

**CRITERIA AND CONCEPTS:
AN ANTI-REALIST APPROACH TO WORD-MEANING**

Bradley William Franks

Ph.D.

University of Edinburgh

1989



Declaration

I declare that this thesis has been composed by myself, and that the work reported herein was carried out by myself, unless otherwise indicated.

Bradley Franks

29th September 1989

Acknowledgements

I should like to thank my supervisors, Terry Myers and Robin Cooper, for their guidance, support and advice. I should also like to thank other friends and colleagues at the Centre for Cognitive Science: Nick Braisby, for discussions in which many of the ideas here presented emerged, although he may no longer recognise them; George Dunbar, Scott McGlashan and other members of the Mental Lexicon workshop at the Centre for Cognitive Science, for numerous comments and discussions that have shaped my thinking on many areas; Philip Roberts for thoughts on the philosophy; and Richard Cooper for help with the formatting of the thesis. I should also like to thank the following for encouraging comments and constructive criticism: Renate Bartsch, Jo Calder, Nick Chater, James Hampton, Ilse Karius, Ewan Klein, Karen Lyon, Charles McCarty, Jim Miller, Doug Medin, Leo Noordman, Mike Oaksford and Pete Whitelock. And Sam for helping to keep The Furies at bay.

A different order of gratitude is owed to my family, for their unfailing tolerance, care, and appropriate irony during my time in the academic wilderness: deepest thanks to my mother, Joan, to Tony, Ian, Joanne and Derek. Most of all, I should like to thank Netty, for things that I can neither specify nor quantify. If a thesis can be dedicated, then I dedicate this to her.

The research was carried out whilst I was in receipt of a Doctoral Programme Award from the Economic and Social Research Council.

Abstract

The central aim of this study is to articulate a view of word-meaning that respects constraints from philosophical and semantic anti-realism, and from the psychological study of concepts. The central philosophical argumentation concerns the relationship between agents, language and the world, and gives rise to a cognitive Manifestation Constraint on the individuation of semantic content ascribed to word-meanings. This constraint is operationalised through a consideration of psychological phenomena regarding concepts and their combination. The philosophical arguments, together with the Manifestation Constraint motivate a discussion of extant views of adjectival modification and concept combination. This then supports three subgoals of the thesis. The first is to articulate a plausible position on word-meanings and concepts. The second is to propose a mechanism for the combination of concepts. And the third is to provide an account of the combination of concepts, and the resulting categorisation behaviour, for two troublesome problem-classes: "Privative" (e.g., *fake gun*, *stone lion*), and "Attributive" (e.g., *attractive ballet-dancer*) combinations.

The presentation can be divided into four parts. The first, Chapters I and II, outlines the general orientation adopted. Chapter I is concerned with the philosophical debate between semantic realism and anti-realism, and proffers a particular view of anti-realism. This results in a Manifestation Constraint on the ascription of semantic content to language-users. Chapter II operationalises this Constraint by discussing and interpreting cognitive studies regarding concepts. The second part concerns the detailed articulation of the position on concepts and word-meaning. Chapter III considers various views of concept combination and adjectival modification in the light of the Manifestation Constraint and their treatment of the problem classes. Particular aspects of various theories are endorsed. Chapter IV discusses some of the constructs of the position in detail, and synthesises the considerations highlighted in earlier chapters. A position on the generation of senses for combinations, and the perspectival-relativity of categorisation, is advocated. The anti-realist notion of a Criterial relation plays a central role in the account. Chapter V applies this account to the problem classes, and challenges the assumptions of some earlier treatments. This gives rise to a typology of privative combinations, and a scale of privative behaviour. Chapter VI considers some connections between the current position and several other areas of research in cognitive science, and suggests some consonant directions for future research. It also seeks to reflect some of the psychological considerations back onto the philosophical issues of Chapter I.

Contents

Introduction

I. Philosophical Animadversions

1. A Sketch of Realism	5
2. Anti-Realism: Negative Theses	11
3. Anti-Realism: Positive Theses.....	15
3.1. Manifestation, Verification and Cognition.....	15
3.2. Defeasibility and Meaning-relations	22
3.3. Holism and Molecularity	27
3.4. Anti-Realism and Partiality	37
4. Summary	42

II. On the Manifestation Constraint: Concepts and Cognition

1. The Problem Stated	44
2. Cognitive Phenomena.....	47
2.1. Intensional vs. Extensional Views.....	47
2.2. Sense Generation vs. Sense Selection	48
2.3. Context-Sensitivity.....	53
2.4. Compositionality and Monotonicity.....	58
2.5. The Role of Knowledge	61
2.6. On the Partitioning of Content	66
2.7. Concept Coherence.....	71
3. Summary: the Manifestation Constraint	75
4. Partiality, Relativity and Sense Generation	77
5. Summary	81

III. Some Views of Concept Combination

1. Introduction	83
2. Mediating Relations Theories.....	85
2.1. Some Views of Mediating Relations.....	85
2.2. What are Mediating Relations?	89

2.3. Difficulties with Mediating Relations Theories	90
3. Combinations Theories	93
3.1. Set-theoretic Approaches	94
3.2. "Intensional" Approaches	101
4. Levels of Modification	109
4.1. Dual-Level Modification.....	110
4.2. Multiple-Level Modification.....	111
5. Types of Modification	115
6. Summary.....	119

IV. On Defeasibility, Perspectives and Property Inheritance

1. Introduction	120
2. On Defeasible Semantic Relations	124
2.1. Types of Defeasibility	124
2.2. Defeat and Molecularity	127
3. On Perspectives	130
3.1. What is a Perspective?.....	130
3.2. A Problem and a Reply	134
4. On Analogy	136
4.1. Central Aspects of Analogy.....	136
4.2. The Motivation for Analogy.....	138
5. On Relativity, Perspectives and Partiality	142
6. On Concepts, Senses and Word Meanings.....	147
7. Summary: the View of Concept Combination.....	151
8. Summary.....	153

V. On Privatives and Attributives: the Current View

1. On Negating and Equivocating Privatives.....	155
1.1. Qualities and Types of Privatives.....	155
1.2. An Account of Negating Privatives.....	161
1.3. An Account of Equivocating Privatives	172
2. On Modal and Pure Privatives	187
2.1. The Qualities of Modal and Pure Privatives	187
2.2. An Account of Modal Privatives.....	190
2.3. An Account of Pure Privatives.....	193
3. A Typology of Privative Combinations.....	194
4. On Attributives.....	197

4.1. An Account of the First Reading	198
4.2. An Account of the Second Reading	199
4.3. An Account of the Third Reading	200
5. Summary: a Scale of Privativity.....	202

VI. Some Implications, Connections, and Suggestions for Future

Research

1. Some Connections and Suggestions for Research.....	204
1.1. Situation Theory, Constraints and Background Conditions.....	204
1.2. Formalised Nonmonotonic Inference.....	209
1.3. Degree and Measure Adjectives.....	213
3. Some Implications of the Current View.....	217
3.1. Privativity, Attributivity and Language Success.....	217
3.2. Some Other Implications	220

References.....	228
------------------------	------------

Appendix: Publications

Introduction

In this study, I attempt to advocate a unified and coherent approach to word meaning, and to certain types of concept combination in particular. The position advanced is that psychological and philosophical approaches to word-meaning are mutually constraining. By considering some fundamental questions regarding semantics, we can articulate more clearly the role of cognition in a view of word-meaning, and determine exactly what kind of a constraint on semantics cognitive considerations might prompt. Perhaps the most important issues in philosophical semantics concern the relationship between our semantics, and the external world. This is the locus of the debate between realism and anti-realism. A discussion of this debate, in addition to the presentation of a particular anti-realist position, is the subject-matter of the first chapter.

In Chapter I, I argue in favour of an anti-realist approach to semantics in general, and in favour of a cognitive construal of the central arguments contra-realism. The crucial aspect of this is Manifestation, to which there are two sides. The first is the semantic skills attributed to agents in their exercise of semantic knowledge, which gives rise to a particular construal of verification and warrant. The second is the ascriptions of content to agents that we can make as theorists, which gives rise to the Manifestation Constraint, which emerges as a cognitive constraint on the individuation of semantic content. This Constraint has far-reaching effects on both the form and the content of an anti-realist view of word-meaning. Anti-realism also carries a commitment to a certain view of semantic relations as defeasible Criterial relations, and as inhering in molecular or local holistic networks. A final implication is that anti-realism can find a natural expression not only in defeasibility but also in partiality, and that concepts might be viewed as partial objects.

In Chapter II, an attempt is made to flesh out the Manifestation Constraint, through a consideration of cognitive phenomena pertinent to word meanings and concepts. Special emphasis is placed on the problem classes of Privative (e.g., *fake gun*, *stone lion*) and Attributive (e.g., *attractive ballet-dancer*, *good dog*) combinations, as being those most suited to examining the individuation, rather than the exhaustive specification, of semantic content ascriptions. Major aspects of the operationalisation of the Manifestation Constraint include a commitment to sense generation, to nonmonotonicity and knowledge-drivenness in concept combination, to a theory-driven account of concept coherence, and to an "intensional" account of concepts. Further, the interaction of sense and reference is emphasised, manifesting the interaction of theoretical and situational knowledge, and distinctions between

different kinds of contents ascribed to concepts are motivated. The overarching phenomena considered concern the context-sensitivity of specificity and the flexibility of generated senses. An interpretation of these phenomena in terms of lexical concepts as partial objects, and senses as their extensions, is proffered. Finally, the relativity of categorisation judgements, both criterion- and sortal- relativity, is advocated.

In Chapter III, I discuss various views of concept combination, or adjectival modification of nouns, in the light of the operationalisation of the Manifestation Constraint, and the appropriate treatment of the problem classes. These are discussed under two orthogonal distinctions: between Mediating Relations, and Combinations theories, and between multi-level and multi-type modification theories. In terms of the former, I argue against Mediating Relations theories and in favour of Combinations theories. Within Combinations theories, I argue against extensional approaches (classical and fuzzy set theory), and in favour of intensional approaches, in particular more psychological views. Concerning the second distinction, I argue in favour of multi-level and multi-type modification, with no a priori restriction on numbers of types or levels. In particular, I advocate an amalgamation of implicit semantic attachment and context-sensitive specification of pre-properties as properties, as a way of accounting for the problem classes.

In Chapter IV, I lay some of the further groundwork for offering a treatment in this vein. I begin by arguing the case for defeasible concept extensions in a view of motivated property inheritance. I then make the types of defeasibility necessary for an anti-realist position more precise, distinguishing between Undercutting and Type I and Type II Rebuttal, and note that Undercutting is the type that relates to Criterial relations. Then I comment on the interaction of types of defeat and types of molecular network. The next step is to be more precise about the relativity of categorisation judgements: I define the nature of Perspectives, and distinguish between two types, noting their independence. I then discuss the role of analogical mapping as the particular kind of motivation for property inheritance, distinguishing its stages, and postulating its role vis a vis partiality. I then synthesise perspectival-relativity, partiality, and analogy, resulting in a redefinition of some basic constructs concerning partiality, and some conditions on particular kinds of perspectival categorisation. The next section draws together many of the foregoing points, in discussing the relations between concepts, senses and word-meanings. I argue that they are independent, though senses and word-meanings are related by perspectives. Finally, I briefly summarise the position I have advocated regarding concept combination. The treatment of the problem classes offered in Chapter V involves outlining a process for the generation of a sense for a combination (requiring flexibility and context-sensitive precision), and then

examining the categorisations and perspectives under which a referent that possesses the properties described in the sense, might be brought.

Having motivated and sketched the mechanism for concept combination, Chapter V goes on to apply the mechanism to the problem classes. Preliminary to this are a series of distinctions between different subtypes of Privatives: firstly, a distinction between Proper Privatives (in which privativity is a result of the type of the A), and Functional Privatives (behaviour resulting from the semantic interaction of the head and modifier). Secondly, orthogonal distinctions between subtypes of privative behaviour: Equivocators and Negators (for which I suggest both Proper and Functional types), and Modal and Pure Privatives (for which no Functional correlates appear to hold in English). I propose different, though related, treatments for each subtype. The treatments are seen to deny three fundamental assumptions of standard accounts. I then apply the mechanism to Attributives, providing a treatment for three discriminable readings of such modifications. The outcome of the chapter is a Typology of Privatives, and a Scale of Privative behaviour.

In the final chapter, I discuss some connections between the current study, and other ongoing work in the cognitive sciences. In particular, I note the parallels with certain aspects of Situation Theory, with the formalised study of nonmonotonic inference in Artificial Intelligence, and with a particular formal semantic analysis of the semantics of degree and measure adjectives. My argument is that there are important differences, but that these very differences form the basis for future research. In the second part of the chapter, I discuss various other implications of the approach, and note some of the ways in which the psychological arguments of foregoing chapters reflect back upon the philosophical questions noted in Chapter I.

CHAPTER I

Philosophical Animadversions

This chapter has two motivations. The first is to outline some central arguments against semantic realism, and in favour of anti-realism. I will not cover the gamut of arguments which have been advanced to counter realism: my approach will be to select those aspects which bear the greatest relevance to the issues which arise in considerations of word meaning. The second goal is to advocate a particular view of anti-realism, in order to motivate the rest of the study. This will indicate that an anti-realist approach must, on penalty of incoherence, be a cognitive approach. I argue this case through a consideration of some of the Negative and Positive Theses of anti-realism.

I.1. A Sketch of Realism

Since the origin of anti-realism is its opposition to realism, we need some idea of the provenance of "realism". I should emphasise that what is *not* at issue here is the intuitive and pre-theoretic "realism" that characterises our everyday thought and talk. Rather, semantic realism is a peculiar and pervasive theoretical position, that connects only loosely with the intuitive idea. The cluster of attitudes which comprise realism is exemplified by several quotations. Luntley (Luntley 1988) comments:

Realism is the belief that the world exists independently of our knowledge of it. What this means is that for any proposition P, to the question 'Is P true or false?' the world has a determinate answer irrespective of our ability to calculate the answer.

(Luntley, 1988: 1).

Platts (Platts 1979) states:

[Realism] embodies a picture of our language reaching out to, connecting with the external world in ways that are (at least) beyond our present practical comprehension. It embodies a picture of an independently existing, somewhat recalcitrant world describable by our language in ways that transcend (at least) our present capacities to determine whether those descriptions are true or not...The world, austere characterised by our language, will always outrun our recognitional capacities.

(Platts, 1979: 237-238).

And Wright (Wright 1987) distinguishes

four non-equivalent elements in realism as Dummett has described it:

- i. endorsement of bivalence for statements which are not effectively decidable;...
- ii. admission of the possibility of verification-transcendent truth;
- iii. acceptance that our understanding of all statements fits the truth-conditional schema;

iv. acceptance that for any true statement, there must be something in virtue of which it is true.

(Wright, 1987: 319).

These attitudes have some corollaries, that should be noted. One is that the central semantic notion - truth - is radically non-epistemic. For effectively decidable domains, an agent might grasp whatever content a sentence has, but for non-decidable domains, she may not. The agent is, then, peripheral to the communication of meaning: what is crucial is the world's determinate constitution, part of which is the determinate relation between words and referents. So there are certain aspects of meaning, or whole domains of discourse, which are constitutively beyond the grasp of language-users.

A second implication is the "independence thesis", which McGinn proffers as the "fundamental thesis of realism" (McGinn 1980). Consider the relation between our evidence for the application of a concept (e.g., the categorisation of an entity) and the actual facts of the matter. One case is the relation between statements about material objects (**M**-statements), and the (perceptual) evidence used in making such statements (**E**-facts). Suppose we possess a range of **E**-facts that are not in doubt, and which are used to support **M**-statements. Now, realism has it that there can be recognition-transcendent facts: so it is possible for **M**-facts (and statements) to obtain without the respective **E**-facts (statements) being true. And since no set of **E**-statements ever entails the truth of an **M**-statement, it cannot be a sufficient condition for the truth of an **M**-statement that some related **E**-statement be true. Exactly the same argument holds for our ascription of mental states to third parties on the basis of behavioural evidence. The independence thesis issues in a two-way independence: since no evidence is ever necessary for the truth of a (recognition-transcendent) fact, then no fact is ever sufficient for the evidence; and to say that the former is not sufficient for the latter is to say that the latter is not necessary for the former. The independence thesis has crucial implications for the way in which categorisation might work. It also provides a point at which a sceptic might obtain a footing, and one which the realist is powerless to deny (Grayling 1985, Strawson 1985). The thesis cannot deny the sceptics' claim that,

since it is never a contradiction to assert the conjunction of the best evidence we have for some epistemological claim **p** with the denial of **p**, we can never be fully justified in believing what we normally take ourselves to believe about the world.

(Grayling, 1985: preface).

A third issue is the metaphysical imagery of realism. Realism allies a semantic modesty with an overwhelming immodesty. The modesty lies in the minimal role of the cognitive agent, the immodesty in the role of the theorist, who has privileged access to the way the

world "really" is. This is an atemporal and non-perspectival position which allows the realist to claim that there is a determinate fact of the matter, and indeed to claim that the fact of the matter lies in a particular way. As Dummett (Dummett 1978) comments, regarding the determinate "reality" of the past:

What the realist would like to do is to stand in thought outside the whole temporal process and describe the world from a point which has no temporal position at all, but surveys all temporal positions in a single glance: from this standpoint...the different points of time have a relation of temporal precedence between themselves, but no temporal relation to the standpoint of the description.

(Dummett, 1978: 369).

The realist has the same aim with respect to the states of affairs and the evidence of different agents. How are the modesty and immodesty reconciled? Dummett's chief argument against realism is that they are irreconcilable.

The quotations above indicate that there are at least two discrete theses at work under the rubric of "realism". These have been labelled by Luntley as the "objectivity-of-content" and "objectivity-of-truth" issues. The former is the issue of whether we grasp semantic contents whose characterisation requires our conception of a world beyond that which is experienced: if it does we are committed to the idea that if there is a fact of the matter it is of a particular kind. On this characterisation of the realist view, there is no commitment to bivalence, only to the unproblematic assumption that contents can be true even when unverified. The second issue is whether the contents we grasp are ones that correspond to a determinate reality beyond our investigation of it - whether those contents have a recognition-transcendent truth-value. The realist response commits one to the assumption that there is a determinate fact of the matter, and that it is this which we grasp when grasping the content of any sentence, recognition-transcendent or otherwise. The crucial point, however, is not just that the contents we grasp *have* a determinate truth-value, but that the grasp of contents *consists* in a grasp of truth-values (through knowledge of the appropriate truth-conditions). That is, semantic realism appears to require a combination of Wright's thesis (iii), with either (i) or (ii), depending upon the species of bivalence adhered to (more on this below). However, that there are two separable components is important both to an explication of the relations between anti-realism and classical logic, and to the relations between semantic and scientific realism. It is to the first issue that I now briefly turn.

What exactly is the principle of bivalence, and how does it relate to laws of inference in classical logic? We need to consider the following classical laws (Dummett 1978) which state the truth of every instance of one of four different schemas:

- (i). A or not A: $A \vee \neg A$

(ii). It is not the case that neither A nor not A: $\neg \neg (A \vee \neg A)$

(iii). Not both A and not A: $\neg (A \& \neg A)$.

(iv). If it is not the case that not A, then it is the case that A: $\neg \neg A \rightarrow A$

At the level of logical laws, these are labelled (i) the law of excluded middle, (ii) the law of excluded third (in Dummett's terms), (iii) the law of contradiction, and (iv) the law of double negation. Now, as Dummett notes, there are four semantic principles, each corresponding to one of these logical laws:

(i'). **Bivalence**: every statement is either true or false

(ii'). **Tertium non datur**: no statement is neither true nor false

(iii'). **Exclusion**: no statement is both true and false

(iv'). **Stability**: every statement that is not false is true

Why distinguish between these two levels of description? The short answer is that acceptance of (i') normally entails acceptance of (i), whilst the converse does not hold. So the touchstone of realism is (i') and not (i). Further, as Dummett comments, a rejection of the principle of bivalence does not imply a rejection of tertium non datur (Dummett, 1978)

In the principles noted above, bivalence is "Strict Bivalence": the principle that every statement is determinately either true or false (Wright, 1987). A weaker version is what Wright notes as "Generalised Bivalence", and Rasmussen & Ravnkilde as DEBivalence (abbreviating "Determinacy Bivalence"): the principle that every statement is determinately true or not; this version tolerates (determinate) truth-values other than truth and falsity, and (determinate) truth-value gaps (Rasmussen and Ravnkilde 1982). On both principles, whatever truth-values hold do so regardless of our having the capacity to recognise their holding. Wright's claim is that realism is tied to Generalised Bivalence, whereas Luntley's claim is that Strict Bivalence is defining of realism. This distinction is no mere quibble: it may be the key to the relation between realism and classical logic, and hence to whether anti-realism is committed to a logical revisionism. We must note the possible domain-relativity of the application of these principles. There is no reason why an anti-realist should have any qualms about bivalence for restricted and effectively decidable domains. It is the assumption that such principles operate in all domains that is questioned, along with the (apparent) correlative commitment to classical logic for those domains.

We must be clear regarding the kinds of commitment which any notion of bivalence involves. If truth is regarded as bivalent, it thereby transcends our abilities to determine its extension in our language. But such transcendence does not flow, contingently, from our being semantically lazy or unskilled; it is not that we have not yet determined the truth, or

that it is too complicated for us to work out. As Tennant (Tennant 1987) puts it, truth transcends our capacities

in a de-stabilising way...It is not merely medical impossibility, in Russell's immortal phrase, or a historical handicap, that denies us access in full to the One True Story. Rather, bivalent truth is a property...that by its very nature could escape detection in eternal hideouts.

(Tennant, 1987: 128-129).

This is no mere epistemic possibility; it is rather that truth can and must really outrun evidence - as Wright comments, "that the availability or otherwise of evidence is no conceptual constraint on the capacity of a statement to be true" (Wright, 1987). The independence thesis implies that it may be the case that we could never know the real truth-value of statements for whole domains of discourse, and we will never know that what we take to be true is or is not true.

We must now ask, which of the species of bivalence is allied to realism - Determinacy or Strict Bivalence? The answer must be that, in its loosest form, realism carries no commitment to strict bivalence. As McDowell comments (McDowell 1976), a commitment to bivalence necessarily implies one to transcendence, but a commitment to transcendence carries no obvious commitment to the precise constitution of the truth-values of the determinate world; that is, transcendence does not imply strict bivalence. So strict bivalence implies determinacy, but not vice versa. Wright labelled the two options "classical realism" (as deriving from strict bivalence) and "bare realism" (deriving from determinacy), respectively (Wright 1980).

