- Miller, G.A. and McKean, K. (1964) 'A Chronometric Study of Some Relations between Sentences', *Quarterly Journal of Experimental Psychology*, 16: 297-308.
- Montgomery, M. (1995) *An Introduction to Language and Society*, 2nd ed. London: Routledge.
- Montgomery, M., Tolson, A. and Garton, G. (1989) 'Media Discourse in the 1987 General Election: Ideology, Scripts and Metaphors', *English Language Research* 3: 173-204.
- Morrow, D.G. (1985) 'Prominent Characters and Events Organize Narrative Understanding', *Journal of Memory and Language* 24: 304-319.
- Myers, J.L., Shinjo, M. and Duffy, S.A. (1987) 'Degree of Causal Relatedness and Memory', *Journal of Memory and Language* 26: 453-465.
- Newell, A. and Simon, H.A. (1972) *Human Problem Solving*, Englewood Cliffs, N.J.: Prentice Hall.
- Nida, E.A. (1975) *Componential Analysis of Meaning*, The Hague: Mouton.



Oaksford, M. and Chater, N. (1991) 'Against Logicist Cognitive Science', *Mind Lang* 6: 1-38.

O'Brien, E.J., Shank, D.M, Myers, J.L., and Rayner, K. (1988) 'Elaborative Inferences During Reading: Do They Occur On-Line?', *Journal of Experimental Psychology: Learning, Memory and Cognition*, 14: 410-420.

Ogborn, J., Kress, G., Martins, I, and McGillicuddy, K. (1996) *Explaining Science in the Classroom*, Buckingham: Open University Press.

O'Halloran, K.A. (1997) 'Why Whorf Has Been Misconstrued in Stylistics and Critical Linguistics', *Language and Literature* 6 (3).

Orwell, G. (1946) 'Politics And The English Language' in G. Orwell (1969: 143-157).

Orwell, G. (1969) *Inside the Whale and Other Essays*, Harmondsworth: Penguin Books.

Outhwaite, W. (1987) *New Philosophies Of Social Science: Realism, Hermeneutics And Critical Theory*, Basingstoke: Macmillan.

Pennycook, A. (1990) 'Towards a Critical Applied Linguistics for the 1990s', *Issues in Applied Linguistics*, 1 (1): 8-28.

Pennycook, A. (1994) 'Incommensurable Discourses', *Applied Linguistics*, 15 (2): 115-138.

Potts, G.R., Keenan, J.M., and Golding, J.M. (1988) 'Assessing The Occurrence of Elaborative Inferences: Lexical Decision Versus Naming', *Journal Of Memory And Language* 27: 399-415.

Putnam, H. (1967) 'The Nature of Mental States', in Putnam, H. (1975).

Putnam, H. (1975) *Philosophical Papers, Vol.2, Mind, Language and Reality*, Cambridge: CUP.

Pylyshyn, Z. (1984) *Computation and Cognition*, Cambridge, Mass.: MIT Press.

Quine, W. (1953) 'Two Dogmas of Empiricism', in *From a Logical Point of View*, Cambridge: Harvard University Press.

Quine, W.V.O (1960) *Word and Object*, Cambridge, Mass.: MIT Press.

Rayner, K., Carlson, M. and Frazier, L. (1983) 'The Interaction of Syntax and Semantics During Sentence Processing: Eye Movements in the Analysis Of Semantically Biased Sentences', *Journal Of Verbal Learning and Verbal Behaviour* 22: 358-374.

Reder, L.M. and Kusbit, G.W. (1991) 'Locus of the Moses Illusion: Imperfect Encoding, Retrieval, or Match?', *Journal of Memory and Language* 30: 385-406.

Regier, C. (1975) *Conceptual Memory and Inference: Conceptual Information Processing*,

Amsterdam: North-Holland.

Regier, T. (1996) *The Human Semantic Potential: Spatial Language and Constrained Connectionism*, Cambridge, Mass.: MIT Press.

Richardson, K. (1987) 'Critical Linguistics and Textual Diagnosis', *Text* 7 (2): 145-163.

Rifkin, A. (1985) 'Evidence for a Basic Level in Event Taxonomies', *Memory And Cognition* 13: 538-56.

Rorty, R. (1980) *Philosophy and the Mirror of Nature*, Oxford: Blackwell.

Rosch, E. (1975a) 'Cognitive Reference Points', *Cognitive Psychology* 7: 532-47.

Rosch, E. (1975b) 'Cognitive Representations of Semantic Categories', *Journal Of Experimental Psychology, General*, 104: 193-233.