The implication is that any anti-realist argument must be able to challenge the bare realist. And Dummett's chief arguments seem, as Grayling (Grayling 1982) notes, to be directed in this way.

Having noted something of the connection between realism and classical logic, we turn now to the relation between semantic and scientific realism. Anti-realism's quarrel is with semantic realism, and any view in which the central semantic notion is both constitutively transcendent and held to be grasped in language understanding. The comments on bivalence relate to semantic and not scientific realism. Semantic realism is logically independent of scientific realism, and semantic *anti*-realism is perfectly consistent with scientific realism. According to Boyd (Boyd 1987), we can isolate four distinct theses in scientific realism. The most important of these are ontological theses: firstly, that existence is univocal, so that that entities of science are "as real" as those of everyday experience; and secondly, an ontological independence thesis, which states that the reality that scientific theories describe is

"largely independent" of our thoughts or theoretical commitments. The first thesis is logically independent of the question of the determinacy of truth-value of statements concerning objects in the ontology. And the ontological independence thesis is clearly separate from the semantic independence thesis of McGinn (see I.1). And it is the ontological thesis with a semantical cast that is the locus of anti-realist arguments. There is no quarrel with the independence of the reality that theories describe; but the assumption that the semantic contents of statements of our theories are such as to make the latter bivalent (on whatever reading) concerning reality, is the difficulty. It is then argued by Tennant (1987) that the content bestowed on our scientific language, by our methods skills and abilities, offers no justification for the view that in general our claims about the acknowledged external reality are bivalent. Bivalence may be adjoined to, but actually extends, the (ontological) doctrine of scientific realism.

So it is that in the anti-realism espoused by Dummett, the crucial issues concern not whether certain classes of entities exist independently of human thought (ontological realism), but with what it is in virtue of that statements are true, when they are true; and whether such a notion can consistently be utilised in natural language semantics. The crucial issue is then the nature of truth-determination, and anti-realism is concerned not with whether the entities to which we refer exist, but with whether the statements that we use to refer are bivalent, hence whether reference to them is determinate as regards our language.

Having surveyed some of the central aspects of realism, we may adumbrate the main arguments in favour of anti-realism. This is the negative programme; the positive one involves articulating appropriate semantic constructs. The rest of the thesis forms an attempt to do that within a highly restricted domain. The anti-realist arguments will be concerned with bare realism since, if this is rendered problematic, so then must classical realism. Another reason for not concentrating on the arguments against classical realism is that the focus of this thesis is not the proper treatment of the logical constants - which is necessarily the object of such arguments - but the treatment of content words. However, the epistemology of doubt and certainty which is attendant upon both kinds of realism will be canvassed, for this will be crucial to the treatment of word meanings.

I.2. Anti-Realism: Negative Theses

Dummett (Dummett 1976) holds that philosophical questions about meaning should be

interpreted as questions about understanding: a dictum about what the meaning of an expression consists in must be construed as a thesis about what it is to know its meaning (Dummett, 1976: 69).

And he comments,

to suppose that, in general, a knowledge of meaning consisted in verbalisable knowledge would involve an infinite regress: if a grasp of the meaning consisted, in general, in the ability to state its meaning, then it would be impossible for anyone to learn a language who was not already equipped with a fairly extensive language...[Such knowledge] must be implicit knowledge. Implicit knowledge cannot, however, meaningfully be ascribed to someone unless it is possible to say in what the manifestation of that knowledge consists (Dummett, 1978: 217).

If truth is the central concept in a theory of meaning, the theory must explain how knowledge of truth-conditions connects with actual language use. If a theory of meaning is viewed as one of understanding, the requirement that this connection be explicated is justified by an appeal to two separable requirements on a theory of semantic understanding. The first is that we must state what would count as a speaker's manifestation of her knowledge of the meanings of the statements of the language; and the second is that the kinds of contents which we ascribe to an agent's knowledge of meaning must be acquire-able. The first requirement is the central focus of Dummett's arguments contra realism. Prior to a detailed consideration of this, I shall first outline the general argument.

The second quotation from Dummett indicates the import of the manifestation requirement. If all of the knowledge of meaning were explicit knowledge - that is, could be formulated in terms of other sentences of the same language - then there would be no difficulty. But to assume that this is the case is to entertain an infinite regress of meaning-dependence, in which the justifying sentences or concepts themselves are ultimately dependent upon the sentences whose meaning is being explicated. We must then postulate implicit semantic knowledge, and this demands a criterion for the manifestation of such knowledge.

The two central questions are then: what is it to know the truth-conditions of a statement?, and: is truth appropriate as the central notion for a theory of meaning? As Grayling (1982) comments, asking what analysis should be given to the notion of truth, is in effect to ask where, in the process of mastering a language, a grasp of the notion of truth comes in? If learning a language is learning what it is for sentences of the language to be true, then we

must be able to state what it is to know that a sentence is true without assuming any prior understanding of the sentence - on pain of circularity. It is not disputed that, in many cases, the meaning of a sentence can be given in linguistic terms, but then such knowledge must be explicit knowledge, thus presupposing prior mastery of much of the language. And this implies that such a mechanism would be inappropriate for the early stages of language learning, and for what Dummett considers as the the "lower" or "more primitive" levels of language. However, it must be clear that such a mechanism is not tenable even for all domains of discourse which are "non-basic" (providing we had a principled manner of demarcating the basic and non-basic).

Let us be clear about what is at stake here. The issue is not how to specify what we might accept as an indicator of a speaker's recognition that a statement's truth-conditions hold. This is not difficult since, providing the statement's truth-value is in principle decidable, it seems that we could elaborate some indicator of her recognition of the fact. But natural language use is replete with statements that are not effectively decidable. Standard instances include quantification over infinite or unsurveyable domains (such as the future), subjunctive conditionals, references to inaccessible regions of space-time, and ascriptions of mental states to third parties. In such cases, there is no effective procedure for determining whether the truth-conditions of the statements in question are satisfied or not. Now, we might chance upon a way of recognising whether its truth-conditions were satisfied, but Dummett's point is that, for any sentence not known to be effectively decidable, "we cannot equate a capacity to recognise the satisfaction or non-satisfaction of the conditions for the sentence to be true with a knowledge of what that condition is" (Dummett, 1976). The argument is this. By hypothesis, the truth-condition is one which may hold or fail to hold, and in both cases, we may be unable to recognise that fact: so although knowledge of what it is for the condition to hold or not to hold might depend on an ability to recognise either or both states of affairs when we are appropriately situated, such a condition could not be exhaustively explained in terms of that ability.

This is the argument in outline. To prise some of the issues apart, Dummett suggests that we consider what is involved in ascribing truth to statements. Consider the following principle (Dummett, 1976):

Principle C: If a statement is true, there must be something in virtue of which it is true.

This principle underlies the correspondence theory of truth (which forms the - often tacit - base of semantic realism). In order to appreciate the force of C, Dummett suggests a consideration of counterfactuals alleged to be true despite there being no grounds which might

prompt this conclusion. Consider an ascription of the ability, *Netty is good at learning languages* (p). To the question, "must p be either true or false?", we might envision three responses:

(i). It is not necessarily either.

(ii). Linguistic ability must correlate with or consist in some aspect of cognitive neurological structure; Netty must either have or not have this, so p must be either true or false.

(iii). p must be either true or false, regardless of whether linguistic ability correlates with neurological structure.

Now, the assumption is that the evaluation of p depends upon the truth of an associated subjunctive conditional, q : *if Netty were trying to learn a language, she would easily succeed*. That is, p would be true if q were true, and false if false. The question, then, of whether q is true coincides with whether bivalence holds for p . Response (iii) implies that q , if true, is barely true - that is, true in a way which violates principle C; response (ii) shares the assumption of bivalence but is a reductionist account; and response (i) enjoins a denial of bivalence for p (dictated by the considerations that there may be nothing which might help in determining p 's truth-value, and that to allow p 's bare truth would be equally inappropriate). As Dummett comments, this response allows one to "escape the dilemma by rejecting the law of bivalence" (Dummett, 1976: 91). Responses (ii) and (iii) embody a reductionist and a naive realism respectively, whilst (i) is an anti-realism.

The notion of reducibility employed is a loose one: for any statement p in class M , there is a family p^* in a related class R , such that, for p to be true, it is necessary and sufficient that some set of the family of p^* be true. So "bare truth" is then definable as follows: statement p is barely true if p is true; and there is no class of statements, not containing p , to which any class containing p can be reduced. A similar distinction might be made regarding the application of concepts, and reduction-sets of properties. Naive realism is thus the combination of realism for a class of statements, and the belief that those statements are barely true.

We can now re-examine what it means to ascribe a knowledge of truth-conditions to a speaker. There will be two different cases, depending upon whether a particular sentence (S) is used to make a statement which is or is not capable of bare truth. If S is not capable of bare truth, then there must be an appropriate reduction class, some members of which are true. So a grasp of S 's truth-conditions will depend upon an implicit grasp of how S 's truth relates to that of the subset of the reduction class. This would be straightforward for effectively decidable domains. However, two different problems arise. The first is that it is

difficult to see how an acquisition mechanism for such reductions could begin: for it presupposes an extensive prior language mastery. The second is that, for undecidable domains, we are faced with the independence thesis: if S is not capable of bare truth, and has transcendent truth-conditions, it is not clear that we could specify an appropriate reduction class, since the latter must be a necessary and sufficient condition for S - that is, must support an entailment of the truth of S. And this very possibility is ruled out by the independence thesis.

In contrast, if S can be used to make a barely true statement, its semantic translation would be trivial. The model for this case is taken to be the reporting of observations: for example, if someone is able to discern, by looking, that one tree is taller than another, then he knows what it is for such a state of affairs to hold: that is, he knows what conditions must be satisfied for the truth of *this tree is taller than that tree*. But this model also fails on its application to non-decidable domains, since these are ones whose truth, by definition, cannot be observed.

So it seems as if there are two distinct models that spell out what a knowledge of truth-conditions might consist in; and both appear to be unable to address the kinds of non-decidable domains whose bivalence is a constitutive assumption of realism. It remains, then, to see how a realist might make a convincing case for the extension of either model to the transcendent cases. Dummett claims that we come to think of our mastery of the meaning of non-effectively decidable domains by appeal to the observational model, the link with which may be only surreptitious. That is,

we try to convince ourselves that our understanding of what it is for undecidable sentences to be true consists in our grasp of what it would be to be able to use such sentences to give direct reports of observation.

(Dummett, 1976: 99).

This is linked to a second regulative principle governing the concept of truth:

Principle K: If a statement is true, it must be in principle possible to know that it is true.

And Dummett's argument is that principles C and K are closely intertwined, since we could not be said to grasp what it is for a statement to be true if we had no conception whatever of how it might be known to be true. This point, he claims, even a realist would have to concede; and it is this which underpins realist attempts to extend the observational model beyond its means. Any theory which utilises a transcendent notion of truth at its heart must, as a result, fall foul of this paradox - that to grasp the meaning of a sentence is to know its truth-conditions, even where such conditions are unknowable. We have arrived

at an answer to the question of precisely what a knowledge of truth-conditions might look like; and this has led to a response to the question of whether truth is appropriate as the central concept in a theory of meaning. The answer must be negative: in abandoning the assumption that bivalence holds for all of natural language semantics, we are brought also to the conclusion that such a semantics must needs be fashioned along non-truth-conditional lines.

I.3. Anti-Realism: Positive Theses

I will consider several aspects of the positive programme of anti-realism: those relating especially to word meaning. Firstly (I.3.1), I consider the nature of manifestation, and its connection with verification and recognition; and the kind of constraint a Manifestation Constraint might be. Secondly (I.3.2), I consider a particular anti-realist conception - that part of a word's meaning can be expressed through defeasible, Criterial meaning-relations. This issues in a particular view of doubt, certainty, and defeasibility. Thirdly (I.3.3), I discuss some kinds of meaning-networks within which such relations might hold, and argue in favour of restrictions on semantic holism, and in favour of a view of molecularity. The fourth aspect (I.3.4) is a discussion of a view of word meanings as partial objects, a view which can be motivated through anti-realism.

I.3.1. Manifestation, Verification and Cognition

In this section, I will argue that an anti-realist view must adopt a naturalistic epistemology. Hence, an appropriate semantics must interact with the cognitive sciences.

Concerning manifestation, we must distinguish between the kinds of evidence used by theorists in ascribing semantic knowledge, and the kinds of information used by the agents themselves in communication. It seems that Dummett's arguments have made an insufficiently sharp distinction, with the result that he has often been charged with a "behaviourism" or "semantic externalism" (Craig 1982). Related to this distinction is the question of the precise nature of manifestation, and its relation to verification. A further issue concerns what the manifestation constraint is a constraint upon. Hence the two issues concern, firstly, the content of the manifestation requirement, and, secondly, its form and function as a constraint on semantic knowledge ascription.

I.3.1.1. Publicity and Behaviourism:

In the accounts of both Dummett (1976, 1978) and Tennant (1987), manifestation and

publicity are closely linked. Tennant notes the Publicity Principle thus:

when we are mastering language, and when we are exercising that mastery, all that is available to us for gleaning or conveying meanings is the overt, observable behaviour of fellow speakers.

(Tennant, 1987: 3).

He divides this principle into an acquisition and a manifestation component. A comment from Dummett underscores the connection:

the meaning of...a statement cannot be, or contain as an ingredient, anything which is not manifest in the use made of it, lying solely in the mind of the individual who apprehends that meaning.

(Dummett, 1978: 217).

According to Craig, this position results in "semantic externalism":

no item which is epistemically private to the speaker - which no one other than the he can know the nature of - can be essential to the meaning of any symbol, word or phrase he uses.

(Craig, 1982: 541).

However, we must distinguish between two different theses:

- (i). semantic knowledge must be observably manifest for agents' communicative success.
- (ii). semantic knowledge must be observably manifest for ascriptions of knowledge in a theoretical account of agents' communicative success.

Where we take "observably manifest" to mean a strictly behavioural criterion, we might assent to (i) but not to (ii): publicity may underpin our everyday manifestations of semantic understanding, for example in making ascriptions to the mental states of third parties; but a restriction to such evidence for a theoretical account of semantic knowledge produces an impoverished form of behaviourism. And it is with respect to (ii) that Dummett's manifestation argument operates. That is, when considering whether we can ascribe semantic knowledge to an agent, we are not doing so from the position of an agent who is in everyday communication with that agent, but from the perspective of scientific inquiry. The manifestation question is thus separable from the issue of publicity.

With respect to a scientific investigation of manifestation, the question is whether all manifestable knowledge must necessarily be manifestable in observable behaviour. The answer to this question must be no. One way of motivating this response is by appeal to evolutionary epistemology, a line advanced by Tennant (1987). Under anti-realism, a theory of meaning is a theory of understanding, which ties it to questions of epistemology, especially

a naturalised epistemology. And a naturalised epistemology must ask after the functions of linguistic communication, and so must take account both of cognition and evolution. For example,

Could it be that we are...somehow equipped with an innate capacity to understand certain statements without their full conferment of meaning being available in what the learner has observed of his fellows during his lifetime?...there might be ways of responding to it which go essentially beyond the immediately available evidence.

(Tennant, 1987: 16).

There might, then, be more information in the communication that someone is in pain, than is publicly manifested by, or consciously available to, the partners to the communication. Either the information may not be open to conscious inspection, or it may form part of the background conditions relative to which a communication is taking place. Tennant gives the example that we might emit certain kinds of pheromones when we are in pain - and these would form as much a part of the evidence as would overt behaviours. Hence there could be an evolutionary explanation as to why we link the (impoverished) behavioural stimuli to our very rich cognitive ascriptions. This challenges Craig's definition of semantic externalism: that a content is epistemically private does not necessarily mean that it is open to conscious inspection by the bearer. A content may be private in language use, but it may be inferred as the best explanation by scientific investigation. Semantic externalism, as defined by Craig, effectively prohibits the study of semantics from utilising any resources which are not, or could not be made, available to the conscious introspection of individuals. In contrast, my argument is that the agent's conscious awareness of such knowledge is irrelevant to the manifestation of that knowledge. It may, for example, be made explicit through experimental means.

The notion of manifestation, then, requires a broader range of potential indicators than has usually been considered. If a theory of meaning is to be one of understanding, then this demands a central role for the cognitive sciences: we must examine language in use in order to ascertain the kinds of knowledge we might ascribe to agents.

I have not yet, however, considered what manifestation might consist in. In abandoning bivalence, the anti-realist approach requires some notion to replace truth as the central concept in the theory of meaning. We need, that is, some concept whose applicability does not constitutively transcend human capacities. Grayling (1982, 1985) has argued that an appropriate concept is warrant, resulting in a verificationist account that "ties grasp of the sense of sentences to the ability to recognise when it is correct to use them" (Grayling, 1985). We can note the distinction between "assertibility" and "warranted assertibility" as follows. Warrant is what attaches to an agent's use of a sentence when she has grounds,

yielded by a verification procedure, for asserting that sentence. Grayling (1982: 289) gives the example of the statement, *the dome of Radcliffe Camera will be*visible from All Souls in 2084 AD*. Such a statement is assertible, since we know what verifying conditions it can be employed in, but until 2084 it will not be warranted, since the actual verifying situation will not be available until then. Since it is probable to some degree that the particular situation will become available, the statement has current warrant to that extent.

Manifestation, then, has two separable components. The first concerns the kinds of capacities that we may ascribe to agents in order to account for their understanding of language. This is discussed below, regarding verificationism. The second component concerns the kinds of contents which we can ascribe to agents' semantic knowledge. The second component is most crucial to this thesis, and is discussed in Chapter II.

I.3.1.2. Manifestation, Verification and Recognition:

Anti-realist verificationism holds that one does not have to actually *effect* a verification so much as be able to *recognise* one. And this places my grasp of meaning within the constraints of a naturalised epistemology, linking it with memory, conceptualisation and cognition. Hintikka (Hintikka 1983, Hintikka 1987) has argued that a covert and unwarranted assumption of most semantic theories is that verification can be and is effected in a once-and-for-all fashion, and at the time of comprehending an utterance. His claim is that verification is best construed as an ongoing interaction, whereby an individual will attempt to verify assertions with reference to a recalcitrant world. And this would again urge a central role for a broad range of cognitive faculties. Hintikka claims that verification is neither all-at-once nor once-and-for-all. I may hear a sentence, but not seek a verification there and then. I may do so later. The verification is then largely dependent upon my cognitive abilities. Or indeed, the actual verification might be carried out by other people (by specialists, as in Putnam's notion of the division of linguistic labour), or may not be carried out at all.

A difficulty for the anti-realist is to outline a non-solipsistic verificationism which is nonetheless constrained. There are three aspects to a verificationist semantics which require comment: does it, firstly, mean that I have some way of verifying a statement every time and place that I understand it? The second issue is, how does verification link with my capacities as a language user? This devolves into the issue of performing versus recognising a verification. The third issue is whether the evidence relation upon which verification is based is defeasible or not.

Appiah (Appiah 1984) distinguishes four species of verificationism:

α : Indefeasible verification: as in logical positivism.

β : Identification of the possibility of verification with the physical possibility of verification.

γ : Weak verificationism: there has to be some spatio-temporal position, whether or not it is one that anyone actually (timelessly) occupies, at which verification would be possible.

δ : Strong verificationism: the quasi-solipsistic assumption that verification must be done by us, here and now.

It will be clear that neither α nor β would be consistent with anti-realism. Realist critics of Dummett have attempted to claim that the manifestation argument implies such a procrustean framework. For example, McGinn assumes that, for an anti-realistic language-user, "understanding a sentence...consists in an ability to determine its truth-value in some canonical way" (McGinn, 1980: 21). As Tennant (1987: 113) comments, this claim fails on two grounds. The first is in its use of "truth-value" as that which an agent must grasp for understanding. And the second is the assumption that an agent can have a "canonical way" of determining such a semantic value: whatever verification takes as its goal (i.e., truth, warrant or whatever), a determinate procedure will yield an indefeasible decision. But this would result in the manifestation argument denying the coherence of both truth and warrant as the central semantic concept. That is, on either view, the notion of a conclusive decision-making skill results in our ascription of semantic knowledge beyond the agents' capacity to possess, since it must hold for both non-decidable and decidable cases. It seems, then, that McGinn's argument is misplaced.

The question then becomes: to which of the remaining options can, or must, anti-realism subscribe? I shall argue that anti-realism requires a position intermediate between γ and δ , the essential point being that it is the *quasi-solipsistic* concept which requires explication. I shall then claim that the only way of making γ coherent is to bring it closer to δ .

What underpins the locutions "here and now" and "some spatio-temporal position" is the issue of the relative similarity of conceptual schemes/schemes of individuation, or of recognitional capacities. Hence we might modify δ thus:

δ' : Strong verificationism': the quasi-solipsistic assumption that verifiability holds iff a verification is carried out by someone with a conceptual scheme identical to our own.

It is now noticeable that the term "quasi-solipsistic" plays no role unless we assume a

radical form of conceptual scheme relativism. And indeed, the "our own" here is yet to be pinned down. If it really means "myself", then the outcome is solipsism. This means that we might have to modify δ' thus:

δ'' : the assumption that verification holds iff it is carried out by someone whose conceptual scheme is sufficiently similar (in ways to be spelled out) to our own.

I shall now argue that some concept such as δ'' underpins any coherent interpretation of γ . The demand in γ was that the verification be done by someone, at some time. Two complementary points may be noted. Firstly, she could not recognise the verification of the sentence either as a verification in general, nor as one for this sentence in particular, if the verifier did not have a conceptual scheme which were sufficiently similar to ours. Secondly, neither we, nor the verifier would be able to claim, with any measure of justification, that the particular assertions under scrutiny at, say, time t and time $t + 1$, were in fact one and the same assertion. So even if his verification held, and was perhaps conclusive, there is just no way in which we could know. So the "someone", "somewhere" of γ must be a verifier as in δ'' .

If we now add in the anti-realist premise, that "verifiability" consists only in recognising verifications when they obtain, then we have the necessary link between δ'' and γ . That is, if any verification is to count, then it has to be recognisable as such, and such recognition requires commensurability of conceptual schemes. So γ is reduced to δ'' , and anti-realism takes its place as a moderate verificationism. In a case where someone else makes the verification, some criterion of similarity to the current conceptual scheme is required. In a case where no verification is possible at all, the requirement of recognising as opposed to making a verification, is crucial.

It may seem as if this discussion has simply replaced one terminological wrangle - concerning verification - with that of recognition. This sub-section therefore addresses the second question of the three noted regarding verificationism, noting one way of fleshing out the idea of the similarity of conceptual schemes. Clark (Clark 1985) has addressed the issue of "recognitional" capacities. He argues in favour of a strong distinction between the following two propositions:

ϵ : Grasp of meaning amounts to the ability directly to recognise *the circumstances* in which an assertion is warranted;

ζ : Grasp of meaning amounts to the ability to recognise (perhaps indirectly) *when* an assertion is warranted.

Ordinarily, he argues, a conflation of the two propositions is harmless: indeed the capacity to satisfy ζ will be dependent upon that to satisfy ϵ . But sometimes they can come apart: Clark gives the possible example of a blind man who has a machine which beeps three times for red and twice for mauve - satisfying ζ - we would not attribute to him an ability to fully grasp the meaning of, for example, *the book is red*.

Clark argues that even if the circumstances which warrant the two sets of claims (on the basis of mediated and unmediated information respectively) do happen to be co-extensive, the two sets of assertion-conditions must be recognised as different. This is clearly the case, but the point is that anti-realism requires only that we can recognise when a verification obtains (when an assertion is warranted), and not the precise circumstances. I can recognise when a friend is in pain, but I cannot recognise the precise (neurological) circumstances. The two conditions parallel what Tennant (1987: 116) labels the "distinction between being able to decide and being able to check that which *purportedly* decides" (my italics). Anti-realism is wed only to the latter, although for certain assertions, we may be able to recognise the circumstances of verification (self-referential assertions such as *I am in pain*, for example).

The anti-realist accepts that we are in a very real sense limited by our own recognitional capacities: terms associated with recognitional capacities other than our own have warrant-conditions which are necessarily transcendent as regards any community which lacks them. But this does not mean that we do not know when they are warranted, only that the circumstances which make them warranted are unavailable to us.

The above hopefully approximates an essay at the second question of the three regarding verification; the first question also received some consideration. A consideration of the question of defeasibility is contained in I.3.2. I have attempted to advocate a unitary thesis: that an anti-realist semantics must involve a broad range of cognitive faculties, and that as a result the concept of verification includes the recognitional and manifestation activities which might intervene before (or in place of) any actual verification. "Manifestation" thus involves a broader range of faculties than envisaged by Dummett, and a less strenuous skill than that promulgated by McGinn.

If recognition-of-verification is the skill to which manifestation gives rise, what kind of a constraint on semantic knowledge is the manifestation constraint? Luntley (1988: 40) distinguishes between three different "schemas for constraints on the ascription of content":

Schema 1: If P does not satisfy constraint C the content fails in significance.

Schema 2: If P does not satisfy constraint C there is no sense to supposing that P is determinately true or false.

Schema 3: If P does not satisfy constraint C P is not distinguishable from some apparently different content Q.

Schema 1, being concerned with meaningfulness, is a constraint which would emerge from the view that verificationism provides a specification of semantic contents: that is, for a particular content to be meaningful, it must be capable of verification-recognition. However, merely demanding that contents be verifiable in order to be meaningful does not, of itself (pace Luntley), commit one to the claim that the contents will be specified by the verification-conditions. And if they contribute to the content, there is no necessary assumption that they will exhaust the sense. Clearly, the exhaustive specification of content is an empirical and problematic issue, and any cognitive and naturalised account of semantics will probably have to accept a more modest goal, especially given Tennant's comment regarding "semantical pheromones" (I.3.1.1).

Schema 2 is the one advocated by Luntley (1988). However, if we bear in mind the kinds of issue discussed above with respect to McGinn's conception of manifestation, I think that Luntley's version is also problematic. This will be seen if we replace the terms "true" and "false" with "warrantedly assertible" and "not warrantedly assertible". This leads us to the very same paradox as Tennant envisaged on McGinn's view. In addition to this, to assume such a narrow purview for the constraint is to adopt an inappropriately constricted picture of realism (as classical realism) and, as a result, of anti-realism, as I argued earlier.

The kind of schema I shall adopt parallels that offered by Peacocke, which is of type 3. This aims not to exhaustively identify contents, but to individuate them. Luntley (1988: 42) reports Peacocke's (unpublished) promulgation of the following constraint (fulfilling C in schema 3):

For any content P there is an adequately individuating account of what it is to judge that P rather than some other content Q.

The way in which the content's individuation-conditions are specified will depend upon the kinds of meaning-relations employed. The current ones will be spelled out in forthcoming sections.

I.3.2. Defeasibility and Meaning-Relations

In this section, I discuss the approach to defeasibility required by a naturalised anti-realism.

This will indicate the way in which the "theory of error" operates, and relate it to the ideas of criterial and symptom relations. Henceforth, "C-relation" and "criterial relation" will be used interchangeably, as will "S-relation" and "symptom relation"; "C-support" will be taken to mean the support or warrant conferred on an assertion by the obtaining of the appropriate C-relations.

Quinton (Quinton 1973) has distinguished between five different senses of the term "certain":

η : Certain is equivalent to "logically necessary"; hence all contingent propositions are inherently doubtful;

θ : Certain is equivalent to "psychologically indubitable", in the sense that S is subjectively convinced that p;

ι : Certain as "self-authenticating", in the sense of propositions whose truth is guaranteed by the way in which they are expressed;

κ : Certain as "incorrigible": a statement is infallible or incorrigible if its truth follows from the fact that it is believed; it is wholly verified by the experience that prompts its assertion;

Thus if "I am in pain" is incorrigible, it is possible both that I can be in pain and not believe "that I am in pain".

λ : Certain as "beyond reasonable doubt".

Classical logic, and its attendant semantics, has conflated the options described in η and θ : X can be certain if and only if it stands in a relation of logical necessity (i.e., of deduction) to some evidence. So certainty is equated with deduction; the other major evidence relation open to classical realism is induction, which lends no certainty. Hence all assertions based upon less than clear deduction are inherently doubtful. The conceptualisation of "conclusive evidence" utilised in classical realism fails to differentiate two different senses of the phrase, as identified by Baker (Baker 1974, Baker 1977). The first is: evidence is conclusive if and only if it is not logically possible that it can be overturned; and the second is that evidence is such if and only if it is not logically possible that it be improved. In classical semantics, if the evidence for a proposition can be overturned, then it can be improved; and vice versa.

In contrast to this, anti-realism can facilitate a distinction between the two senses of conclusive evidence, and thus derive a parallel distinction between necessity and certainty. Baker claims that the provenance of the term "necessary" is already established thus: X is necessary relative to Y if and only if X follows from Y; that is, if and only if evidence

making Y certain is incontrovertible evidence for X. So "certain" might be applied to the other sense of conclusive evidence: X is certain relative to Y if and only if Y is evidence for X that cannot be strengthened or improved upon. And it is just this distinction that is at the heart of the notion of C-support: C-support is conclusive in that it gives grounds for certainty, even though it is weaker than entailment, and hence can be undermined. Hence, certainty might be construed as an epistemic attitude denoting the practical belief that the evidence cannot be improved upon, and indeed that there is no point in attempting to do so.

The twin of the classical notion of conclusive evidence is the classical notion of doubt. Here, the *possibility* of doubt is treated as equipollent with actual grounded doubt. As Baker notes,

If certainty is conflated with necessity, then a demonstration that Y *may* be true and X simultaneously false is sufficient to show that the knowledge that Y is true does not make X certain. Consequently, the mere possibility of doubting X while admitting Y defeats the claim that certainty for X can be grounded in the knowledge of Y;
(Baker, 1974: 181).

C-support constitutes evidence which confers certainty upon a claim - even though it does not result from an entailment; to doubt C-support, it is necessary to adduce further evidence which serves to negate that C-support. The logical possibility of doubt cannot connote an actual doubt. To rephrase, the mere fact of defeasibility (the *possibility* of defeat) cannot justify either doubt or the denial of certainty. The presence of defeating conditions justifies doubt, but the mere intelligibility (or imaginability) of defeating conditions does not. Indeed, the supposition that defeasibility justifies doubt is tantamount to the absurd proposal that possible doubt is a kind of doubt; that imaginable reasons for doubt in *other* circumstances are, in *these* circumstances here and now, actual reasons for doubting.

Harman (Harman 1986) notes another way of phrasing this distinction. The Principle of Positive Undermining states, "One should stop believing X whenever one positively believes one's reasons for believing X are no good." This is essentially equivalent to the anti-realist notion of doubt. The Principle of Negative Undermining states, "One should stop believing X whenever one does not associate one's belief in X with an adequate justification." This is equivalent to the classical notion if "adequate justification" is read as "adequate deductive support".

The question then arises: which, if any, of the concepts of certainty does C-support trade upon? I think that the separation of certainty from necessity cautions against an unmodified η , yet it does allow a role for an interpretation of λ . And ι and κ are also inappropriate. Hence, it seems that the notion of certainty in play in the C-relation incorporates something

of types θ and λ .

It might be objected that a defeasibly necessary relation can readily be captured by classical means, so C-relations are superfluous. That is, a defeasible concept could be represented by the conjunction of the necessary and sufficient conditions for the term's application, with the negation of the disjunction of the various defeating conditions. There are difficulties with this method. The first problem is that it misplaces the burden of proof: a defeasible concept would be applicable if and only if we could positively show that none of the defeating conditions obtained. But this is simply an application of the law of excluded middle. The lack of evidence verifying the defeating conditions cannot be taken as verifying anything else: a positive verification of the C-support conditions is necessary. And note that this then implies that the concept will be considered verified unless there is evidence to the contrary: but there is no *requirement* to search for defeating evidence. The second problem is that it interposes context of "absence of specified defeating conditions" between the evidence and the claim. And this is once again inappropriate if it is an attempt to describe the way in which people reason. It is further shaken by the possibility, canvassed by Harman, that the attempt to capture reasoned revision of belief (i.e., defeasibility) in terms of logics represents a category error. His argument is that the process of fixating and changing belief operates on a non-logical basis; the process of argument is the proper arena for logics.

A further difficulty is that entailment is transitive, whereas, when part of the meaning of a word, the C-relation is not. That is, using \rightarrow to note the material conditional, where $(A \rightarrow B)$ and $(B \rightarrow C)$, the inference to $(A \rightarrow C)$ does not reduce the level of warrant for C: it preserves truth and hence can be completely certain. But, on an anti-realist view, the sense of a term is not exhausted by its C-relations: it must also involve, minimally, symptoms (or S-relations: inductive support). Now consider the situation where information p stands in a C-relation to concept A (which latter comprises at least criterial and symptom information), and where A functions as criterial information for B (which will again have its own S-relations), and B as criterial for C. So the criterial information for C is partly composed of B's symptom information. The further away from the initial p we move in the chain, the more deeply-embedded (and hence, as it were, "less criterial") becomes that first p . Since warranted assertion is based solely on the C-support, warrant is not transmitted along the chain by S-support. So where the chain of justification is initiated by such a p support is not transmitted in the same degree of strength: hence a *complete* transitivity cannot be said to hold. This partial transitivity is a direct result of the fact that the C-relation represents a content-driven inference. A further difference between the C-relation and entailment is

discussed by Phillips (Phillips 1978). In an entailment ($A \rightarrow B$), the truth of the evidence (i.e., of A) guarantees the truth of the claim (B): true evidence entails true claim. But it is possible for A to be (true) criterial evidence for a false claim: the example Phillips gives is that there may be evidence that a man has committed a crime even though he is actually innocent. In addition, the C-relation does not support contraposition, either: if particular C-support is successfully challenged (and hence we cannot infer B from support A), this does not imply that evidence A no longer holds: it is just that, with the addition of further information, it ceases to perform as C-support for B. In contrast, entailment, captured through the material conditional, supports "modus tollens" as the contrapositive of "modus ponens".

The C-relation is an internal relation, and hence could not be expressed through a context-sensitive deduction. Consider Baker & Hacker's (1984) comment (Baker and Hacker 1984):

One deeply rooted misconception of the sceptic consists in searching for grounds supporting what is in effect an internal relation. He looks at what are called grounds for saying that a person is angry, and demands *further grounds for taking these grounds* to support the statement that this person is angry.

(Baker & Hacker, 1984: 98; italics theirs)

The C-relation is just such an internal relation. But what does this amount to? Most crucially, an internal relation between two entities cannot be decomposed or analysed into a pair of relations with some independent third entity. Nothing external to the two related entities can mediate between them, since this would make the existence of the internal relation depend upon the existence of a suitable third entity, and its relations with each of the given internal relata. And this would mean either that the third entity is externally related to the internal relata (so the internal relation would be distorted), or the third entity is internally related to them: hence it would be redundant. The C-relation is an internal relation in that, at the level of phenomenological experience, some state of affairs is taken *directly* as evidence for another state of affairs' being the case. However, in terms of a theoretical explanation of action on the basis of the C-relation, the notion of context-sensitive inference may be the most apt description.

There is nothing in the exposition so far to indicate that the internal relation cannot be a deductive one: and indeed, this may be the case. However, the particular focus here is on the C-relation. And this has the added complication of being a *defeasible* internal relation.

As already stated, C-relations both carry certainty (they are internal) and are defeasible. They are the observable conditions for the application of a concept. But how can such non-

entailment relations be necessary? Baker & Hacker suggest that the certainty of the relation holds independently of its defeasibility. Defeasibility concerns the conditions under which we judge its application to be unwarranted. But when its application *is* warranted, the relation carries certainty. The evidence is not indefeasible but it is incontrovertible when it does obtain. In being evidence which cannot be improved upon it renders the claim certain.

The major aspects of the C-relation and defeasibility that I have mentioned are then as follows. The conditions of application of a defeasible concept can be given by outlining the set of conditions *normally* sufficient for its application, together with the conditions which could defeat its application. The latter will be an open list of instances of a list of *types* of defeating conditions. The default assumption is that the application is appropriate wherever the C-support is met; in these cases, the onus of proof is on the challenger of the application. And the multiplication of support cannot strengthen the certainty of a C-supported judgement. Defeasibility is one way in which the C-relation departs from the law of excluded middle (i.e., evidence for the claim does not necessarily mean that the defeating conditions can be denied outright); the requirement of positive evidence for the the C-support is the twin of this: denial of the conditions of defeat is not equivalent to support of the conditions for application.

I have commented that C-relations hold between conditions of application of a concept, and the conceptual content inferred. In the ensuing sections, I will pursue the idea that the same kind of semantic relation may obtain between a term (or its use) and its semantic content (or justification/warrant). That is, there may be C-relations between a term and its support, between a concept and its properties. Having looked briefly at C-relations and defeasibility, the next step is to examine the kind of network of relations that can support the individuation of semantic content. This will require a diversion into the issue of semantic holism and molecularity.

I.3.3. Holism and Molecularity

In this section, I consider the issues of semantic holism and molecularity. After sketching the main aspects and problems of holism, I discuss an approach to molecularity, and illustrate its relation to C-relations.

I.3.3.1. Holism:

The two major exponents of semantic holism are Quine and Davidson: For current

purposes, I treat their positions as of a piece. A quote from Davidson sets the scene (Davidson 1984):

we can give the meaning of any sentence (or word) only by giving the meaning of every sentence (and word) in the language.

(Davidson, 1984: 22).

This is the essence of holism, and it may stem from one (or both) of two sources. The first source (labelled "constitutive holism" by Tennant) is founded upon Quine's notion of the web or network of language, in which

there is no adequate way of understanding the statement short of knowing the entire language...it is not that a statement or even a theory has, as it were, a primal meaning which then gets modified by the interconnections that are established with other statements and other theories; rather, its meaning simply consists in the place which it occupies in the complicated network which constitutes the totality of our linguistic practices.

(Dummett, 1978: 218).

In the second source, as Tennant notes, "the emphasis is on how it is impossible to separate belief from meaning, to separate grounds for assent or dissent from shift in content of the statement at issue." (Tennant, 1987: 47). Dummett labels this Quine's "inextricability thesis" (Dummett, 1978: 349): the thesis that meaning is always and everywhere "contaminated by theory", that there can be no distinction between analytic and synthetic, between knowledge and word-meaning.

What is the relation between the two types of holism? As Tennant (1987: 50) argues, we could deny inextricability whilst maintaining some form of constitutive holism. But notice the different implications of these types for a theory of lexical semantics. Under inextricability, there can, *in principle*, be no way of separating out elements that are more or less inherent in the word, since knowledge and meaning are not distinguishable. In contrast, under constitutive holism, the problem is simply one *in practice*: there is no a priori reason why principled distinctions should not be made. And this opens up the possibility of a molecular theory of meaning, which I consider below.

This undifferentiated view of holism is countered by Dummett. The main objections he cites are as follows. Firstly, it cannot account for how we do, in fact, understand new statements. Secondly, holism by inextricability does away with the "periphery/interior" metaphor - one of the main motivations for accepting the network model as useful. Thirdly, holism cannot give an account of how we manifest our knowledge of language, for I cannot know anything that someone knows until I know (or guess) everything that he knows. And fourthly, holism cannot account for language acquisition, since this involves learning what justifications are required for sentences of various kinds. A similar range of arguments

is adduced by Fodor (Fodor 1987).

We have arrived at an impasse: it must be conceded that some form of holism is inevitable for at least some domains, yet on the terms of Dummett's criticisms holism appears incorrigible. We could, however, adopt constitutive holism and disavow inextricability; a consideration of the above criticisms indicates that the impasse stems from inextricability: without it, the "interior/periphery" metaphor remains tenable, and some way of addressing acquisition and manifestation might then be proffered. In examining this possibility, I shall lean heavily on Tennant's analysis of molecularity.

I.3.3.2. Molecularity:

How, then, are we to characterise a molecular theory of sense? Consider some comments made by Dummett:

Individual sentences carry a content which belongs to them in accordance with the way they are compounded out of their own constituents, independently of other sentences of the language not involving those constituents.

(Dummett, 1978: 222).

A grasp of the meaning of any sentence must...depend upon a mastery of some fragment of the language, a fragment which may, in some cases, be quite extensive.

(Dummett, 1978: 304).

Molecularity is an extension of the "context principle" of Frege, noted by Tennant (1987: 42) as: "never enquire after the meaning of a word in isolation, but only in the context of a whole sentence".

Tennant begins by considering the notion of a conceptual scheme, as a structuring of conceptual dependencies. Let R^* be the ancestral relations of R , where R is a two place relation indicating a semantic or conceptual dependency of the first item of the relation on the second. The idea is that, in order to understand the first item, we need also to understand the second. Thus, for example, if

$$\{C_1, C_2\} R C_3$$
$$\{C_4, C_5\} R C_6$$
$$\{C_3, C_6\} R C_7$$

then $\{C_1, C_2, C_4, C_5\} R^* C_7$.

The crucial question in respect of such dependencies has been phrased by Weir (Weir 1985) as one of whether the process of tracing along these lines is decidable - whether it comes to an end after a finite number of steps. If it does, the relation is "well-founded". So

a relation R is well-founded iff

(i). there are no R -cycles (or loops):

$$x_1 R x_2, x_2 R x_3, \dots, x_{n-1} R x_n, x_n R x_1; \text{ and}$$

(ii). there are no infinitely descending R -chains:

$$x_{n+1} R x_n, \dots, x_3 R x_2, x_2 R x_1,$$

For an inextricability holist, the relation of justification could not be well-founded. Tennant adds the notion of separability. Intuitively, this concerns whether:

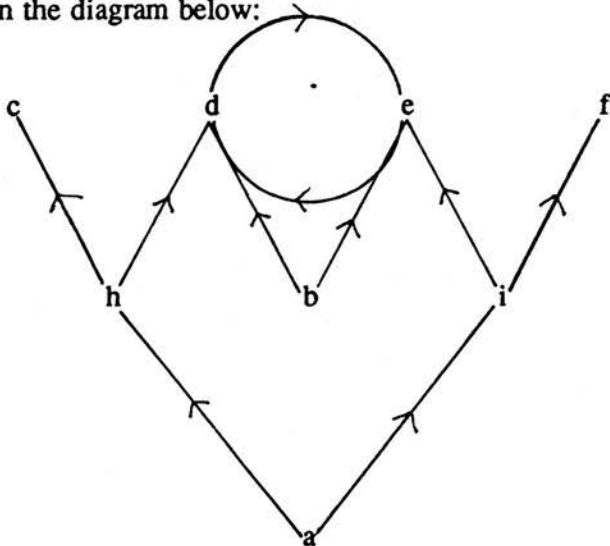
As we trace along the dependency relation R ...does the "covering umbrella" of ancestral concepts reached from any given concept mushroom out, well-foundedly or not, in such a way as take in all the concepts in...the basis of the whole conceptual system?
(Tennant, 1987: 56)

If it does, then the network is holistic; if not, molecular; that is, for molecularity, the relation is separable, whilst for holism it is not.

Tennant elaborates on this as follows. Consider any domain, D . Consider a concept $a \in D$. The *pedigree* of a will be the set of all R -ancestors of a ; the *foundation* of a is the subset of the pedigree consisting of R -terminal points; the *ground* of a will be the intersection of all the pedigrees of points in the pedigree of a . Intuitively, it is the set of all those non-terminal points on which a , via R , ultimately depends. The *basis* of a will be the union of the foundation of a with the ground of a . The *basis of D* will be the union of all bases of points in D .

Now, R is separable with respect to a in D iff the basis of D properly includes the basis of a .

I will reproduce an example of Tennant's to clarify this. Consider the relation R as indicated in the diagram below:



holism leaves as a practical problem the possibilities of soft molecularity and/or soft holism.