Rosch, E. (1978) 'Principles of Categorization' in E. Rosch and B.B. Lloyd (1978) *Cognition and Categorization*, pp. 27-48. Hillsdale, NJ: Lawrence Erlbaum Associates.

Rosch, E. (1981) 'Prototype Classification and Logical Classification: The Two Systems' in E. Scholnick, (1983) (ed) *New Trends in Cognitive Representation: Challenges to Piaget's Theory*, pp. 73-86. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Rosch, E., Mervis, C.B., Gray, W.D., Johnson, D.M. and Boyes-Braem, P. (1976) 'Basic Objects In Natural Categories', *Cognitive Psychology* 8: 382:439.

Roth, E. M. and Shoben, E.J. (1983) 'The Effect of Context on the Structure of Categories', *Cognitive Psychology* 15: 346-78.

Rumelhart, D.E. (1975) 'Notes on a Schema for Stories' in D.G. Bobrow and A.M. Collins (eds) *Representation and Understanding: Studies in Cognitive Science*, San Diego, C.A: Academic Press.

Rumelhart, D.E. (1979) 'Some Problems with the Notion of Literal Meanings' in A. Ortony (ed) *Metaphor and Thought*, Cambridge: CUP.

Rumelhart, D.E., Hinton, G. and McClelland, J.L. (1986) 'A General Framework for Parallel Distributed Processing', in D.E. Rumelhart, J.L. McClelland and the PDP Research Group: 45-76.

Rumelhart, D.E. and McClelland, J.L. (1986) 'PDP Models and General Issues in Cognitive Science' in D.E. Rumelhart, J.L. McClelland and PDP Research Group (eds): 110-146.

Rumelhart, D.E, McClelland, J.L. and the PDP Research Group (eds) (1986) *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol 1: Foundations*, Cambridge, MA, MIT Press.

Rumelhart, D.E. and Norman, D.A. (1982) 'Simulating a Skilled Typist: A Study of Skilled Cognitive-Motor Performance', *Cognitive Science* 6: 1-36.

- Russell, B. and Whitehead, A.N. (1910) *Principia Mathematica*, Cambridge: CUP.
- St. John, M.F. (1992) 'The Story Gestalt: A Model of Knowledge Intensive Processes in Text Comprehension', *Cognitive Science* 16: 271-306.
- St. John, M. F. and McClelland, J.L. (1992) 'Parallel Constraint Satisfaction as a Comprehension Mechanism', in R.G. Reilly and N.E. Sharkey (eds) *Connectionist Approaches to Natural Language Processing*, pp.97-136. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Salasoo, A. and Pisoni, D.B. (1985) 'Interaction of Knowledge Sources in Spoken Word Identification', *Journal of Memory and Language* 24: 210-231.
- Sanford, A.J. (1987) *The Mind of Man*, New Haven: Yale University Press.
- Sanford, A.J. (1990) 'On the Nature of Text-Driven Inference', in D.A. Balota, G.B. Flores D'Arcais and K. Rayner (eds) (1990) *Comprehension Processes in Reading*, pp. 515-535. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Sanford, A. J. and Garrod, S.C. (1981) *Understanding Written Language: Explorations of Comprehension Beyond the Sentence*, Chichester: John Wiley and Sons.
- Sanford, A.J. and Garrod, S.C. (1994) 'Selective Processing in Text Understanding', in M.A. Gernsbacher (1994) *Handbook of Psycholinguistics*, pp. 699-719. San Diego, CA: Academic Press.
- Sanford, A.J., Moar, K. and Garrod, S.C. (1988) 'Proper Names as Controllers of Discourse Focus', *Language and Speech*, 31: 43-56.
- Savin, H. and Perchonock, E. (1965) 'Grammatical Structure and the Immediate Recall of English Sentences', *Journal of Verbal Learning and Verbal Behaviour* 4: 348-53.
- Schank, R.C. (1980) 'An Artificial Intelligence Perspective on Chomsky's View of Language', *The Behavioral and Brain Sciences* 3: 35-7.
- Schank, R.C. (1986) *Explanation Patterns: Understanding Mechanically and Creatively*, Hillsdale, NJ: Erlbaum.
- Schank, R.C. and Abelson, R. (1977) *Scripts, Plans, Goals and Understanding*, Hillsdale, N J: Lawrence Erlbaum.
- Schön, D. (1993) 'Generative Metaphor: A Perspective on Problematizing Social Policy', in A. Ortony (ed) (1993) *Metaphor and Thought*, 2nd ed., pp. 137-63. Cambridge: CUP.
- Schopman, J. and Shawky, A. (1996) 'Remarks on the Impact of Connectionism on our Thinking about Concepts', in A. Clark and P.J.R Millican (eds) *Connectionism, Concepts, and Folk Psychology: The Legacy of Alan Turing*, vol. 2, pp. 67-74. Oxford: Clarendon Press.
- Searle, J. (1995) *The Construction of Social Reality*, London: The Penguin Press.