In order to highlight the contrasts, I shall give some examples. An example of "holistic" explanation offered by Peacocke (Peacocke 1979) exemplifies soft molecularity: the explanation of action by appeal to beliefs and desires. The assumption is that belief and desire are connected by an a priori principle which links the desire for X with the belief that a set of actions can bring about satisfaction of that desire. The point is that the concepts of belief and desire are wholly interdependent. That is, in Tennant's terms, they form an R-loop; whilst it may be separable, any instance cannot be well-founded. An example of hard molecularity may be semantic/lexical field analyses (for example, of kinship terms). Here, the domain is separable, and Grandy (Grandy clearly implies that it would be well-founded. Other examples include standard analyses of noun meanings in terms of category or class-inclusion relations: there will be definite termination-points (the semantic primitives), and such taxonomies are separable. I am not certain that any reasonable example of hard holism can be found, whilst a good example of soft holism is the kind of explanatory scheme envisaged by those who promulgate a unitary science programme: ultimately all of scientific endeavour is to be reduced to the same (physicalist) primitives. Thus, science is well-founded although non-separable.

In an anti-realist account of word-meaning, we must accept cognitive limitations on holism. Tennant argues (Tennant, 1987: 64) that we must combine a hard molecularity (in my terms) regarding the logical primitives, with at least a soft holism, if not a hard or soft molecularity, on extra-logical content words.

Tennant argues against a non-holistic view of the extra-logical terms by appeal to one example only; this example is one which may be atypical of the semantic structure of many other domains, and with respect to which Tennant disregards data from cognitive psychology which undermines his conclusion. He claims, "The mastery of any one colour-word requires mastery of all (or at least some of) the others in the language" (Tennant, 1987: 63). As the distinctions above indicate, there is a vital difference between the understanding of a term's requiring all, or only some, of the other terms in a language or domain.

Evidence from cognitive anthropology and psychology, for example, that of Berlin & Kay (Berlin & Kay 1967) and Kay & McDaniel (Kay and McDaniel 1978), might support the view that the colour lexicon is at worst a soft holism. The boundaries between the colour

terms are fuzzy; this does not necessarily support the non-separability of colour terms - for an identification or description of an object as blue-green does not require traversing the dependency chains for all other colours. And, within each segment of the colour space (for example, in the range we label red or blue) there is always, across all subjects from all tested cultures, agreement concerning the "focal" colour. And these focal colours are the "anchors" for processing colour terms and information: objects were labelled with colour terms by distance from the focal colour; once this distance has been measured, and a judgement made, the chain of dependence can terminate. It seems then that the colour lexicon is well-founded. But is it separable? I have already claimed that it is partially separable - in the sense of requiring only the concepts of two adjacent focal colours; but there are some concepts which are possessed by all colour concepts - those of hue and brightness. So processing colour terms does not involve processing all of the concepts of all of the terms - only one or two higher-order concepts which are common to all colours. The colour lexicon seems to be somewhere between soft molecularity and soft holism.

My claim is that Tennant is a little hasty in assuming that only the logical operators are amenable to molecular analysis. My argument in the rest of the thesis will be that one route to molecularity lies in the kinds of constraint on the R-chain which psychological studies of concepts have shown to be necessary.

Having put forward a view of a molecular network, my next goal is to argue that C-relations, as the central components of word meanings, can themselves support molecularity. I will approach this through an outline of Peacocke's notion of "canonical links", of which C-relations may be seen to be a particular type.

Peacocke (1986: 48) argues that a canonical commitment of a conceptual content that p is something which, if a thinker finds it fails to hold, constitutively requires her to judge that $\neg p$, for the reason that it fails to hold. The schematic form of a canonical commitment may be represented as a 4-tuple:

<informational state 1, content state 1, informational state 2, content state 2>

So if the particular content ascription is not to be rejected, then *if* (a) certain contents hold, and (b) the thinker is in certain mental states, *then* (c) certain other contents will hold, and (d) the thinker will be in certain other mental states. Hence, an individual may accept content (a) and be in mental state (b), and *only if* there is reason to reject either (c) or (d) will she judge that the content does not hold. Such a 4-tuple can also capture the notion of canonical grounds of a content: where (a) and (c), and (b) and (d), are accepted and "thought" respectively (the scare quotes are to indicate that no assumption of conscious

inspectability is intended), then she is *rationally required* to judge the content that p is the case. Since the same structure is to be assumed for both canonical grounds and commitments, the general locution of canonical links is used. And the notion of a canonical link is the same kind of dependency-in-justification link upon which the definition of molecularity is based.

Now, Peacocke distinguishes between "lower-level" and "same-level" canonical links. A thought has lower-level canonical links if someone could have all the concepts from which the contents in its canonical links are composed, without having all the concepts of a given thought. If all concepts have lower-level canonical links, this issues in a justificational reductionism. Indeed, if concepts had *only* lower-level links, the system would necessarily be non-well-founded. Peacocke therefore introduces same-level links. An implication of this is that two concepts are *evidentially equivalent* if and only if they have the same canonical links. An identity function should be defined over all concepts which are not above any of the concepts in the range of the mapping, Peacocke contends. In this way, the manifestation requirement operates as a constraint on the individuation of contents.

In order to apply these ideas anti-realistically, I shall discuss the way in which C-relations, as particular kinds of canonical links, can support molecular network structures. More specifically, the C-relation must be able to underpin three different networks:

- (i). separability and well-foundedness (Hard Molecularity).
- (ii). if non-well-foundedness, then separability (Soft Molecularity).
- (iii). if non-separability, then well-foundedness (Soft Holism).

Consider an assertion; understanding this generates an appropriate concept, which is composed of both C-related and S-related information. The successful tracing of the C-relations determines whether the whole concept is supportable. The 4-tuple is thus: we have an assertion about a state of affairs, with its attendant mental state (conceptual representation); this state of affairs has certain canonical links, expressed through C-relations; where the canonical links can be traversed without defeat, the transition to the asserted state of affairs (and its conceptual content) is permitted. S-relations, as context-sensitive, inductive relations, are necessarily lower-level, and do not form part of the justification-conditions for the claim in general.

C-relations may be same-level or lower-level: where a justification is well-founded, there is a reduction of level in the C-relation: the justification does not involve a dependency loop, nor does it descend ceaselessly. I argued in the last section, that, as part of a word's

justification, the C-relation is not completely transitive. This means that it can support well-foundedness, since the chain of support must end somewhere, with the initial warrant becoming less and less pertinent. The possibility of local holisms, or loops in dependencies, does not result in global non-well-foundedness.

So, can the C-relation support a separable justification chain? It seems that separability is at the behest of the particularities of the semantic content used on any occasion. So, different domains may be separable or not, as may particular content relations, and a single domain or relation might be either depending upon the particular processing context. There is nothing in C-relations, since they are contentful dependency links of concepts, to necessitate their being separable; but nor is there anything to stop them.

So it seems that condition (i), well-foundedness and separability, may be a natural attendant of a C-relation. But what of condition (ii)? This is the case where there are justification-loops, and hence C-links obtaining between same-level conceptual contents. But can this situation support separability? The answer must be in the affirmative. For since they *are* in a loop, all that can be taken into account in analysing the conceptual dependencies involved are the very concepts which form the loop; since a loop is by definition finite, the justification could not be non-separable. There must always be other concepts lying either outside the perimeter of the loop, or beneath it - and which therefore could not form part of the loop's justification.

The final condition, that if the justification is non-separable then it must be well-founded, seems equally amenable to criterial interpretation. Consider the two sources of non-well-foundedness: infinitely descending chains and justification loops. The latter simply is not possible in a non-separable structure, for a loop is, by definition, separable. The former is again answered by the incomplete transitivity of transmission of C-support.

It seems, then, that the C-relation in anti-realist semantics can provide a reasonable interpretation of the functioning of the three types of network above. The question remains: could a realist construal of lexical meaning, with deductive/meaning postulate type links achieve the same effects? I think the response to this must be negative. For the deductive relation is inherently transitive, and hence some added constraint would be needed to promote well-foundedness.

Given the above discussion, I should like to suggest that the concept of separability be redefined, and also that it be viewed from a slightly different perspective from that

evidenced in Tennant's analysis. The definition as it stands allows that concept *a* is separable with respect to domain *D* even if there is only one justifying concept more in the basis of *D* than there is in the basis of *a*; and there is no constraint on how remote that concept can be from *a*. I shall label this "Weak Separability".

"Strong Separability", as I shall label it, adds two constraints to Tennant's definition. Firstly, a concept *a* is separable with respect to domain *D* iff the basis of *a* is properly included in *D*, and the complement of *a* with respect to *D* contains the basis of at least one other concept which is of the same or higher level than *a*. Such a *Contrastivity Criterion* is critical to any analysis of the cognitive representation of concepts, as the abundant work on cue validity has indicated. Secondly, an *Accessibility Criterion* would require that the concepts which form *a*'s basis be part of the agent's explicit beliefs, or a part of her implicit beliefs which is C-related to the current explicit beliefs (Harman 1986). And the contrastive concept's basis should be at least embodied in implicit beliefs which are S-related to the ongoing explicit thoughts.

Strong Separability entails Weak Separability, but not vice versa. If a particular concept is not weakly separable, it cannot be strongly separable, although if it is not strongly separable it may nonetheless be weakly separable. In the discussions above, I have presupposed strong separability; and the content-driven nature of the C-relation is ideally suited to satisfy the Contrastivity Criterion. Henceforth, I shall assume strong separability rather than weak. Although we could define different types of molecularity and holism depending upon whether they utilise the strong or weak versions, the current commitment to cognition requires that we adopt the strong version.

So much for the redefinition of separability. A possible re-orientation on the idea of a molecular network involves construing such networks in a dynamic framework. For example, the meaning relations may be (largely) generated on a particular processing occasion; the particular contents which are linked to a concept as justifications are highly dependent upon the particular task. This has two aspects. Firstly, where only a somewhat cursory comprehension is necessary, very few of the possible semantic links of a term will be traversed. Such "quick skimming" is just the kind of processing which promotes well-foundedness. The second aspect is somewhat more interesting: different tasks will provide different perspectives on a conceived object; these are analogous to different "modes of presentation" of the object. And these will focus on different possible qualities of the object, determining different justificational subnets, and hence different possibilities for separability.

Molecularity (or soft holism) is, therefore, both cognitively and anti-realistically necessary. This is directly captured by C-relations operating under a cognitive constraint. One way of construing the possibilities for different individuations of content is through different "Perspectives" on the same entity (with the same canonical links). And the chief way of addressing this, on the current view, is through considering word meanings as partial objects.

I.3.4. Anti-Realism and Partiality

The goal of this section is to indicate the relationship between anti-realism and a conception of word meanings as partial objects. In this, I draw upon the work of Landman (Landman 1986). Landman claims:

the basic notion in semantics...[will be]...truth/falsity on the basis of the (partial) information. To some this may seem like a harmless change of terminology, but that is only an illusion: it is...realism that is at stake here.

(Landman, 1986: x).

And he makes explicit the connection between partiality and a rejection of certain aspects of classical logic:

[Semantic] objects...are partial objects, that may grow into total, real objects. Our semantic theory has to regard them as such, and...that is impossible if we accept the static ontology that is imposed by classical logic (in particular, the classical theory of identity).

(Landman, 1986: x).

How does this cluster of notions regarding partiality relate to anti-realism in general, and C-relations and defeasibility in particular? The answer is surprisingly straightforward.

Anti-realism is as much about information-lack as information-possession: that is, what we can and cannot do when our information is incomplete. In denying the "independence thesis", anti-realism commits itself to proffering at least some position on the relation between our evidence (as partial information) and what we can be justified in saying the world is like (as less partial information). Criterial relations were advanced as a means for effecting precisely this. That is, they are relations between the conditions of application of a concept (i.e., partial information constituting evidence for the classification), and the conceptual content which is inferred about the object being classified (again, a state of greater information concerning the object).

And a defeasible semantic relation just expresses that the transition we make in satisfying the relation is one based not upon total information but upon partial information. That is, it

expresses our having only partial information about the connection between our evidence and the claim for which the evidence is used. The only way of having total information concerning such transitions would be if the relation were a deductive one. And the anti-realist view is that, even though our information about the transition falls short of totality, this does not thereby automatically cast doubt upon the transition.

Another way of indicating the connection between partiality and anti-realism is simply to look at the aspects of classical logic which anti-realism is forced to reject. Consider double negation elimination: $\neg \neg A \rightarrow A$. Anti-realism, requiring that we have evidence for assertions, as much as for denials, denies that, because we have insufficient information to negate A, we thereby have sufficient information to assert A. And again, excluded middle, $A \vee \neg A$: the claim that, if we have insufficient evidence to assert A, we thereby have sufficient evidence for the denial of A. Both laws depend upon our having total information concerning the assertion or denial of A: only in such a limit case could these laws apply.

So there is a clear connection between anti-realism in general and in particular, and the concept of partiality. We need, now, to specify what we mean by "partiality" and "partial objects". On classical logic, we can typically equate an object with the set containing all of its properties. These will be:

(i). Maximally Consistent;

(ii). Total Sets: for every property P, either P or $\neg P$ is in the set.

Classically, a metalanguage is turned into a domain (structure) of total objects, which are identical to themselves and to nothing else. Consider Discourse Representation Theory (Kamp 1981). DRT postulates a system of levels of representation: the syntactic level is mapped onto the discourse representation, which is then embedded in a model. It employs a classical view of identity statements: classical logic is the metalanguage for identity. So, at each level, the structure for identity is a classical one. We may then give identity conditions within the DR level (thus making too many distinctions); or we may give them in the embedding conditions, in terms of referents (thus making too few distinctions). This dilemma is a result, claims Landman (1986: 112), of the fact that classical structures for identity utilise total objects. Such a classical view is denied by the anti-realist perspective outlined above: firstly, concepts (of, for example, referents) cannot be construed as total objects; secondly, in II.2.3, I will argue that the understanding of a term may involve the interaction of a discourse referent and the sense; thirdly, and crucially, in IV.6, I argue that word-meanings are to be defined in terms of *relations* between a (mentally represented) categorisation and a referent; so identity is defined on the relations between levels, rather than within levels.

Landman argues that we need an idea of objects under "guises" (Heim 1982): we may have two non-identical concepts for the same referent. And these are best construed as partial objects: "Objects in disguise are typically *partial objects*; with respect to their identity conditions they are typically objects that can *grow into the same total object*". (1986: 112; his italics). Partial objects can, minimally, be construed as sets of properties which are not total. It looks, then, as if we need to adopt a non-classical view on partial objects, and thus a non-static account of their change of state, resulting in a non-classical construal of identity.

Partiality is construed in an informational sense: partial objects are not the real objects, but rather those real objects as they appear in our information: propositions/properties carry information about the real domain; partial objects, as sets of properties/propositions, are *approximations* of real objects.

Landman defines an Information System as a "structure of propositions about a domain", which is based on two informational relations between units of information: the relation of *information containment* which indicates when a proposition/property gives the same or more information about the domain as does another proposition; and that of *incompatibility*, which indicates incompatibility of two properties of an object: they cannot both be properties of the same referent. The resulting structure is a partial order on information states. The relation of information containment, \sqsubseteq , defines a partial order on objects or information states. This order is transitive, reflexive and anti-symmetric. The growth of partial objects is seen as the replacement of approximations to an entity by successively better approximations to that element: that is, the better approximations define information-states which include the initial partial objects.

Landman considers two kinds of function: application (of, e.g., a predicate to a concept), and identity. A crucial question is then, whether the functions are monotonic. If the function maps a partial object p into a value, say 1, and it maps extensions of p , $p' \supseteq p$ into the same value, then the function is monotonic. That is, having once fixed the value of a partial object with respect to a particular function, whatever way that partial object may change under changes of information state, the new object's value with respect to the function should not change. So the requirement of monotonicity is equivalent to the demand that extension should involve only information addition, and not correction - that extensions of partial objects should be indefeasible.

Landman claims that application is monotonic, and then argues that identity should be

considered so on the basis of this. In Chapter II, we will find that, if we construe application as a type of concept combination, then it is not monotonic: combination may involve the defeat of some properties of either or both of the concepts, and also the generation of new properties not present in the inputs. Such partial objects are the objects we assume in conversation, and which we may follow through changes in information.

Landman holds that partial objects are best viewed as "pegs". I shall, in Chapter II, argue that concepts are best viewed as pegs, but I will also claim that the evidence militates in favour of a non-monotonic view of their identity. Landman's central claim is that partial objects are Pegs. Pegs are objects in our information, which we "hang properties on". Pegs do not *have* any properties in the usual sense: rather, properties are *ascribed* to them, on the basis of the information available. Pegs have identity conditions only after they have properties ascribed to them. It is possible for different pegs to have the same property ascribed to them, but it is not possible for a single peg to have incompatible properties. So properties map pegs on facts, and information states contain facts. Since pegs obtain identity conditions only after having properties hung on them, different pegs often cannot be distinguished on the basis of an information state - the latter may not contain sufficient relevant information. This then leads to the issue of discernability, which is the key to Landman's account of the identity of pegs. A minimally necessary (though not sufficient) condition for the identity of two partial objects is that they be indiscernable; if they are discernable, they cannot be identical, and if they are indiscernable, it is possible that they could be identical.

The basic conception of discernability is intuitively straightforward. Consider two partial objects, d and d' : they are discernable if d has property P (in information state s) and d' having P is incompatible with some fact in s ; or vice versa. The two objects are indiscernable iff they are not discernable. Discernability, then, is always relative to an information state. Two objects which are indiscernable in state s may well be discernable in state $s' \supseteq s$. Landman employs a notion of *strong indiscernability*, which is defined as follows:

d and d' are **strongly indiscernable on the basis of s** iff for all properties P :

$s \Vdash P(d)$ iff $s \Vdash P(d')$, and $s \dashv\vdash P(d)$ iff $s \dashv\vdash P(d')$; and

$s \Vdash \text{must } P(d)$ iff $s \Vdash \text{must } P(d')$, and $s \Vdash \text{may } P(d)$ iff $s \Vdash \text{may } P(d')$.

This amounts to the requirement that d and d' be indiscernable with respect to both the properties available in the information, and with respect to the properties which they *must* and *may* have on the basis of the information. In a later chapter, I will argue that the account of indiscernability requires some modifications in order to fulfill the Manifestation Constraint. Landman claims that we cannot define identity conditions on the basis of the current

information state, s , since partial objects which are indiscernable with respect to s may be discernable with respect to some $s' \supseteq s$. So, in order to preserve monotonicity of identity statements, we need to discount this possibility. And this is simply achieved by asserting that two pegs are identical on the basis of s iff they are (strongly) indiscernable on the basis of *every* extension of s .

In addition to pegs, Landman advances the concept of an "alec", which allows for even more flexibility. Alecs are a kind of peg, one which plays more than one role in a discourse. So, "in the domain of pegs we will find alecs and the characters they play" (p. 143). Landman asks us to consider an actor in a film: he may play more than one role. If we chart the development of a film's plot as different stages of information growth, there will be certain points at which the actor, d , plays first one character, a , and then another, a' ; when he plays a' , he ceases to play a . But this cannot be seen as different stages on the same "branch" of information growth: for this would require that d change from a to a' and that some extension of that information could contain them both. And this would imply that d could change its properties whilst remaining the same object - and this would mean that d 's identity would be non-monotonic. And of course, pegs are not pegs if they are not monotonic.

But, to backtrack a little: what is the relationship between d (the alec), and its characters, a and a' ? They cannot be identical, for that is part of the point: d must be able to play a and a' , whilst retaining its "d-ness". This also means that they cannot be strongly indiscernable. Landman therefore utilizes a "one-way" version of indiscernability: the idea of an *indiscernable approximation*:

d is an indiscernable approximation of d' in s , iff for all properties P :

if $s \Vdash P(d)$ then $s \Vdash P(d')$, and if $s \dashv\vdash P(d)$ then $s \dashv\vdash P(d')$; and

if $s \Vdash \text{must}P(d)$ then $s \Vdash \text{must}P(d')$, and if $s \Vdash \text{may}P(d)$ then $s \Vdash \text{may}P(d')$.

The idea, then, is that d can be an indiscernable approximation of both a 's; both characters will have all of the alec's properties, in addition to their own individuating qualities.

I commented above that, in order to preserve monotonicity of identity, we cannot consider the alec to play the two roles on the same "branch" of information extension. Let me clarify this a little. Landman assumes that we have a fixed starting-point and a fixed end-point for the information-growth, but that there may be many different ways of getting from the former to the latter (i.e., many different branches). These alternative branches *do not* contain *different* information; rather, they denote different orders of acquisition of information; to pursue the film analogy, they are different ways of telling the same story. If an alec

plays more than one character, there are thus different ways of following him through the same story. One way of thinking about this is that pegs are (or can be) different ways of viewing the same referent; alecs are different ways of viewing the same peg.

A brief comment will suffice to indicate the kinds of interaction between Landman's view of partial objects and the anti-realist picture outlined herein. If a word meaning is construed as a peg, then its identity is defined through the relation between the current information ascribed to the word, and the other informations which could or must be so ascribed. The question must then be: what is the nature of this relation? My claim will be that the most important relation is the Criterial relation; this is defeasible and hence yields a non-monotonic account of identity.

I.4. Summary

I have argued in favour of semantic realism by a consideration of Dummett's Manifestation argument. This in turn depended upon a proper consideration of the role of cognition in manifestation. I argued that the manifestational capacities necessary for an agent's grasp of language were those of recognition-of-verification; or rather, ability to recognise a verification of an assertion. And these capacities were crucially related to cognition. Such capacities are open to error in their operation, and this was considered through Criterial relations and defeasibility. Criterial relations yield certainty in recognition/verification procedures, even though they are defeasible. The qualities of C-relations, as crucial aspects of word-meanings, promote justifications of the application of concepts that may be separable and well-founded. That is, they can support a network of molecularity or soft holism, which is necessary for a cognitively decidable view of semantics. The final claim was that anti-realism is crucially dependent upon a view of semantic objects as partial objects: that C-relations embody relations between partial information states, and less partial states; a relation which is itself partial-informational, and therefore defeasible. This consideration gave rise to a view of word-meanings as "pegs". That is, they are semantic objects that have content ascribed to them during processing, which can play the role of more than one types of entity during a particular discourse.

The particular kinds of content ascribed to a peg (and thus to an agent understanding a word corresponding to that peg) is the crucial issue of the Manifestation Constraint. This constraint is, I argued, a constraint on the individuation of content, and so does not yield a complete specification of all the content ascribed to a peg. Rather, it allows distinctions between types of content. Its effect on a cognitive theory of word meaning is the subject of

Chapter II.

CHAPTER II

Manifestation: Concepts and Cognition

In this chapter, my main goal is to provide an operationalisation of the Manifestation requirement in general, and the Manifestation Constraint in particular, as it applies to word-meanings and their combination. Within this goal, there are three sub-goals. Firstly (II.1), I outline the particular classes of combinations with which I shall be concerned. Secondly (II.2), I consider cognitive phenomena regarding concepts and their combination, and word-meanings. In doing so, I shall advance a particular position on each set of phenomena. After outlining and summarising this position (in II.3), I proceed to the third sub-goal (II.4). This is an interpretation of the position and phenomena as the expression of concepts' being partial objects. Sense generation, I argue, implies partiality. I then argue that there is a read way of interpreting the problem classes, under the Manifestation Constraint, if we allow for the (perspectival) relativity of categorisation judgements.

II.1. The Problem Stated

The basic question to be addressed is: what semantic knowledge may we ascribe to an agent when she understands (either comprehends or produces) the use of a word in an utterance? That is, how might we describe the mental representation of a word's meaning? In accordance with the Manifestation Constraint, the goal is not to provide an exhaustive account of all aspects of this knowledge: rather, any content specifications arrived at are intended to provide only conditions of individuation of one content from another. I will assay the issue through the cognitive psychology of concepts and their combination. Such phenomena will motivate the basic constraints on the account of word meaning and combination to be promulgated. An interpretation of these phenomena will provide an operationalisation of the Manifestation Constraint.

My main focus will be the modification of noun concepts by adjective concepts. Two reasons for making this decision are pertinent. The first is that, as Murphy (Murphy 1988) has argued, theories tend to make similar predictions about the behaviour of single concepts, but they may not be equally able to countenance combinations. Thus, "an account of complex concepts may be crucial for evaluating the many theories of concepts now extant" (Murphy, 1988: 529). The second reason concerns an implicit goal of an account of word meaning: this is to provide an answer to the question of the relative mutability of the

various aspects of the knowledge ascribed to the agent. This is directly considered by examining the behaviour of concepts in different combination contexts. •

We may rephrase the problem as one of giving an account of the knowledge which we ascribe to an agent when she understands a noun phrase; in particular, what are the relations between this knowledge and that ascribed in understanding of the head noun alone? For example, if an object can be described by a noun phrase, can it also be categorised as a member of the head noun? In order to address such questions, I shall be considering two classes of combination types. The first class comprises combinations such as

Fake gun
False promise
Apparent friend
Possible solution

Combinations involving adjectives such as *fake* and *false* are labelled "Privative" A's by Kamp (Kamp 1975). In this study, I will refer to combinations involving both *fakelfalse* and *apparent/alleged* as privative; and I will later make a case for labelling as "Functional Privatives" such analogues as:

Stone lion
Square basketball
Blue orange
Straight banana

The second class comprises combinations involving non-predicating adjectives, as in

attractive ballet-dancer
excellent swimmer
nice paint

Within the second class, I will restrict my attention to A's which do not have any obvious measure component: I will not, therefore, discuss *large*, or *hot* or any other overtly dimensional quality. Before indicating why I have chosen to focus on this range of constructions, I shall note some of their qualities.

The particular qualities of the problem classes can be expressed in the different relations between these sentence-types:

Type X: This is is a (Adjective)(Noun)

Type Y: This is (Adjective)

Type Z: This is a (Noun)

For Privatives, at least at a prima facie level, the inference from Type X to Type Z sentences cannot go through. That is,

This is a fake gun
cannot support an inference to

This is a gun

For certain uses of *fake and false*, the Type Y sentence may make sense, as in *fake gun*. However, for a restricted group of Privatives, the Type Y sentence may scarcely make sense; Montague (Montague 1974) notes examples such as *false friend*, *reputed millionaire*, *possible president*, *ostensible ally* and *alleged intruder*. The crucial point about these combinations, for the moment, is that, if an entity can be categorised as a member of the noun phrase, it seems that it cannot be categorised as a member of the head noun.

The non-predicating adjectives, which I henceforth label "Attributives", involve combinations in which the inference from Type X to Type Z is straightforward (that is, they are "Affirmative" combinations), but where the inference to Type Y sentences is problematic. Typical examples include

Arnold is a large fish

from which we can infer that

Arnold is a fish

but not that

Arnold is large

And similarly for *attractive*, and so forth. The intuition behind this class is that, whereas the adjective property may be ascribed to the noun entity as a member of the noun category (so, Arnold is large qua a fish), such an ascription to the entity simply as an entity (with no obvious restriction on the domain of individuals for comparison) is cast into doubt.

Why were the two problem sets chosen? Privatives were chosen for two complementary reasons. The first is that, if we are interested in the relative mutability of the different aspects of the knowledge ascribed to an agent in understanding a word, then it is important to choose a range of examples in which the interactive effects are marked, and upon which intuitions are likely to converge. So, for example, whatever the subtleties of *fake gun*, it is intuitively very clear that the adjective has priority over the noun: it is this effect which underlies the difficulty of the transition from Type X to Type Z sentences. So the problem classes were chosen so as to allow the analyses to conform to the Manifestation Criterion: the more overt the interactive effects, the greater the likelihood that content differences might be discernable. The second reason for choosing the Privatives is linked to this. Since we will not be concerned to provide an exhaustive specification of content, but rather to sketch some of the individuation conditions of the knowledge, then we need to investigate effects which are not on the periphery of the word's meaning, since these latter just may

not be relevant to the canonical individuation conditions. And it appears that Privatives do involve some direct effect upon the canonical links of the nouns in question.

For this reason, the study concentrates on Privative combinations. However, it does attempt some examination of Attributives, so as to gauge how far the proposed account for Privatives can generalise to other cases, in particular Affirmatives. This comparison is crucial since it will be argued that the Affirmative and Privative types are simply two extreme poles on the dimension of the effects of the semantic interaction of the noun and modifier.

II.2. Cognitive Phenomena

In this section, I discuss what I take to be crucial phenomena regarding cognition and word-meaning. I do not directly consider particular theories of concepts and their combination: this is deferred to Chapter III. Rather, I consider aspects of the behaviour of concepts that, under a particular interpretation, can go to operationalise the Manifestation Constraint; the Manifestation Constraint then forms the basis of the discussion of the various theories in Chapter III. I discuss seven different aspects in total: dichotomies between (i) "intensional" and "extensional", and (ii) Sense Generation and Sense Selection views; (iii) the extent and nature of context-sensitivity (in concepts/senses and reference);(iv) monotonicity of concept combination; (v) the roles of different types of knowledge vis a vis concepts and word-meaning; (vi) the crucial question of the individuation of different types of conceptual content; and, finally, (vii) the issue of concept coherence. Regarding each aspect, I comment on relevant research, and adopt a stance consonant with the naturalised anti-realism of Chapter I.

II.2.1. Intensional vs. Extensional Views

I discuss some intensional and extensional approaches in Chapter III: here, I will make some general comments. Extensional views are typified by set-theoretic accounts, in which concepts are construed as sets, and their combination is effected by some variant on set-intersection. The rejection of the utility of classical sets in modelling the cognitive processing of concepts and concept combinations forms much of the recent history of the psychology of concepts, and I shall not go into this here (e.g., (Rosch 1973, Rosch 1975, Rosch 1978, Smith and Medin 1981)).

An extensional alternative is in prototypes or fuzzy sets. There is, however, evidence that prototypicality structures are themselves highly context-sensitive: we might require, for

example, a different typicality structure for every different context (Roth & Shoben 1983). And it has been argued that typicality structures are not properly evidence for the nature of the representation of concepts at all (Lakoff 1987). Further problems with prototype theories stem from attempts to model the combination of concepts in fuzzy set theory (Zadeh 1965) (see III.3.1.2). The implication is that we must take account of the content of the concepts. And if content is important, we need some account of the organisation of that content, and the ways in which combinations are influenced by world knowledge.

The main conclusion of this area is that no strictly extensional theory - operating over concepts as unanalysed units by means of formal or quasi-formal rules - will be adequate.

An "intensional" view of concept combination is therefore necessary. Such an account need not adopt the realist assumption that "intension determines extension", that is, that the former can determinately circumscribe the set of entities to which a term may correctly apply. Rather, as a cognitive approach, it must question the determinacy of reference upon which the classical realist accounts are based. As Johnson-Laird (Johnson-Laird 1987) comments, a cognitive semantics cannot assume that there will be either a transparent or a uni-directional relation between the semantic knowledge which constitutes a word's meaning and the referents which the word might denote through the aegis of that knowledge. I make some comment on these issues in II.2.3..

Influential intensional theories of concepts and their combination have been advanced by Miller & Johnson-Laird (Miller and Johnson-Laird 1976), and more recently, and in the same spirit, Smith & Osherson (Smith and Osherson 1984). Both views propose a partitioning of the content associated with a term, and concomitantly differing effects in combination. Although the latter model can account for much of the data which confounded fuzzy set approaches, it appears to share one crucial problematic assumption with the set-theoretic views. That is, it assumes that the processing of concepts occurs through sense selection.

II.2.2. Sense Selection vs. Sense Generation

Clark (Clark 1983) states that sense selection accounts assume that senses for words are represented as stable, isolable packets of information, which make a constant contribution to any complex expression of which they form a part. Combination of concepts occurs through putting together two such whole packets. In contrast, the sense generation hypothesis assumes that the sense of a term is constructed on-line, during processing. The sense is generated in order to make an appropriate contribution to the discourse: the

generation of sense for a concept combination will follow a form of mutual constraint satisfaction for the two concepts.

Most of the recent psychological evidence militates against sense selection. I divide this into three distinct components: concerning *sense-invariance*, *post-access combination*, and *explicit semantic attachment*, respectively.

Sense-invariance is the assumption that we can and do access the same stable packet of sense regardless of the context. All of a term's sense is utilised on every occasion of use. Such a view has received criticism from many directions. For example, the evidence of Roth & Shoben indicates that the typicality of exemplars varies as a function of context. The work of Barsalou has been central to the downfall of sense-invariance. Barsalou (Barsalou 1982) has repeatedly found that two components of a noun's sense - context-independent and context-dependent - could be differentiated. So not all of a term's sense is invariably employed.

Parallel findings were reported by Barclay, Bransford, Franks, McCarrell & Nitsch (Barclay et al. 1974). For example, they argued that the representation of *ball* in these contexts

the man hit the ball

the man hid behind the ball

differs, in that the latter object must be seen to be larger in size. Their claim was that the individuation of senses must be relative to a particular context of use, rather than absolute. Similarly, Half, Ortony & Anderson provide evidence for the variation of the sense of *red* according to the noun with which it was combined (Half, Ortony and Anderson 1976). And Medin & Shoben demonstrated such an effect for a wider variety of adjectives. They found, for example, that the sense of an adjective is not independent of the noun context in which it is processed, and that the changes in the N concept were not restricted to the properties associated with the modifying N. They were, rather, propagated through the correlated attributes of the N (Medin and Shoben 1988).

An interesting proposal concerning such context-dependence has recently been mooted by Bartsch (Bartsch 1987): she has claimed that, for many terms, the precise "thematic dimension" which it involves is given by context: standard examples include *good* or *successful*. She argues that these terms encode a "pre-property", of a general level which the thematic context operates upon to yield a *bona fide* property. For example, *good* has many potential properties, which are narrowed down in *good paint*; but these also need to be narrowed down by the thematic context (a paint which is good from one viewpoint - say, a graffiti

artist's - may not be good from another, such as the person who has to clean the wall). Depending upon context, the semantic knowledge ascribed on the basis of *good* might involve either a general property, or a more precise one. This idea will be developed and utilised in the account of Privatives offered in later chapters. The role of viewpoints in determining which sense is constructed has also been supported by the findings of Barsalou & Sewell (Barsalou and Sewell 1984). Subjects regularly gave different typicality structures for concepts, depending upon the "viewpoint" or "perspective" they were asked to adopt (as a businessman, housewife, and so forth). The sense of both the adjective and the noun, then, is crucially dependent upon the other, and the role of individual properties will be equally context-sensitive.

A further challenge to sense-invariance is that we can easily process concept combinations in which the adjective properties would not ordinarily be considered as (possibly) true of the noun's referent. Gerard & Mandler found that people could easily interpret predications of properties to an entity which were in direct contradiction of the entity's ontological type (Gerard and Mandler 1983). They investigated Keil's (Keil 1979) claim that there is a hierarchical and one-to-one mapping of the predicate-object relations of language onto the basic structure of ontological knowledge. Keil argued that objects (representing the noun terms of natural language) are differentiated by type on the basis of the predicates that span them: if two predicates span one term in common, they must span all terms in common. A predicate spans a term if its application to the term makes sense - that is, can be adjudged true or false. For example, whereas applying the predicate *red* to *apple* should be perfectly reasonable, the view rules out its application to, say, *performance*, since the latter is simply of the wrong ontological type. However, Gerard & Mandler found no such obvious constraint on predication. Clearly, if we always access the same information for each term, there is no way in which such apparent ontological anomalies could be understood. This finding has far reaching effects, in at least three different areas. Firstly, it implies that the sense associated with a noun term is not inflexible: even a denial of some ontologically crucial aspect of the noun object is interpretable, and may not prevent the inference from the object's being of NP type to its being of head N type. Secondly, the organisation of knowledge which feeds the understanding process cannot support a single, simple hierarchy: there are considerably more complex relations at play, promoting flexibility of property ascription. Thirdly, any account of the coherence of concepts and concept combination must allow for the possibility of crossing ontological boundaries, whilst preserving coherence. That is, coherence of concepts is not simply a matter of reflecting the way the world is. If coherence is linked to theories of the world (Murphy and Medin 1985), these theories must allow for the interplay of language use and world knowledge. These issues

will be picked up in this and later chapters.

There is also evidence that we may generate very scant and inapt content in association with a term, though this content is nonetheless thematically appropriate. Erickson & Matteson found that, in replying to the question, *How many animals of each kind did Moses load aboard the ark?*, subjects consistently failed to spot the error. Such a failure persisted even when subjects were aided in various ways. (Erickson & Matteson 1981). What was important was that the inappropriate entity be semantically similar to the correct entity, as might be the case for Moses and Noah. It seems, then, that subjects simply generated sufficient content to support the coherence of the sentence. The precise nature and the amount of content ascribed to a word in a sentence may depend upon the role of that word relative to the agent's goals in understanding the sentence. Here, the thematic interest was the activity of the agent: the subject just was not interested in any detailed information concerning the biblical character himself.

The second aspect of sense selection - the assumption of post-access combination - is also countered by the findings of Medin & Shoben, Gerard & Mandler, and Halfp et al. Further, Potter & Faulconer (Potter and Faulconer 1979) found that the retrieval of information for the A and N was interactive and mutually constraining: the NP was constructed during retrieval, rather than subsequent to it. These results have been replicated recently by Lyon (Lyon 1989).

Similarly, Rips & Turnbull (Rips and Turnbull 1980) found that the senses of "absolute" adjectives (such as colour terms) and "degree" adjectives (including linear adjectives such as *tall*) were constructed and varied according to N context. Again, the arguments against sense-invariance appear to deny post-access combination.

The third aspect of sense selection - the assumption of explicit semantic attachment - is a special case of the sense-invariance assumption. Sense selection implies that the sortal type information with respect to which the A and the NP will be evaluated will inevitably be that of the N. That is, whatever referent the NP might have, it will nonetheless be construed as of the type of the N term's null-context sortal type (Platts, 1979). So any inferences which can be made on the basis of the N's sortal type can be made on the basis of the NP; and the interpretation of the A properties will take place with respect to that N sortal type. Consider an example given by Platts:

Rudy is an attractive ballet-dancer

On a sense selection approach, this would be represented as something like:



Rudy is attractive as a ballet-dancer (dances ballet attractively)

So we can then infer that Rudy can pirouette with the best of them.* This interpretation is based upon an assumption of explicit semantic attachment between the A and the (explicit) N. However, there seems to be absolutely no reason why the sentence should not be read as:

Rudy is a ballet-dancer who is attractive to look at

The ability to pirouette well cannot be inferred on this reading. Here, the semantic attachment of the A is implicit - to a N class which is not given in the sentence (the group of *people who are attractive to look at*). Evidence which contradicts the assumption of explicit semantic attachment comes from several sources. Firstly, Rips & Turnbull asked subjects to verify sentences such as:

Insects are big

Here, subjects vacillated between using two different N classes as reference or comparison points for size. Subjects used either the immediate superordinate of the explicitly given N term (say, *animals*), or a general "anthropomorphic" reference class. If the assumption of explicit semantic attachment were true, they should have invariably used the N type of *insects* (or, at worst, *animals*) as a source for evaluation of the sentence. Secondly, Barsalou (Barsalou 1983) found, regarding "goal-directed" categories, that, if we have a particular goal or motivation for categorising a referent, this goal-based category may be used instead of the term's usual sortal type. So, for example, my favourite novel by Samuel Beckett will be construed first as one of the *things to take with me in case of a fire*, rather than as *a book* or *a novel*: the properties which would be ascribed to my mental representation of *my favourite novel by Samuel Beckett* will involve its personal value, its flammability, and so forth. The more mundane qualities of a book will be of less import. And even though, for many reasons (stylistic, number of pages) this particular novel is not a typical member of *novels* or *books*, it is highly typical of this particular goal-directed category, and even more typical of the category of *books to take with me in case of a fire*. So the sense generated with respect to the phrase describing this particular novel is crucially related to the current goal. The numerous possibilities for implicit semantic attachment clearly caution against simple adherence to the sense selection view. For example, my goal might be to classify *things with which to swat a fly*, or *things to stop a chair wobbling*, or *things for abstruse and elitist entertainment*: each category will prompt a different sense for the nominal phrase, *my favourite Samuel Beckett novel*.

It seems then that a sense generation approach is required. However, we need to be somewhat more precise about what this involves. For example, how many distinct representations in the mental lexicon are necessary? Another way of putting this is: how would a

sense selection account approach the issues of vagueness and context-sensitivity?

II.2.3. Context-Sensitivity

How might we construe the variability of a word's sense? In Myers Franks & Braisby (Myers, Franks and Braisby 1989), we argued that this amounts to a tension between two complementary aspects of the behaviour of sense: flexibility and specificity. A consideration of this phenomenon will facilitate a clearer statement of the contrast between sense generation and sense selection.

Consider the following example:

Rudy is at the ballet watching a single ballerina dancing on stage. Excitedly, he whispers to his friends, "Isn't she delightful?". His friend Ron agrees, replying, "Yes, she's very beautiful". His other friend, Reg, agrees too, saying, "Yes, what an exquisite pirouette". But Rudy replies, "I simply meant she's a wonderful person".

Delightful, here, appears to mean different things for the different individuals. From Ron's perspective, it means "physically delightful" ; from Reg's, it means "dancing delightfully", while for Rudy *delightful* means "delightful as a person". From the perspective of an usher who only hears Rudy's first comment, the sense is indeterminate with respect to these specific senses.

In order to characterise the semantic component of the lexical entry for *delightful*, we might propose that the lexical entry is sufficiently general to encompass all the different observed senses. Even if we could specify such a general lexical concept, however, this would not then allow the possibility for it to convey, say, the sense "delightful as a person". Yet, in use, *delightful* can convey just this. And any of the senses indicated above could be made more specific if more contextually provided information were available. For example, Ron's sense for *delightful*, "physically delightful" could be "physically delightful with respect to coiffure" if we know that Ron is a hairdresser. In principle, there is no bound to this degree of specificity. Clearly, we would also want to allow that senses can be indeterminate: for example, the sense for the usher is non-specific with respect to the senses for Rudy and his friends (or any other more precise sense). In conclusion then, our treatment of the lexical entry as general, being motivated by the observation of flexibility, renders the lexical entry unable to provide for the specificity of sense that we observe. Seemingly, we can't treat the phenomena of flexibility and specificity independently of one another.

Henceforth, I will refer to the semantic contribution of the lexical entry as the "lexical concept", and the intended meaning as the generated "sense". Now, we might distinguish between two versions of sense selection: the weak and the strong. One approach to characterising lexical entries is to treat the different senses of a word as different lexical concepts. For example, the lexical entry for *mother* might include several lexical concepts, corresponding to *foster mother*, *step mother*, and *biological mother*, in the same way as, say, *bank* is assumed to have a number of independent lexical concepts. A strong version of this view would hold that for each and every different sense a different lexical concept is accessed from the lexical entry. It is the strong version which Clark (1983) showed to be unworkable. A weaker version maintains that there may be independent, unrelated lexical concepts from which, on different occasions of use, different senses may be elaborated. For example, under the weak view, we could, in the context of a fairy tale, arrive at a sense for *mother* corresponding to "step-mother who is wicked" by elaborating the lexical concept for *mother* labelled "step-mother". Presumably, on the strong sense selection view we would select the appropriate sense, "step-mother who is wicked", which could not be further elaborated.

In the strong version, we have a granular picture of flexibility and specificity. Flexibility is granular in that the possible senses a word may have cannot alter and we have an impoverished view of flexibility. Specificity is also fixed by the set of senses in that senses can be no less or no more specific than those represented. The weak version allows senses to vary in terms of their specificity but we have only an apparent flexibility: the range of possible senses is restricted by the initial choice of lexical concept.

It seems, then, that in addition to the empirical problems for sense selection noted in II.2.2, such views cannot do justice to our intuitions concerning some of the basic phenomena of word meaning. In section II.2.7, I shall consider the failings of a sense selection account in relation to conceptual coherence.

What exactly does the sense selection/generation debate involve? It is not, properly, an issue of ambiguity, except insofar as true ambiguity is probably the only uncontroversial motivation for proposing multiple lexical concepts as the origin of divergent senses. This is important, since a possible rejoinder to the sense generation case might be that there is evidence that all of the different senses of an ambiguous word are accessed initially, and then a single one is chosen on the basis of contextual constraints (e.g., (Swinney 1979)). The argument might be that this process also operates for other words. But our intuitions on this are quite firm: it just is not the case that the difference between the two senses of *ball*

noted by Barclay et al, is commensurate with that between the two senses of *bank*. Rather than the former two senses being ambiguous alternatives, they are better seen as different elaborations of a single lexical concept.

The question to be considered concerns "whether or not it is necessary to postulate more than one semantic representation for a word in order to account for the interpretations of the sentences in which it occurs" (Johnson-Laird, 1987: 196). On present terms, this has two components. Firstly, how many different lexical concepts are necessary to account for the range of senses which a word may have? Secondly, how many different lexical concepts (and senses) are necessary to account for the range of referents which a word can have in different contexts? The first issue has received some consideration already, but we will not be able to outline a more detailed position until the end of this chapter. It is to the second that I now turn.