- Sejnowski, T.J. and Rosenberg, C.R. (1987) 'Parallel Networks that Learn to Pronounce English Text', *Complex Systems* 1: 145-68.
- Sharrock, W.W. and Anderson, D.C. (1981) 'Language, Thought and Reality, Again', *Sociology* 15: 287-93.
- Shepherd, V. (1994) *Literature about Language*, London: Routledge.
- Skinner, B.F. (1957) *Verbal Behaviour*, New York: Appleton-Century-Crofts.
- Skinner, B.F. (1976) *About Behaviourism*, New York: Knopf.
- Simpson, P. (1993) *Language, Ideology, and Point of View*, London: Routledge.
- Singer, M. (1979a) 'Processes of Inference in Sentence Encoding', *Memory and Cognition* 7: 192-200.
- Singer, M. (1979b) 'Temporal Locus of Inference in the Comprehension of Brief Passages: Recognizing and Verifying Implications about Instruments', *Perceptual and Motor Skills* 49: 539-550.
- Singer, M. (1980) 'The Role of Case-Filling Inferences in the Coherence of Brief Passages', *Discourse Processes* 3: 185-201.
- Singer, M. (1990) *Psychology of Language: An Introduction to Sentence and Discourse Processes*, Hillsdale, N.J: Erlbaum.
- Singer, M. (1994) 'Discourse Inference Processes' in M.A. Gernsbacher: 479-515.
- Slobin, D. (1982) 'Universal and Particular in the Acquisition of Language', in E. Wanner and L.R. Gleitman (eds) *Language Acquisition: the State of the Art*, pp. 128-170. Cambridge: CUP.
- Smart, J.J.C. (1959) 'Sensations and Brain Processes', *Philosophical Review* 68: 141-56.
- Smolensky, P. (1987) 'Connectionist AI, and the brain', *Artificial Intelligence Review* 1: 95-109.
- Smolensky, P. (1988) 'On the Proper Treatment of Connectionism', in C. MacDonald and G. MacDonald (eds) (1995) *Connectionism: Debates On Psychological Explanation, Volume 2*, pp. 28-89. Oxford: Blackwell.
- Sontag, S. (1988) *Aids and Its Metaphors*, London: Penguin.
- Sperber, D. and Wilson, D. (1995) *Relevance: Communication and Cognition*, 2nd ed. Oxford: Blackwell.
- Stevenson, R.J. (1993) *Language, Thought and Representation*, Chichester: John Wiley and Sons.

Stillings, N.A., Weisler, S.E., Chase, C.H., Feinstein, M.H., Garfield, J.L., and Rissland, E.L. (1995) *Cognitive Science: An Introduction*, 2nd ed. Cambridge, Mass.: MIT Press.

Stubbs, M. (1983) *Discourse Analysis*, Oxford: Blackwell.

Stubbs, M. (1997) 'Whorf's Children: Critical Comments on Critical Discourse Analysis', in A. Ryan and A. Wray (eds) *Evolving Models of Language: British Studies in Applied Linguistics* 12: 100-116.

Swinburne, R. (1986) *The Evolution of the Soul*, Oxford: OUP.

Taraban, R. and McClelland, J.L. (1988) 'Constituent Attachment and Thematic Role Assignment in Sentence Processing: Influences of Content-Based Expectations', *Journal of Memory and Language* 27: 597-632.

Taylor, J. (1995) *Linguistic Categorisation*, 2nd ed. Oxford: OUP.

Thomason, R. (ed) (1974) *Formal Philosophy: Selected Papers of Richard Montague*, New Haven: Yale University Press.

Thompson, J. (1984) *Studies in the Theory of Ideology*, Oxford: Polity Press.

Toolan, M.J. (1988) *Narrative: A Critical Linguistic Introduction*, London: Routledge.

Touretzky, D.S. and Geva, S. (1987) A Distributed Connectionist Representation for Concept Structures. Paper Presented to the 9th Annual Conference of the Cognitive Science Society, Seattle.

Trabasso, T. and Sperry, L. (1985) 'Causal Relatedness and Importance of Story Events', *Journal of Memory and Language* 24: 595-611.

Trabasso, T., van den Broek, P., and Suh, S.Y. (1989) 'Logical Necessity and Transitivity of Causal Relations in Stories', *Discourse Processes* 12: 1-26.

Trask, R.L. (1999) *Key Concepts in Language and Linguistics*, London: Routledge.