The issues of flexibility and specificity are largely concerned with the context-sensitivity of the concept or sense. The complement of this concerns the vagaries of meaning and reference. The examples which I mention here concern reference to a particular entity by a noun or noun phrase which might seem inappropriate.

I have various familiar phenomena in mind. Consider the example of metonymy noted by Nunberg (Nunberg 1978). A waiter uses the phrase *ham sandwich* to refer to a customer who has ordered a ham sandwich. The difficulty for any theory of word meaning is that the object referred to (the customer) has none of the properties which we would normally associate with a ham sandwich, and vice versa. Yet, by some pragmatic method used in that situation, reference is achieved.

Another kind of example involves reference to an entity which possesses some of the properties of the noun, but also has a denial of some of its properties. Consider someone talking about a lemon, when all that is visible is a pulpy mass on the road. Such a squashed lemon exhibits virtually none of the distinctive properties of a lemon. We can allow the defeat of many of the properties which we would normally ascribe to the sense represented by someone who understood the word *lemon*, and yet still apply *lemon* to an object possessing the remaining properties. Such a degree of defeasibility is implied by the findings of Gerard & Mandler. Later I shall argue that the degree of defeasibility outstrips the kinds of transformations we can ordinarily perform on the referents. And although a squashed lemon is still a lemon, we also know that, where a categorisation is made with the goal of counting whole lemons, or making a cocktail, a squashed lemon would not be categorised

as a member. This is taken up in the discussion of perspectives in Chapter IV.

Another kind of example involves Privative combinations, and is exemplified by what Braisby (Braisby 1989) labels the "lion puzzle":

Fred is sitting on a park bench in London. He knows that at the other end of the park there is a statue of a lion. A schoolgirl approaches him and, explaining that she has been given an assignment to sketch a lion, she asks Fred if he has seen one. Fred replies that he has and points her towards the statue. A little later an exhausted zoo-keeper appears and, explaining to Fred that a lion has escaped from the zoo, he also asks if he has seen one. Fred replies that he hasn't.

How should we account for this apparent contradiction in Fred's behaviour? In both instances, Fred appears to be telling the truth. A similar kind of scenario could be envisaged for something like a fake gun. We would not want to claim that *lion*, *gun* or *lemon* were ambiguous, and that Fred selects one sense to respond to the schoolgirl and another to respond to the zoo-keeper. The crucial point is that, although a stone lion or a fake gun or a squashed lemon all differ significantly from their respective head nouns, they are also closely related to those nouns. Later I shall argue that categorising an entity might be seen as adopting a perspective on that entity; and that, only if that perspective is supported by the entity is the categorisation warranted.

A similar kind of case has been discussed by Donnellan (Donnellan 1966). He argues that descriptions can be put to two uses - "referential" and "attributive". In the latter case, the speaker is attributing something to whatever fits the description: so the attribute of having a particular quality is all-important. On a referential use, by contrast, a speaker uses a description to enable her audience to pick out the object she is talking about. The description is a tool for effecting reference: it need not be strictly true of the entity. For example, a referential use could pick out a person as *the man drinking champagne*, even though he is drinking fizzy water. The application of the description to the man would thus fail attributively but succeed referentially. The crucial implication is that a misdescription can function satisfactorily in the referential use but not in the attributive. Some difficulties in Donnellan's account can be seen from the kinds of examples noted above. In the *lion* example, Donnellan would have to argue that the zoo-keeper's use of *lion* is attributive, whilst the school-girl's is referential. So although *lion* is a misdescription (the statue is not a lion), and hence cannot be used attributively, it can nonetheless succeed referentially. But my claim is that there is a limited sense in which the use is attributively true: in the same way that there is a restricted way in which a fake gun is a gun. Consider an example: a

cashier in a shop is confronted by a menacing intruder who brandishes what looks like a gun: so the cashier hands over the day's receipts. A customer, however, believes that the intruder is carrying only a fake gun, and hence overpowers him. The crucial point about this is that the object in question can be taken to satisfy the description of *gun* in ways which suggest that there is something more than merely coincidental in the ability to refer to it as a gun. The agents can act towards the object in ways which suggest that a description of it as a gun might be attributively true. That is, the object satisfies the description in a way which allows that more behaviours than simply the referential one can be effected on the basis of it. It may do so even though it does not satisfy an attributive use from a perspective incorporating all of our knowledge about guns. I think the same kinds of consideration might apply in the "lion puzzle". What this amounts to is a claim that attributivity, rather than being all-or-none, as Donnellan assumes it to be, is actually deeply perspectival.

So the relation between a particular word and its referent is not mediated solely by the sense of that word; rather, a crucial role is played by the local context as well as the knowledge and goals of the participants. The goal of specifying the number of different lexical concepts and/or senses required for the range of different referents of a word would seem, then, to be doomed - for first we would need to denumerate the possible referents for a given word. And given the range of contexts, desires, times, participants and other variables, such a denumeration is not possible.

The third aspect of context-sensitivity concerns the possibility of reciprocal interaction between the generation of a sense for an expression and the construction or assignment of a (discourse) referent to that expression. I shall focus on the notion of "instantiation" (Johnson-Laird 1983, Johnson-Laird 1987). Let us again approach this through a brief allusion to ambiguity. Johnson-Laird (1987: 195) gives the example:

The client received the cheque on Tuesday. He banked it.

Here, the verb *banked* is ambiguous. Johnson-Laird claims that the sentence is disambiguated easily because the reference of *it* is plainly the cheque of the prior sentence. The argument is that it is the referent, or knowledge of the referent, of the expressions which has a crucial role in disambiguation - rather than the particular sense of the word in question.

How far could such a view extrapolate to other words? That is, does the referent have an important role to play in the content ascribed to a word? There is much evidence to support this possibility. For example, Anderson et al (Anderson et al. 1976) found that, when

subjects were presented with target sentences such as

A fish attacked a swimmer

recall of the sentence was better when cued with *shark* than with *fish*. They argued that the sense of *fish* instantiated is here equivalent to *shark*. Garnham (Garnham 1979) has reported similar effects for verbs. The view of Anderson et al is, however, strongly akin to the sense selection view, since interpretation centres on the selection of a particular kind of exemplar from a list of possibilities. However, it simply is not clear that we would want to discriminate senses for a word like *bird* at the level of detail which differentiates *chicken* from *sparrow* or *thrush*. Johnson-Laird argues that

Instantiation is ... a process, not of eliminating senses from a list in a lexical entry, but of imagining a more specific situation than is warranted solely by the meanings of words.

(Johnson-Laird, 1987: 197).

Such situations then provide access to knowledge which supports the elaboration of the lexical concept. So *shark* does not *become* the meaning of *fish*, nor does it displace it: rather, it provides access to knowledge supporting an extension of the meaning of *fish*. On the terms employed here, the lexical concept is extended into a sense for the word on the basis of the instantiated entity. And a crucial possibility is that the instantiated entity may not bear a direct relation to the noun in a NP: that is, it may not, unlike the examples considered by Johnson-Laird, simply be a subordinate of the noun. This would be a direct manifestation of explicit semantic attachment; the possibility of implicit attachment opens up far more flexibility regarding the entities which can act as an instantiation of a noun or NP. This will be taken up later.

A similar notion has been promulgated by Hampton (Hampton 1987). Subjects generated many features for a complex concept which were not present in the lists for either of the constituents. He accounted for this by arguing that our knowledge of the referents of the complex concepts could be used to support ascriptions to the concept. This parallels instantiation: after the initial combination, the referent is employed as a source for elaboration of the combination's sense. Such post-combination elaboration, though utilising a wider range of knowledge, has been demonstrated by Murphy (1988).

II.2.4. Compositionality and Monotonicity

Another question concerns the cognitive plausibility of a compositional and/or monotonic account of concept combination. A standard definition of compositionality, from Dowty, Wall & Peters (Dowty, Wall and Peters 1981) is that the meaning of a complex expression is given by the meaning of its parts plus their grammatically correct mode of combination.

Each input must make a stable, iterable contribution to any complex expression in which it features. This is the "fusion" conception of compositionality, which should be separated from the "fission" conception, whereby the complex expression must be decomposable into the appropriate constituents. A combination of the two would be the strictest form of compositionality possible, in which there could be no contextual influence. In practise, virtually all semanticists would advocate some version of Frege's "context principle". In formal semantic systems, any contextual influence must be expressible through the "mode of combination" - through a general rule, which bears at least a homomorphism to the relevant syntactic rule. Such a rule must needs be too general to speak to the kinds of semantic interaction involved in concept combination. The arguments against sense-invariance, in favour of sense generation, all count against a compositional account. This is because compositionality assumes sense selection as a premise.

A direct test of compositionality in property inheritance has been provided by Hampton (1987). He found that Boolean mechanisms of conjunction, disjunction, and so forth, had no reflection in concept combination. There were several important restrictions on a simple Boolean summation of the sets of properties. These included the possibility that new properties, not in either of the inputs, might occur in the combined concept; and that if properties were either "impossible" or "necessary" for either of the inputs, then they would be absent or present, respectively, in the conjunction. The latter finding - of inheritance of necessary and non-inheritance of impossible attributes - might, as Murphy comments, be partly a function of the kinds of constructions which Hampton used. Other constructions might yield an even less compositional result.

The thesis of monotonicity is somewhat weaker than compositionality. Monotonicity is the requirement that any predication or identity statement which is true of concept **P** is also true of any complex expression of which **P** forms a part. It does not, therefore, discount the possibility that there may be new properties in the complex expression which were not in the inputs - with the proviso that the new properties do not render untrue any of the predications/identity statements which were true of the inputs. Hence, whereas sense selection is one of the premises of compositionality (and so to ask whether concept combination is compositional in a sense generation view does not strictly make sense), monotonicity is agnostic on this issue.

Monotonicity is implied by, but does not imply, compositionality. That is, if a combination is compositional, it is necessarily monotonic; but if it is monotonic, it need not be compositional. Nonmonotonicity, however, implies noncompositionality. And only a certain type of

noncompositionality implies nonmonotonicity: this is the noncompositionality which stems from a denial of some of the properties of the inputs in the combination: what was true of the inputs is no longer true of the combination. Nonmonotonicity is not implied by the non-compositional emergence of new properties.

In a nonmonotonic system, we can truthfully apply a predicate to a term under one state of information, and yet find that the predication is false under different circumstances. If predication (concept combination) is nonmonotonic, then concept identity must also be nonmonotonic. That is, we may ascribe some sortal type to an entity only to revise that ascription later. Or we may ascribe a sortal type on the basis of one state of information, and maintain that type even though the information changes, if we adopt a more coarsely-grained criterion for categorisation. So nonmonotonicity of concept identity and concept combination allows that we may add new properties, and deny old ones - without having to deny the appropriateness of the use of the term. Under a monotonic account, if we wish to deny some of the properties which we have ascribed to a term, we would simply have to abandon the use of the term - since a requirement on the use of the term is that we increment content monotonically.

It seems we can claim that both identity and combination of concepts is nonmonotonic. It is clear that this issue depends partly on what we take to be the starting point of generation (i.e., the lexical concept). The relationship between this, and the generated sense, will provide a clue regarding monotonicity. However, even if we argue (and I shall not) that *this* relationship is monotonic, the relationship between the generated sense, and that sense as it appears later in the discourse (after the addition of information concerning the themes and goals), may well be nonmonotonic. However, I claim that the first relation (between the lexical concept and the sense) may also be nonmonotonic. It seems to be clear that, under *any* theory of privatives, there will necessarily be some defeat of some of the the N's properties: and this will inevitably vouchsafe the combination as nonmonotonic. A further point is the assumption which I make concerning the lexical concept. I assume that this consists of knowledge relating to the usual referent of the N input to a NP: that is, I assume that we start from the premise of explicit semantic attachment. This content may then be overridden in the generation of a sense. Such overriding may take the form of a selection of only some of that content, or a shift to an implicit semantic attachment. Hence, both the sortal identity of the concept and concept combination would be nonmonotonic. A third reason for nonmonotonicity is the findings of Hampton. There is no *a priori* reason why his results should not apply to a sense generation view. And these support the non-monotonicity of concept combination.

II.2.5. The Role of Knowledge

Does world knowledge have an effect on concepts? That is, does our understanding of concepts/concept combinations make use of knowledge which might be "outside" the concept?

The properties ascribed to a concept may be influenced by our theory of the kind of object which the concept represents. There is evidence from the literature on how concepts change with expertise, to suggest that there are strong relations between the concepts in, and theories governing, a particular domain (Murphy and Wright 1984). And evidence from developmental studies suggests that theories play a crucial role in the acquisition of concepts. For example, children have rudimentary theories which support inferences about a domain, and thereby provide the basis for the learning of new concepts (Keil 1987). And, just as experts appear to make a shift from "characteristic" to "defining" aspects of an entity, as a result of elaborating and refining their theories, so children have been found to develop into using more and more sophisticated criteria for categorisation (Keil 1987). Similar evidence has been provided by Carey (Carey 1982). So world knowledge, often domain-specific, is implicated in the development of concepts.

Adult understanding of concepts is also knowledge-rich. In their highly influential paper, Murphy & Medin (1985) argued that hypothesised bases of categorisation - such as the similarity of exemplars, or correlated attributes within a concept - could not receive any theory-free analysis. Such constructs could not explain the structure or coherence of our concepts because the properties of those concepts themselves only inhere in them as a result of the theory of the domain. Similarity and other judgements are thereby parasitic on those theories. They argued, then, that the coherence of concepts, and therefore the coherence of the ascription of properties to concepts, is founded upon the relations between the concepts and the others in the domain. As a result, the theory of a domain and the conceptual content are intertwined.

That world knowledge is crucial to the understanding of concept combinations has recently been demonstrated by Medin & Shoben (1988) and Murphy (1988). Medin & Shoben hypothesised that, if combination is a knowledge-free process, then the properties of the head noun should be independent of each other. Hence, modification of one property should have no effect on the rest of the noun's properties; if modification were knowledge-rich, we should expect that modification would be propagated through correlated attributes. Their results supported the latter: the effect which an adjective has on a noun is not restricted to

the property that it ostensibly modifies.

Murphy argued that, concept combination were knowledge-free, then it should be a closed operation: no properties not in either of the inputs should be found in the combined concept. So, if a property is atypical of both the head and the modifier, it should not be typical of the combined concept. He found no support, from typicality ratings for properties, for the hypothesis that combination is a closed operation.

Hence, it seems that world-knowledge is crucially implicated in the generation of senses during the processing of words and phrases. The explanations of both Medin & Shoben and Murphy lean heavily on the role of sense relations (as the meaning connections in a lay theory). Such relations support inferences which elaborate on unspecific initial interpretations. They may, however, be insufficiently flexible to allow for the degree of context-sensitivity noted. This is because of the assumption that theories are theories of the world, reflecting the objective aspects of the world; so transformations of concepts must be legitimised by parallel transformations in the world. For example, crossing ontological boundaries might be problematic. I take this up in more detail in section II.2.7.. However, one way of circumventing this difficulty - of allowing for conceptual flexibility without requiring that the world be equally flexible - is by allowing for a different kind of knowledge, one that we might label "situational knowledge".

The kind of knowledge I have in mind links with instantiation. The idea is that, by fleshing out the discourse scenario in more detail, we might be able to support further inferences about the entity in question, and thus ascribe more properties to the sense. So a crucial aspect of the elaboration of the sense will be the kind of scenario with which an agent is presented, either physically or in the discourse. Consider the statement

He hit the man with a fake gun

Here, *fake gun* is fully understandable, but not fully specified. For example, we do not know precisely what kind of a fake it is. Now, if the statement had been preceded by

James Bond crept into the hotel room

then we might instantiate the combination *fake gun* with a replica of a gun, and this could be used as the basis for the inheritance of properties. Now consider the statement

The schoolgirl sketched the stone lion

Again, it seems that we do not know precisely how the stone object looks like a lion - whether it is a household ornament, a sculpture, or even a bizarre rock-formation. We do not know precisely which properties to ascribe to the NP sense. But in a context such as

In the park there were two statues

we might be able to instantiate *stone lion* to a statue of a lion.

Notice what is happening in both cases. They contrast with the *fish/shark* and *bird/chicken* examples of Halff et al, since the latter involve an instantiation being a subordinate of the noun type given. In the current examples, the instantiation is an entity which is of a type entirely distinct from the head noun entity (indeed, in the second case, there is a crossing of the ontological "animate/inanimate" boundary). So instantiation may support a shift in the way in which the entity is construed: a shift across types and even across domains. The superordinate-subordinate instantiation depends upon explicit semantic attachment, whereas the other is based upon implicit attachment.

The argument I am making is that, in processing a concept combination, an agent constructs a scenario which represents the discourse situation. In doing so she may construct a situation which resembles the discourse one, but in which a more precise instantiation of the entities and actions involved are represented. This latter can then be used as the basis for the inheritance of properties by the expressions in the discourse (the concepts). This might be done either by a direct "reading off" of properties from the instantiation to the sense, or by switching the placing of the concept in the relevant theories. The discourse situation facilitates sense generation in at least two ways. Firstly, it will provide a set of constraints on the selection of an instantiation for the combined concept: the instantiation must be able to fulfill the role and effects of the entity in the situation. Secondly, the situation will set up thematic viewpoints on the entity, indicating which kinds of information are to be considered.

We will need a little more structure on our conception of a "scenario" or "situation", in order to be able to be more precise in later sections. This will be especially important for the notion of a perspective, which I will develop later. For this purpose only, I will borrow some notions from Situation Theory, as developed by Barwise (Barwise 1989), although I will use slightly different terminology.

In understanding a discourse, we can conceive of the agent as relating the words she hears to a situation: this situation might be the current situation "out there" (when the interlocutors are talking about the immediate environment), or it might be some described situation. The most basic distinction which Barwise makes is between different Perspectives on a situation; I shall label my (roughly equivalent) notion "Orientations", and reserve the term "perspective" for another meaning which I have been developing (Franks, Myers and McGlashan 1988, Franks, Myers and McGlashan 1988, Myers, Franks and Braisby

1989, Peacocke 1986). A single physical situation can support innumerable Orientations: these may be different in a physical or psychological sense. To take Barwise's example, consider two people sitting across a table from each other; they share an environment, a situation, and probably individuate the same range of objects in the situation. But they have radically different Orientations on what is before them. From one's Orientation, the salt is left of the pepper, and so she tells the other that the salt is left of the pepper. Clearly, from the other's Orientation, this is not the case. So the understanding of concepts must take place from the particular Orientation of the agent: this will not only limit the range of concepts and combinations (as descriptions of entities and predications) which can be understood; it will also determine whether or not a concept or combination is being applied accurately. It is only under an Orientation that agents use and understand language. This reflects the anti-realist concern to wed semantics to the temporal, physical and intellectual context in which language is used, as noted in Chapter I. Within any one Orientation, different agents will have different Viewpoints. A Viewpoint is the range of entities and relations, a very small proper subset of those possible, which the agent is taking into account on any occasion (Barwise's "focus situation"). It is therefore variable from one moment to the next. Different agents bring different world knowledge, different memories of episodes, different informational requirements, expectations and susceptibilities to a particular situation. As a result, they extract different information from that situation. So the Viewpoint has primacy in any understanding of a discourse: it concerns the kinds of information which an agent can extract from the situation; it encompasses all of the resources that an agent brings to bear on a discourse. We can partition a situation (as cognised through a the Viewpoint on the Orientation on the situation) along two lines. Each way of partitioning is a way of providing a more specific focus within our Viewpoint. One is in terms of the particular categorisation of the entities in the situation; this is the central question of this thesis, and the crucial aspect of this is the Perspective or Perspectives being adopted on the entities. I shall pick this up again in a moment. The second partitioning is of a thematic nature. That is, the current theme or subject-matter of the discourse will influence which aspects of the entities in the situation are taken into account. So properties of the entities individuated in the Viewpoint will be thematically pertinent or not: those which are might be said to define the current Thematic Dimension. Now, the entities in the Viewpoint under the particular Thematic Dimension can be categorised in many different ways, through the aegis of Perspectives. A Perspective is the categorisation of an entity as a certain type on the basis of a specific range of properties. Now, we must be careful to distinguish between the instantiation of a concept as a particular kind of entity, and the use of a Perspective to categorise the entity which a concept represents: the two may not necessarily coincide. The instantiation of an entity for a concept, under the particular Thematic Dimension, will

provide content for the concept. If an agent were then to attempt to categorise an entity of the concept type, then the appropriateness of that categorisation must be evaluated with respect to the information in the concept. So the instantiation provides an initial categorisation of the entity, which then supports inferences to the concept content. The latter then is the benchmark against which other Perspectives must be evaluated. So although the instantiation-categorisation can be seen to define a perspective through associating a description (of sortal type plus a range of categorising properties) with an entity, other Perspectives are possible, so long as they do not contradict the content ascribed to the concept. The important point is the degree of flexibility provided for: within any one described situation, there may be more than one Orientation; within any Orientation, the number of Viewpoints will be variable; within a thematically partitioned Viewpoint, there may be many Perspectives. And any categorisation perspective may fail to be appropriate.

The argument is that we utilise two aspects of world knowledge: the "encyclopaedic" and the "situational". The manner of their interaction will be discussed when the model of concept combination is outlined later.

I should like also to mention at least one further type of knowledge which I think has received little or no attention in the cognitive science literature. This concerns what we might call our "theory of sense". This might be said to have two relevant aspects. The first concerns our theory of error (Grayling, 1985: 2): we use words to categorise the world, to claim that things are thus and so; but such claims are defeasible. That such defeasibility does not render our claims open to sceptical doubts is a result of our theory of error: in order to possess the notion of being mistaken in our categorisations, we need to know what counts as being mistaken. This means that we need to know exactly where and when to doubt our claims, and how they might be put right. This issue was addressed in Chapter I, and will be picked up again in more detail in section IV.2.1. The second aspect is a "theory of language use", which must inevitably cover a vast range of knowledge - from the principles of co-operativeness and conversational implicature (Grice 1975), through pragmatics factors such as presupposition, the given/new distinction, and so forth (Levinson 1983), to the particular "language game" in which we are operating. The crucial point is that such knowledge leads us away from the assumption that semantic content must needs be veridical: that is, the way we use language is replete with instances where the content ascribed to a concept differs widely from the kinds of properties ordinarily ascribed to the entities labelled though the word. In our discourses and texts, we perform transformations and operations on objects, and have objects perform various actions, which could not, through historical, conventional or medical constraint, have any parallel in the behaviours of the real

objects in the world. We can construct local or ad hoc theories or explanations which "suspend" our ordinary theoretical and explanatory principles. And metaphor, metonymy and analogy, which is central to our language use violate, although in regular ways (Lakoff 1987), the usual real-world behaviours and connections between objects. And this means that we know that language can be used in ways that violate the usual property-ascriptions of the concepts we use: that such ascriptions are flexible and defeasible. We might hypothesise that a possible motivation for allowing such defeasibility is preserving the Principle of Co-operation, noted by Grice. It will be my claim in a later chapter that the principles of defeasibility that underpin our theory of error (in the categorisation of entities), are the very ones that operate in our theory of language use (in the ascription of properties to concepts in sense generation). Of course, if we allow that the theory of sense can take such priority over the theory of the world, then this means that the resultant concepts will have a different relation to the world; that categorisation must be more flexible than is usually countenanced. This is the motivation for the perspectival-relativity of categorisation, into which I delve later.

II.2.6. On the Partitioning of Content

In this section, I discuss partitioning the content of a concept, reflecting the nature of the Manifestation Constraint as one on the individuation of content. I have referred to the "properties" of a concept. In doing so, I am adopting the convenient assumption that we can consider the concepts as at least partly comprising of sub-lexical units. We should be a little clearer about what is meant by a "property". As I am using it, the term is agnostic as between binary and dimensional attributes, and between perceptual and non-perceptual attributes. It is best construed along the usual featural lines as a <feature, value> pair. Further, I do not assume that properties are necessarily any more "simple" than the concepts in which they inhere: indeed, properties themselves may be concepts which play a role in another concept. Where a property-analysis does not result in a reduction of level of complexity, a local holism occurs. Finally, I do not assume that all properties are expressible in language: they may manifest the ineffability suggested in I.3.1.

II.2.6.1. Conceptual Core vs. Identification Procedure:

Miller & Johnson-Laird (1976) advocated a distinction between the "core" and the "identification procedure" associated with a concept. The core is an inchoate lay theory, comprising essential information associated with a domain. Miller & Johnson-Laird are not entirely clear regarding what a core might contain, although they do indicate that its content will be somewhat abstract, not directly concerned with the identification of objects. The

core structures the sub-domains in which identification procedures inhere. The latter are perceptual-functional properties which constitute categorisation-conditions for an entity. So the core determines the nature and functioning of the identification procedure.

A crucial issue concerns the precise relation between core and identification procedure. The argument is that core properties are never utilised in identification. And the identification procedure can be revised or modified without altering the core, and vice versa (Miller & Johnson-Laird, 1976: 300). However, if identification procedures are perceptual-functional schemas, as well as relations to other concepts, then this blurs the line. For such "relations" are surely the stuff of theory: either the core aspects must enter into the procedures or vice versa. And if this is so, revision of the core demands revision of the identification procedure. This problem results from a deeper issue concerning on exactly what basis the distinction is made. Their initial comments suggest that the distinction is content-driven: that identification procedures are intrinsically perceptual-functional, and cores are abstract. Later, their argument centres more on the functions of the components: so the perceptual is not exclusively the province of the identification procedures. If this is the case, then they are still able to preserve the independence of the components, and the precise contents of each becomes an empirical matter.

More recent accounts have utilised the same distinction (as a distinction between centrality and diagnosticity), and assumed the components' independence.

II.2.6.2. Diagnosticity and Centrality:

The label "diagnosticity" derives from the Modification Model of concept combination of Smith & Osherson (1984). The idea is that, amongst the publicly observable aspects of an object, some will be more critical in the categorisation of that object than others. The more diagnostic a property is, the greater the likelihood that an agent will make an appropriate categorisation of the object when the property presents itself. Properties vary in their diagnosticity: whereas colour and shape might be equally diagnostic for the application of the concept *apple* to an entity, colour will be of higher diagnosticity for *red apple*. They argue that adjectives, as naming properties of an object, can be "negatively diagnostic", "positively diagnostic" or "non-diagnostic". Diagnosticity is thus considered with respect to an adjective or property's role regarding an entity describable by the head noun, and not by the noun phrase.

The question to be considered is whether diagnosticity should be viewed as independent of a notion of "centrality", as espoused by Medin & Shoben (1988). They aver that, if we

hold constant both the diagnosticity and the subjective probability of an entity's having a certain property, we nonetheless find that some properties are more "important" in one concept than in another, even though they are equally "true" of both concepts. Such properties are also those which "accept little change". This centrality is explanation-related: if a theory or ad hoc explanation maintains that there is a strong connection between a property and the sortal type of the entity, then that property is deemed central. Some of Medin Shoben's examples of combinations in which central properties are denied, are: *soft diamond*, *straight boomerang*, *wooden skillet* and *polka-dot stop sign*.

It can be argued that diagnosticity and centrality must be closely intertwined. To see this, consider the Criterial relation (see Chapter I). A C-relation holds between the particular conditions of application of a concept and the applicability of that concept. Consider again the case of my friend's being in pain. As a result of his manifesting the behavioural conditions of application for the concept of *pain*, I might claim that he was indeed in pain. Conditions of application may be divided into Criteria and Symptoms. Whereas Criteria are highly diagnostic of an entity, Symptoms are less so. An example is the tendency to cry when in pain: crying *per se* is not an indicator of pain, so if this were our only evidence for the application of the concept, we should have to qualify the generalisation embodied in the application of the concept. That is, whereas when Criteria are true, we can claim that the particular entity is of a certain type in the general case, an application on the basis of Symptoms will be of a restricted generality. So we might say that my friend is manifesting pain behaviour, in crying, but that any categorisation of his being in pain must be restricted to this range of evidence.

This view, then, illustrates the interdependence of centrality and diagnosticity. Depending upon which of the diagnostic properties obtain, we can make different inferences as regards the truth of an ascription of central properties to the object. If the Criterial (i.e., highly diagnostic) properties are true, then we can (defeasibly) infer the central properties with certainty. If only Symptoms (i.e., properties of low diagnosticity) hold, then we can make either or both of two inferences. We could infer some of the non-central properties of the object with certainty; or we could infer the central properties with less certainty.

An examination of Medin & Shoben's account of centrality appears to indicate that they should require such interdependence. They argue that what is central may be dependent upon a particular local explanation. And clearly, one aspect of an explanation of the presence of certain properties is to link them with the presence of a particular type of entity. So, the application of a concept, and the inference to conceptual content, are closely linked,

with the latter helping to explain the former. This means that, given the possibility of different perspectives on an entity - supporting categorisation judgements of greater or lesser restrictedness - we may generate different explanations for the same set of properties, both on the same and different occasions. Once again, then, diagnosticity and centrality are intertwined. In arriving at this conclusion, we have also claimed that diagnostic properties can be divided into properties which are highly and less highly diagnostic.

What, then, of the notion of centrality? Medin & Shoben provided evidence that a property which is equally true of two concepts can be more central in one than the other. However, they do not give any detailed breakdown of centrality. Their stimuli were concept combinations in which a property of the head N is denied by the modifier; they were chosen to reflect the denial of central versus non-central properties. Unfortunately, Medin & Shoben appear to have taken into account only the relative centrality within pairs of combinations: so their account is solely based upon whether being hard is more central to diamonds or knives. They take no account of whether the effect of denying a central property of one N is similar to denying that of another. This makes it difficult to generalise across properties or combinations, and thus to distil a common factor in centrality.

To demonstrate this problem, we will examine their stimuli a little more closely. For example, for a diamond to be soft seems to imply that it cannot perform the usual functions of, or have the usual ontological nature of, a diamond: hence the combination denies some central property of the N. But consider a *triangular record* or a *diamond-shaped yield sign* (both allegedly involving a defeat of central properties): there is nothing in these modifications to prevent NP referents performing the functions of the head N, and thus being categorised as members of that N. Again, there is no reason why green clouds or pink grass should not be clouds or grass. It seems, then, that there are two types of modification here, whilst Medin & Shoben assume there to be only one. There are those combinations which, if true of an entity, would seem to imply that the entity was not, in the general case, a member of the head N: these include

- flying whale
- square basketball
- soft diamond
- waterless ocean
- straight boomerang
- yellow blood.

These modifications involve either a direct or indirect denial of a central function of the head N (for artifacts), or a central ontological essence (for natural kinds). And there are

those combinations where the entity may or may not be a member of the head N, where the information given by the combination cannot support a determinate decision either way. This modification appears to be altering a diagnostic property of the N, in such a way that a NP entity might be either a rather odd N, or not one at all. If we decide that they are the latter, then they will behave in the same way as the first group. This group includes:

- pink grass
- polka-dot yield sign
- triangular record
- striped sun
- two-wheeled truck
- wooden skillet
- diamond-shaped yield sign
- green clouds
- orange coffee

The two groups make up the total of fifteen combinations in which, according to Medin & Shoben, the modifier denies some central property of the head. The distinction between these two groups is borne out by observation of Medin & Shoben's data. All of the combinations in the first group had mean typicality ratings (with respect to their head N's) lower than the lowest mean typicality rating of the second group. So a sign test between the first group and any random selection of the same number of observations from the second group would indicate a significant difference between the two groups. And this would be in the required direction of the second group's being more typical of their respective head N's than the first group. So there is a good prima facie case to suggest that Medin & Shoben's analysis is incomplete in that it makes too few distinctions.

Now, let us look at centrality a little more closely. I will take Central properties as crucial to being of a certain sortal type. Neisser (1987) argues that, in the domain of natural kinds, this involves ontological knowledge concerning either some presumed internal "essence" of an animal, or some ontogenetic history (such as a certain form of birth and rearing). For human artifacts, central properties might be the essential functions which they perform in human beings' interactions with them. The explanatory theory in which such properties are embedded will license the criterial canonical link between the conditions of application and the central conceptual content.

But notice now that we have arrived at a further divergence from Medin & Shoben. Their claim was that the property "curved" was central of boomerangs, that having a restricted range of colours is central to clouds, grass, coffee and the sun. But none of these properties

would, per se, be deemed central on the current view. Rather, they may be properties which support the central properties: diagnostic properties linked to the central ones of function and/or presumed essence. But pink grass may still be grass, and a straight boomerang may still perform the crucial functions of a boomerang.

A second aspect of Medin & Shoben's account is that centrality is closely linked property mutability. That is, the more central a property is, the less mutable - the less subject to change - it is. However, they do not elaborate on this. However, on the current view, Unless the appropriate conditions of application obtain, the ascription of the conceptual content will not be warranted. This amounts to the requirement, already maintained, that central properties are defeasible. I will later argue that such defeat of these properties has important implications for the entity's being categorised as of N type; although it need not deny the categorisation altogether, it will severely restrict its generalisability to other contexts.

II.2.7. Concept Coherence

Myers Franks & Braisby argued that a sense selection approach is incompatible with a theory-driven account of concept coherence. Murphy & Medin (1985) distinguish between internal and external components of concept coherence. Both aspects derive from theories of the world. Internal coherence results from the relations between features for which the theory provides an explanation. For example, the various properties of chairs represented in the concept cohere in virtue of theories concerning the function of a chair. Theories also motivate external relations - as causal and explanatory connections - between concepts. Consider the concept for *cat*. According to Murphy & Medin this is coherent to the extent that it is embedded in our theories of the world: the greater the number of connections between the concept and the rest of our knowledge, the more coherent it is. The connection of *cat* to many different aspects of our knowledge (such as our theories of domestic life, ownership, pleasure and so on) provides it with coherence additional to that provided by our theories regarding *cat* alone, which is the source of the concept's internal coherence. Seemingly, then, this could not allow that complex concepts such as *elephants and lemons* or *stone lions and trout* have the same level of coherence as *cat and dog*. This stems from the fact that such complex concepts have fewer common theories to relate the component concepts.

This view of the theories underlying coherence seems to be intimately connected to metaphysical issues. Accordingly, for Murphy & Medin, concepts result from the way that

theories cut up the world. Consequently, categories that cut across ontological boundaries would not be very coherent. For example, *elephants and hopes* would be considered coherent only if the category were motivated by a theory. If this theory were consistent with one's other theories, the coherence of the concept would be enhanced. In this sense, the view has distinct parallels with Keil's approach. As was commented above, there is evidence to suggest that Keil's picture (and, by extension, Murphy & Medin's) fails to reflect the flexibility of everyday predications.

The claim of Murphy & Medin is that the flexibility of conceptual coherence derives from the flexibility of particular theories. This appears to proffer two mechanisms for allowing coherence to be flexible. Either, in certain contexts, coherence is enhanced by constructing an explanation or theory or, alternatively, the theories that underpin coherence are inherently flexible. It is not clear how these mechanisms for flexibility in coherence would work for examples like the "lion puzzle", noted in II.2.3. In such a case, it seems that the word *lion* can be used quite appropriately to talk of a statue of a lion. However, it is unlikely that many of our theories of lions involve statues: our theories of lions can cut up the world so that we can categorically assert that a statue of a lion is not a lion. However, it seems that there is an essential perspectival aspect to this case. That is, we can adopt a different perspective on a statue of a lion: we may simply view it as a lion. Additionally, both of these perspectives seems equally coherent.

There appear to be two possibilities for Murphy & Medin to account for this flexibility in coherence. One is to invoke their suggestion regarding the flexible nature of theories. Another is to adopt the standpoint of weak sense selection, and thus hold that there are two lexical concepts for *lion*. In what follows I will explore the former possibility and conclude that the latter is their only plausible option. This in turn leads to an inconsistency.

In invoking flexibility in theories, we might envisage three possible mechanisms open to Murphy & Medin. One mechanism that they offer is that of the construction of an explanation in cases such as *elephants and hopes* where, seemingly, ontological boundaries are crossed. In the "lion puzzle" apparently there is a similar crossing of ontological boundaries. Our theories of lions tell us that stone lions are not really lions, though presumably there are some theories of lions that are also theories of stone lions. Perhaps, then, the coherence in this case rests on an ad hoc explanation. However, since our stable theories concerning lions cannot all apply in the case of the stone lion, it must be that the use of *lion* in talking of a stone lion is less coherent than the use of *lion* in talking of a real lion. This follows because coherence in Murphy & Medin's view falls out of the number of

explanatory links which structure the concept. So constructed explanations necessarily lead to less coherent concepts than those stable explanations provided by the knowledge base. This is at odds with our intuitions. Seemingly both uses are, in this case, equally coherent. What we want is to allow that both ad hoc and stable explanations can provide for equal coherence.

A second mechanism to account for flexibility in coherence is to allow that we may have flexible theories. So the use of *lion* to talk of a stone lion can be coherent, in virtue of the fact that the same theories concerning lions can also be theories concerning stone lions. However, this does not accord with the intimate connection that theories have to metaphysical concerns. If our theories concerning lions are also theories concerning stone lions then we have no principled manner of cutting the world such as to differentiate between non-lions (like our stone lion) and real lions. This must be seen to undermine the metaphysical position that Murphy & Medin adopt.

A third possibility that Murphy & Medin may allow for is the option of theory change. This option entails that in order to preserve the coherence of the lion concept when employed to talk of a stone lion, the theories underlying the concept would have to change. This, however, seems problematic. Since they adopt the Quinian position of viewing theories as components of holistic web, a change in one will have ramifications for all others. In this case we would have to change our theories concerning lions so as to allow the possibility of inanimate lions. Changes as radical as this, as they acknowledge, would require a global reorganisation of one's knowledge base. It would also require, under most interpretations, a divorcing of theories from metaphysics.

I have argued that there are good reasons to suppose the kind of flexibility observed in coherence is not to be captured by the notion of flexibility in theories. I now turn to the second possibility noted earlier.

The alternative position that Murphy & Medin might adopt is weak sense selection. That is, the above arguments will not go through if we assume that there are two lexical entries and hence two lexical concepts for *lion*. The arguments already presented would caution against adopting this view: for example, it just is very clearly the case that stone lions and lions are very closely related; and this is something we should want our lexical entries to capture. And it is presumably the case then that since we can relate lions and stone lions we have theories to do just this.

A further issue concerns the possibility that a theory-based view will precipitate a slide into either an inextricability or a constitutive holism. The view is not, however, committed to inextricability: Murphy & Medin's claim is that conceptual knowledge is crucially related to theoretical knowledge, and that concepts gain their coherence through their interconnections. So there may be a practical problem as regards the separation of theoretical and lexical knowledge. However, they propose no criterion for differentiation, nor for restricting the range of interconnections taken into account. And they do allow that their picture has parallels with that of Quine (Quine 1960). In recent developments, there is again no attempt to address this question: their focus changes to what might constitute an adequate form for a lay theory (Medin & Wattenmaker 1986). But that a theory is "adequate" carries no implication for the separability or well-foundedness of processing using that theory. And holistic links are holistic links, whether they involve adequate or inadequate theories. However, they do not appear to view this as a difficulty. But if it is a problem for a theory of semantics *simpliciter*, it is even more troublesome for a theory which claims some connection between semantics and cognition - and perhaps insuperable for a theory of cognition *simpliciter* (Fodor 1987). And it seems that Murphy & Medin are advocating the latter.

So it seems that we would want to accept a view of coherence as theory-driven, with the proviso that we do not thereby allow in any unrestricted hard holism. As commented in Chapter I, the existence of local holisms is not problematic. If we accept the arguments that a global holism is implausible for both a tractable theory of semantics and a processable account of cognition, then what is the alternative? Chapter I argued that this amounts to: how might a molecular semantics be possible? The most important aspect of this is the permissibility of defeasible semantic relations. On a holistic view, for a proper understanding of a concept, all of the relevant connections would need to be traversed, with no guarantee of either separability or well-foundedness. On a molecular view, not all relations need to be traversed, and those that are taken into account may be defeated without rendering inappropriate the application of the concept.

Molecularity is also guaranteed by local theories: that is, provided by ad hoc explanations, situational factors and informational requirements. If Fred were to exclaim, *This is not a lion*, then we can only ascribe to him the perspective we labelled the "real lion" perspective, since only this perspective could make this use of *lion* coherent. If we attribute the "statue of a lion" perspective to Fred, then the use of *lion* would be rendered incoherent. It follows then that the coherence of a sense can only be defined relative to some perspective. And Perspectives may involve local or ad hoc theories.

Let me be clear about the claim being made here. I have argued that we must abandon the assumption that the knowledge which drives coherence is based upon a realist picture of the way the world is, which results in semantic holism. Rather, our theories must allow for flexibility. It is not that we want our theories of the world to alter every time we use a word; but that it is a part of our knowledge of language use that we can override real-world knowledge in favour of an ad hoc theory; this is a suspension, rather than a denial, of world-knowledge.

II.3. Summary: the Manifestation Constraint

In this section, I summarise some of the main points of this chapter. This will issue in an operationalisation of the Manifestation Constraint, which an adequate theory of concept combination must satisfy.

The first point was in favour of an "intensional" account of concepts and their combination: the theory should take account of the content of the concepts. Linked to this was an acceptance that the relation between sense and reference is a complex and interactive one.

The second aspect was the denial of Sense Selection: the three central phenomena noted were that the sense ascribed to a word varies across different contexts, that such ascriptions may be based upon the inheritance of properties from an implicitly attached concept, and that concept combination does not occur after a prior accessing of two whole concepts. Any theory should provide for all three possibilities. An account of sense generation must also allow for both flexibility and context-sensitive specificity.

Some vagaries of reference were then outlined, in which the connection between sense and reference demands a central role for world knowledge. The interaction of the generation of a sense, and a referent, was illustrated by instantiation. Here, an entity is accessed as a referent of a concept whose properties underdetermine that choice; the sense then inherits properties from that referent (or its concept). A crucial possibility is instantiation to an implicitly attached N entity. An account must allow for this ascription of properties to, and definition of the identity-conditions of, a concept with respect to a referent.

The next section indicated that, in the process of sense generation, any of the properties of the concept might be denied, and new properties emerge. Hence, generation, and concept combination in particular, must be potentially nonmonotonic.

It was then argued that lay theoretical knowledge - expressing sense relations - was crucial to the elaboration of a concept or combination. Instantiation - as knowledge of the situation and referent - is also crucial to this elaboration. The latter is also a source for knowledge modification and revision, through implicit attachment, thus providing for flexibility in property ascription. In the generation of a sense, then, these two types of knowledge interact. Also relevant is our meta-communicative knowledge, or theory of sense. This allows that an interpretation might deny world-knowledge and even ontological constraints.

A distinction was also advanced between the conditions of application of a concept and the content ascribed to that concept. In the former, the distinction between diagnostic and non-diagnostic properties was adapted to incorporate the distinction between Criteria and Symptoms. The distinction between central and non-central properties was also modified, and it was argued that ascription of central properties must be defeasible. In the individuation of content, a coarser individuation is achieved by focussing on the central properties alone, and a finer discrimination arises from considering non-central ones also. It was claimed that diagnosticity and centrality are interdependent, and that both must be relativised to a particular use of a concept.

The final issue concerns concept coherence. It was argued that a sense selection view is inimicable to a theory-driven picture of coherence. A modified version of the theory-driven picture was adopted, in which the theories utilised were shorn of any ontological presumption, and where ad hoc theories were able to motivate an application of a concept as coherent as those supported by the more stable lay theories of the world. A further requirement is the specification of a means of circumventing a global holism. The argument will be that the processes which underpin the extension of lexical concepts are the very ones which support coherence.