Trew, T. (1979) 'Theory and Ideology at Work' in R. Fowler et al.: 94 - 116.

Turing, A.M. (1950) 'Computing Machinery and Intelligence', *Mind* 59: 433-60.

Ungerer, F. and Schmid, H.-J. (1996) *An Introduction to Cognitive Linguistics*, London: Longman.

van den Broek, P. (1990) 'Causal Inferences and the Comprehension of Narrative Text' in A.C. Graesser and G.H. Bower (eds) *Inferences and Text Comprehension*, pp.175-196. San Diego, C.A: Academic Press.

van den Broek, P. (1994) 'Comprehension and Memory of Narrative Texts: Inferences and Coherence' in M.A. Gernsbacher, M.A: 539-588.

- van Dijk, T.A and Kintsch, W. (1983) *Strategies of Discourse Comprehension*, New York: Academic Press.
- Varela, F.J., Thompson, E. and Rosch, E. (1991) *The Embodied Mind: Cognitive Science and Human Experience*, Cambridge, MA: The MIT Press.
- von Neumann, J. (1947) 'Preliminary Discussion of the Logical Design of an Electronic Computing Instrument'. Us Army Ordnance Report.
- Vonk, W. and Noordman, L. (1990) 'On the Control of Inferences in Text Understanding' in D.A. Balota, G.B. Flores d'Arcais and K. Rayner (eds) *Comprehension Processes in Reading*, pp. 447-464. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Wales, K. (1989) *A Dictionary of Stylistics*, London: Longman.
- Wallace, C. (1992) *Reading*, Oxford: OUP.
- Waltz, D.L. (1989) 'Connectionist Models: Not Just a Notational Variant, Not a Panacea', in Y. Wilks (1989) *Theoretical Issues in Natural Language Processing*, pp. 56-63. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Wason, P.C. (1966) 'Reasoning', in B. Foss (ed) *New Horizons In Psychology*, Harmondsworth: Penguin.
- Wason, P.C. and Shapiro, D. (1971) 'Natural and Contrived Experience in a Reasoning Problem', *Quarterly Journal Of Experimental Psychology* 23: 63-71.
- Weber, J.J. (1996) (ed) *The Stylistics Reader: From Roman Jakobson to the Present*, London: Arnold.
- West, C., Lazar, M.M., Kramarae, C. (1997) 'Gender in Discourse' in, T.A. van Dijk (ed) *Discourse as Social Interaction: Discourse Studies, A Multidisciplinary Introduction, vol. 2*, pp. 119-143. London: Sage.
- Widdowson, H.G. (1984a) 'Reading and Communication' in C.Alderson and A.Urquhart (1984) *Reading in Foreign Language*, pp. 213-230. London: Longman.
- Widdowson, H.G. (1984b) *Explorations in Applied Linguistics*, Oxford: OUP.
- Widdowson, H.G. (1994) 'Old Song That Will Not Declare Itself: On Poetry and the Imprecision of Meaning' in, R. Sell and P. Verdonk (eds) *Literature and The New Interdisciplinarity: Poetics, Linguistics, History*, pp. 31-43. Amsterdam: Rodopi.
- Widdowson (1995a) 'Discourse Analysis: A Critical View' in *Language and Literature*, 4 (3): 157-172.
- Widdowson, H.G. (1995b) 'Review of Fairclough's 'Discourse and Social Change'', *Applied Linguistics*, 16 (4).

Widdowson (1996) 'Reply to Fairclough: Discourse and Interpretation: Conjectures and Refutations', *Language and Literature* (1996) 5 (1): 57-69.

Widdowson, H.G. (1997) 'The Use of Grammar; The Grammar of Use', *Functions Of Language* 4 (2): 145-168

Widdowson, H.G. (1998) 'Review Article: The Theory and Practice of Critical Discourse Analysis', *Applied Linguistics* 19 (1): 136-151.

Wilensky, R. (1983) *Planning and Understanding: A Computational Approach to Human Reasoning*, Reading, Mass.: Addison Wesley.

Wittgenstein, L. (1921) *Tractatus Logico-Philosophicus*, trans. D.F. Pears and B.F. McGuinness (1961), London: Routledge and Kegan Paul.

Wittgenstein, L. (1953) *Philosophical Investigations*, trans. G.E.M. Anscombe, 3rd ed. (1967), Oxford: Basil Blackwell.

Whorf, B.L. (1956) *Language, Thought and Reality*, J.B Carroll (ed), Cambridge, Mass: The MIT Press.

Wodak, R. (1996) *Disorders of Discourse*, London: Longman.

Wolpert, L. (1992) *The Unnatural Nature of Science*, London: Faber and Faber.