The Manifestation Constraint, then, is operationalised as a cognitive constraint on an account of concept combination in general, and the problem classes in particular. In the following chapters, I develop such an account, and put forward a "mechanism" for combination. We must, however, be clear concerning the underlying cognitive/representational claim that is being made here. My approach carries a commitment only to the cognitive representation (or rather, discriminability within a representation) of different types of content or properties. Although I argue that certain properties are "ascribed" to "lexical concepts", this is simply a locution for referring to the regular contribution that general knowledge makes to the interpretation of a word or phrase. I use the terminology of concepts and senses only

as a convenient way of describing surface manifestation phenomena in a perspicuous manner. The mechanisms I suggest are then only plausible functional means of effecting a transition from this individuation of content types to the observed cognitive behaviour. As functional descriptions, they are to be located at Marr's "computational" rather than "algorithmic" level (Marr 1977). They present means of moving from the use of a term (its reception or production), and its associated content, to an observed categorisation behaviour. This categorisation behaviour gives rise, in Chapter V, to a typology of Privative adjective types, and "scale" of "privativeness". It is thus important to note that such a typology, reflecting what appear to be strong intuitions concerning the categorisation of entities, concerns word-meanings (i.e., the relations between the generated sense and referents), and is strictly independent of the account of concepts, and the proposed mechanism for sense generation. Hence, the two aspects should be evaluated separately. Insofar as the mechanism can produce such discriminations among behaviours, and insofar as those behaviours are intuitively grounded, the structure of the mechanism for sense generation is supported. It is clear, however, that there may be many competing accounts of a mechanism that could adhere to (some form of) the Manifestation Constraint.

II.4. Partiality, Relativity and Sense Generation

The requirement that an adequate theory satisfy the Manifestation Constraint is met by a view of concepts as partial objects. In this section, we will see that it is indeed possible to translate the phenomena into the theoretical terms regarding partiality; however, a mechanism to account for the behaviour of the problem classes within this framework requires additional machinery, some of which will be derived from a consideration of some theories of combination, in Chapter III.

Sense selection assumes that concepts are total objects: they are fixed packets of information with respect to which any property is determinately either true or false. So it is that the identity of our semantic objects must be defined either with respect to concepts/senses, or with respect to referents.

In contrast, on sense generation concepts are akin to Landman's pegs: since properties are ascribed to them during processing, they obtain their identity conditions after such ascription. By definition, this approach is "intensional", since it is concerned with the properties ascribed to a concept. On sense generation, the various senses of a word are extensions of a single lexical concept, where the latter is a partial object. That is, senses are related to lexical concepts via a relation of concept extension. Returning to the example of Rudy at

the ballet, the different senses associated with *delightful* can be ordered in the following way. The usher's sense is only as specific as the lexical concept, and is compatible with any of the possible extensions of the lexical concept. Ron's sense, however, is more specific. It doesn't just indicate that there is something that is delightful, it also indicates what it is that is "delightful", namely the dancer. Thus, for Ron, the information conveyed by *delightful* contains the information gleaned by the usher. Similarly, for Reg and Rudy with respect to the usher. So the ordering is one of informational containment. However, if we compare the information conveyed by *delightful* from Ron's perspective with that conveyed from Reg's, we cannot say that either is more informative. The relation of informational containment is undefined for these two senses, and it follows that the relation is a partial one. Objects ordered by this relation are partial objects.

We must be clear about what is meant here by a "partial object". I do not mean that a lexical concept is composed of "unvalued parameters" This would rule out a possibility I wish to preserve: that the sense of a term on any occasion might just correspond to the lexical concept. For example, the usher's understanding does not involve any extension at all; it seems that such senses will be pertinent where we are adopting a relatively "perspective-free" approach to an entity, for example in simple type-type categorisation (e.g., *a lion is a mammal*). Which means that the lexical concept, though a partial object, is in no sense "incomplete"; it is partial only with respect to the various possible extensions, and referents.

Lexical concepts represent a coarse level of individuation of content, based upon the central properties ascribed to a concept; finer discriminations will be effected by concept extensions involving non-central properties, or different central properties. In this way, flexibility is manifest in different directions of extension of a single lexical concept, and specificity as extensions of different degrees of delicacy.

How does the process of concept extension operate? Taking lexical concepts as partial objects seems to provide us with too many degrees of freedom. So, for instance, we require a constraint to prevent certain co-occurrences of features. A theory-based approach to the elaboration of concepts (manifesting sense-relations) requires that ascriptions can be valid only if licensed by a theory. Again, the process of extending a lexical concept to form a sense is then the very process by which that sense becomes coherent. Knowledge of referent-relations also has a crucial role. Landman has argued that objects partially ordered by the relation of information containment can be viewed as approximations to a referent. An object which is less partial than another is a better approximation to the referent in

question. So the usher's representation of the ballet-dancer is a poorer approximation to the ballet-dancer than is Rudy's. And situational requirements may motivate selection of a referent which is not explicitly attached to the construction. Hence, flexibility is allowed for. A further important aspect of this concerns the definition of identity: the identity of our semantic objects (and hence the individuation of contents) must be given through a specification not only of the property ascriptions, but also the referent in question. That is, identity is defined through relating a description to a referent, where the description (or sense) might "grow into" a fuller description of the referent after the accumulation of information. And this is the same as adopting a perspective on an entity - since this latter amounts to associating a description with an entity. So identity becomes dynamic and flexible, involving not only the content ascriptions but also their relation to a referent. In this way, we avoid the two problems associated with sense selection/total objects, noted above (see IV.6 for an elaboration of this point).

This picture can allow for nonmonotonicity in a very direct way. The emergence of new properties is the very stuff of concept extension, and, in principle, any property ascription provided by a lexical concept can be defeated. As a consequence, we cannot regard the various entities described by a given word as forming a category that "carves the world at its joints". The important point is that such defeat can only be licensed by theories, ad hoc or otherwise, or by an instantiation that involves a shift of referent type.

Hence it appears that there may be a fruitful way of construing both the phenomena and the problem classes through considering lexical concepts as partial objects. The above comments made liberal use of the notion of a perspective. I will be more precise about this in IV.4. Here, I will consider the underlying notion of the relativity of identity and categorisation.

Geach (Geach 1967, Geach 1967) mooted the thesis of the sortal relativity of identity statements. The following comments exemplify this:

I maintain that it makes no sense to judge whether x and y are 'the same' or whether x remains 'the same' unless we add or understand some general term - the same F.
(Geach, 1968: 39).

I am arguing for the thesis that identity is relative. When one says 'x is identical with y', this, I hold, is an incomplete expression; it is short for 'x is the same A as y', where A represents some count noun understood from the context of utterance.
(Geach, 1967 :3).

The count noun is a "covering sortal" for the two entities involved. And where the covering

sortal changes, the validity of the identity statement may alter. For example, in order to evaluate the identity statement, *The Cam is the same as the Rhee*, we need to make explicit the covering sortal. If the sortal is *river*, the statement is warranted; if it is *water*, then it is unwarranted. Geach claims that the particular sortal is usually implicit in, though evident from, the discourse context. Identity statements should then take the form:

I: X is the same as Y qua N

Now, we can view the covering sortal as concealing a particular description - a range of properties used as the basis for the identity judgement. It is a criterion against which the entities are compared, incorporating those properties of the covering sortal that both entities might have. On the current view, we will say that the sortal provides a criterion of indiscernability: if the two entities are indiscernable with respect to this criterion, then the identity statement will be deemed true. So the criterion changes when the sortal changes.

We can now look at the relativity of categorisation. To see this, consider the possibility of making an identity statement with a covering sortal, but making it on the basis of a proper subset of the properties of the sortal. For example, we might want to take into account only some of the properties of *river*: excluding, for example, the precise geographical location, the depth of the water and so forth. The Cam and Rhee might be the same river, but they each exist at different locations along the same course, with correspondingly different depths of water. Although the location and depth of a river could be part of the criterion of indiscernability in most cases (since they form part of the content usually ascribed to *river*), in this case they might not be used. So the expression of an identity statement becomes:

I': X is the same as Y qua N, with respect to C

where N is the covering sortal and C is the effective criterion of indiscernability.

Now, when we categorise an entity, we are not concerned with whether that entity is indiscernable from the categorising type; rather, we are interested in whether the type is an indiscernable approximation to the entity. That is, if a type is an indiscernable approximation to an entity, then all of the properties which the type has, the entity has, and the latter will have additional discriminating properties (see I.3.4). This is exactly what happens in type-type and type-token categorisation: we judge that the central properties of the categorising type are embedded in those of the categorised type/token. Hence, for affirmative categorisations, we have the schema:

AC: X is a Y with respect to C

where we read "is a Y" as "is a member of the category of Y". Again, then, categorisation can be based upon a proper subset of the range of properties ordinarily ascribed to Y. For

example, although a tomato is a fruit, it will not be categorised as one on the basis of all of the properties usually ascribed to *fruit* (such as being a dessert). The picture is complicated somewhat by Privatives. In this case, we have a categorising type which is an indiscernable approximation to an entity manifest through another type: that is, the head N type is an indiscernable approximation to the NP entity only as a result of the NP entity being also of another implicitly attached type, NI. For example, we can say that a fake gun is a gun, since the latter is an indiscernable approximation of the former only with respect to a criterion which is manifest in the fake gun as a replica or a toy. Hence, we can state a schema for Privative categorisations:

PC: X is a Y with respect to C, qua N

Here, N is the implicitly attached covering sortal, NI. But note that Privatives appear to be able to support categorisations satisfying both AC and PC:

AC: The fake gun is a replica with respect to its central properties.

PC: The fake gun is a gun with respect to its appearance, qua a replica.

These schemas indicate the form that perspectives on an entity take, since they relate descriptions (for AC, the criterion of indiscernability; for PC, the criterion plus the covering sortal) to entities. There are, then, two degrees of freedom in such descriptions: the covering sortal (which, for AC is identical to the explicit categorisation sortal; and for PC it is an implicitly attached noun type), and the criterion of indiscernability.

II.5. Summary

In this chapter, I have introduced the particular problem classes to be treated of, and commented on a broad range of cognitive phenomena which should inform an anti-realist view of those classes. My view of these phenomena can be seen as an operationalisation of the Manifestation Constraint, noted in II.3. The crucial facets were the flexibility, and possibilities for context-sensitive specificity, of generated senses. And a sharp distinction was made between the semantic contribution of a lexical entry (the lexical concept), and the sense generated on any particular occasion of use. Section II.4 presented a particular interpretation of the Manifestation Constraint phenomena: in terms of concepts as partial objects and senses as defeasible contextually-determined extensions. Extensions may diverge widely from lexical concepts (thus supporting flexibility), and may add more specific content. A further crucial aspect of the Manifestation Constraint was the idea of a "perspective", whereby categorisations warranted from one perspective may be unwarranted from another. I argued that this could be viewed as an elaboration of Geach's "sortal relativity of identity" thesis, and that perspectives may have two components - a "covering sortal" and an applicable "criterion of indiscernability". This then allowed the specification of two different

schemas for the representation of categorisation judgements - one for affirmative (including attributive) categorisation, and one for Privative categorisation. •

Having provided some detail for the Manifestation Constraint, we are now in a position to discuss various theories of concept combination/adjectival modification in its light; their evaluation will provide a further range of ideas for the current view of concepts and word-meaning.

CHAPTER III

Some Views of Concept Combination

In this chapter, I sketch some accounts of concept combination, and ascertain their ability to satisfy the Manifestation Constraint. I consider both psychological accounts of concept combination, and linguistic-semantic accounts of adjectival modification. I will approach the discussion through two orthogonal distinctions. Sections III.2 and III.3 are concerned with a distinction between theories that address "mediating relations" between concepts, and those that assume a more direct "combination" of concepts. Sections III.4 and III.5 are concerned with views of different "levels" of modification, contrasting accounts in which adjectives may modify a noun at different "levels" depending upon context, with those which assume that modification behaves consistently across contexts.

III.1. Introduction

The argument of III.2 and III.3 will be that Mediating Relations theories are empirically and conceptually flawed, and that we must adopt some variant of a Combinations theory. Within Combinations theories, I argue against set-theoretic views, and in favour of an "intensional" view. The latter is of a cognitive-psychological cast, although I argue that such a view must incorporate insights from a formal-semantic approach. The extant intensional views, then, are insufficiently flexible to account for the problem classes according to the Manifestation Constraint. In III.4 and III.5, methods of overcoming this deficit are discussed, the argument being that a viable approach must allow for multi-level and multi-type modifications.

The discussion, then, results in a more precise picture of the phenomena considered in Chapter II: through a consideration of possible mechanisms, the Manifestation Constraint interpretation of those phenomena is sharpened, and the groundwork laid for the current view to be developed in Chapter IV.

In order to discuss the approaches to combination, we must first be a little more precise concerning the problem classes viewed under the Manifestation Constraint. Firstly, consider Attributives. We can note three possible readings of an attributive combination like:

Rudy is an attractive ballet-dancer

two of which might be seen to be "subsecutive", and one of which could not. The two

subjective readings (in which Rudy, in addition to being attractive in some way, is also a ballet-dancer) are paraphrasable thus:

Rudy is a ballet-dancer and is attractive for a ballet-dancer

Rudy is a ballet-dancer and is attractive in some other way

Note that they are *not* intersective, since "attractive" is not an independent predicate (see II.1). This is exemplified in the invalidity of the following inference:

Rudy is an attractive ballet-dancer

Rudy is a chef

therefore, Rudy is an attractive chef

However, the third reading of the combination satisfies neither a subjective nor an intersective constraint. This reading may be paraphrased:

Rudy is attractive at ballet-dancing

This reading arises from the interaction of an attributive A with a particular kind of N, a member of the class of *nomina agenta*. As Bartsch (1986) comments, such N's (deriving from verbs and denoting an occupation) allow for two distinct meanings. The first regards entities that perform the action described by the verb, and another regards entities that have a profession in which the action is performed professionally. Only in the latter case would we want a subjective analysis. In the former case, in which the A modifies the action of the origin verb, a subjective analysis may or may not be possible. That is, from the paraphrase

Rudy dance ballet attractively

we have no information regarding whether Rudy is or is not a professional ballet-dancer. In Chapter V, I will argue that this particular type of reading may or may not turn out to be affirmative; if it does, then subsection might be appropriate.

This is linked to Privatives. In II.2.6, I distinguished between combinations in which the central properties of the head N were denied (e.g., *stone lion*), and those in which they were cast into doubt by the denial of diagnostic properties (e.g., *polka-dot fire hydrant*). These are, I claim, of the same types, *mutatis mutandis*, as *fake gun* and *alleged intruder*, respectively. There is definitely only a highly limited sense in which a *fake gun* can be categorised as a gun; but, depending upon further information, an *alleged intruder* may be an intruder in either a limited or a general sense. That is, *alleged intruder* may emerge either affirmative or privative. I label these types Negating and Equivocating Privatives, respectively, and provide different treatments in Chapter V. The third reading of an Attributive would then be as an Equivocating Privative. These comments challenge the standard assumptions regarding Privatives (Kamp 1975, Hoepelman 1983): that there is *no* sense in which a Privative NP entity is a member of its head N type, and that this behaviour is a

result of a restricted group of A's. As Kamp comments,

A privative adjective A is one which, when combined with a noun phrase N produces a complex noun phrase AN that is satisfied only by things which do not satisfy N. If A is a privative adjective then each sentence *No AN is an N* will be a logical truth.

(Kamp, 1975: 125).

Although Kamp does allow that a particular privative A may have non-privative results on occasion (e.g., *fake counterfeit*), he does not allow the converse: that particular affirmative A's might be privative in some context of use.

We can allow for the categorisation of a Privative NP entity as a member of its head N (*pace* Kamp), as a result of the Manifestation Constraint-motivated assumption of the perspectival-relativity of categorisation judgements. And the interactive nature of concept combination provides for different effects of a single modifier on different N concepts, even to the extent of a change from affirmative to privative. *Stone lion* and its ilk, I label "Functional Privatives". I have noted some distinctions between types of Privatives here so as to furnish a context for the ensuing discussion. I will discuss and motivate these distinctions further in Chapter V.

III.2. Mediating Relations Theories

Mediating relations have been utilized to solve some fundamental problems in the combination of concepts, and in lexical combination generally. The problematic is compositionality: how can we derive the meaning of a noun phrase from the constituent concepts, coupled with the assumption that word meanings are stable? For example, how does the contribution of *morning* differ across different contexts, such as

morning flight

morning coat?

Compositionally, we should aver that the relation between *morning* and the noun, rather than the meaning of *morning*, differs. That is, the way that *morning* modifies *flight* differs from the way that it modifies *coat*. The specification of this relation is the goal of Mediating Relations Theories, as Warren (Warren 1984) comments:

there exists between a modifier and its nominal head an opaque relation which ... the interpreter must ... decipher in order to understand the construction in question. This relation is not invariable; nor is it endlessly variable, but there is a set of regularly employed relations.

(Warren, 1984: 15)

III.2.1. Some Views of Mediating Relations

In outlining Levi's approach, I will note all of the MR's she postulates. For the other approaches, I will omit such a list; my comments will largely be addressed to MR views on the whole, rather than to specific MR contents.

III.2.1.1. Levi (Levi 1978):

Levi argued that complex nominals (N + N compounds) should be viewed as having the same source as non-predicative adjectives. The latter are derived from the former by an optional adjectivalization rule, transforming N + N compounds to A + N constructions. This would account for the apparent synonymy between pairs such as *atom bomb/atomic bomb* or *hand signal/manual signal*.

Levi aimed to specify the rules deriving the N + N compounds and the non-predicative A + N combinations from a common semantic source. Such a source contains the relevant MR's, which are embodied in the following nine predicates:

CAUSE (e.g., tear gas: gas that CAUSES tears)

HAVE (e.g., picture book: book that HAS pictures)

MAKE (e.g., honeybee: bee that MAKES honey)

USE (e.g., steam iron: iron that USES steam)

BE (e.g., soldier ant: ant that IS a soldier)

IN (e.g., field mouse: mouse IN a field)

FOR (e.g., horsedocter: doctor FOR a horse)

FROM (e.g., olive oil: oil FROM olives)

ABOUT (e.g., tax law: law ABOUT taxes)

The expression of these predicates in terms of MR's leads to twelve different "lexical rules", incorporating all of the above, plus three more involving CAUSE, HAVE and MAKE, in which the "direction" of modification is reversed:

CAUSE' (e.g., drug deaths: deaths CAUSED by drugs)

HAVE' (e.g., government land: land that government HAS)

MAKE' (e.g., snowball: snow which MAKES a ball)

The details of Levi's transformational account need not detain us. The crucial aspect is that the twelve MR's are held to express the modification relations between the modifier and the head N. However, since Levi did not aim to provide a psychological account, she did not attempt to specify conditions under which a combination would use particular MR's. Specifying such conditions has been one goal of other theorists, however.

III.2.1.2. Aarts & Calbert (Aarts and Calbert 1979):

Aarts & Calbert hold that most A + N combinations are essentially predicate structures. For example, *happy man* conceals the predicate structure, "man (who) experiences happiness". So any account of the meaning of such a combination must elucidate:

- (i). The A's predicative force (in this case, "experience");
- (ii). The A's non-predicational - usually referential - content ("happiness", here).

The predicative force expresses the MR between the A and the N. This is not invariable, and may vary depending upon the type of A (for example, *stony* will not occur with the relator EXPERIENCE), and upon the type of N. They postulate thirteen possible MR's: ranging from HAVE (e.g., *clever man*) to UNDERGO (e.g., *felled tree*), and EXPERIENCE (e.g., *sad person*). For example, the conditions on the selection of the MR EXPERIENCE will be: the referential content of the A involves either [+ emotion] or [+ corporeal condition]; and the modified N refers to some animate being.

In addition to relators, Aarts & Calbert postulate negative operators and quantifying elements. The former are implicated in, for example, *blind* (e.g., *blind man*: "man NOT having the ability to see"). Degree A's contain quantifying elements: for example, *hot/cold*. Hence, *hot water* is paraphrased by, "water having high degree of heat".

By specifying such occurrence conditions for MR's, the aim is to give an account of the semantics of non-metaphorical combinations. A metaphorical interpretation is then parasitic on the prior failure of a non-metaphorical reading.

III.2.1.3. Warren (1984):

Warren analyses the semantic contribution of a prenominal modifier into several components. The first is the referential content of the A (or N); the second is the MR between the A and the head N; the third is the function of the modifier in the combination, and the fourth is an optional "classifying component" in the A.

Warren adds an allocation of thematic roles to the modifier and the head in a combination. The scheme is as follows. A combination contains an implicit MR between the referential contents of the A and the N; this MR determines an allocation of thematic roles to the A and N. The allocation of roles constrains the possible function of the modifier; if the roles permit the modifier to have a "classifying" function, then further aspects of the A's referential content, supporting classification, will be ascribed to the NP. As with the other views, the MR is "part of" the meaning of the A on the particular occasion. This then means that any context-sensitivity in A or N derives from variation in either the A's referential content, or the choice of MR. This is because Warren assumes that an A has a "nucleus

meaning", which will be elaborated differently according to context. This contrasts with the head N meaning, which is invariant across contexts.

Warren postulates a set of twelve MR's (labelled "link" relations), which come under five different notional headings. For example, the notional heading "constituted" covers MR's CONSTITUTED BY and CONSTITUTING; the former determines a Source/Result, and the latter a Result/Source role allocation, and the former has a classifying function (e.g., *civic group*), whilst the latter has a Descriptive function (e.g., *nucleus group*). The other notional headings are "Possession", "Location", "Causation" and "Comparison".

Warren envisages three kinds of modifier function: the first is the **descriptive** or **non-defining** function, in which the A serves to add some information concerning the referent of the head. They are non-classifying in that they are non-restrictive:

warm shelter ("shelter EFFECTING warmth") (Result/Causer)

The second type is the **classifying** or defining function of the modifier. This has two sub-types: one in which the referent is *directly specified*, and one in which the modifier serves as a *complement* to the referent of the head. Consider the latter case:

metal handle ("handle CONSTITUTED BY metal") (Source/Result)

In contrast, specifying classifiers indicate directly what kind of referent is under consideration:

Beckett's novels ("novels CAUSED BY Beckett") (Causer/Result)

Only if the modifier is allocated the role of "Part", "Result", "Comparant" or "Norm" may it have a non-classifying function. The classifying function requires that the interpretation of the NP contain an aspect of the A's referential content that supports the classification. Consider

stony path ("path HAVING stones on it") (Part/Whole)

This paraphrase omits the sense that a stony path has more stones than a normal path, even though it does indicate the nature of the MR. Adding a classifying component, Warren claims, suffices to specify the classification adequately.

III.2.1.4. Cohen & Murphy (1984):

Cohen & Murphy's position differs from the previous ones in that it is motivated by cognitive psychological considerations. Cohen & Murphy eschew the possibility of a finite list of MR's. They argue that an understanding of the implicit MR is necessary for syn-categorematic A's, the so-called "absolute" A's, and perhaps even more so for "contradictory concepts", such as

singer who is not a singer

fish that is a mammal (i.e., whale)

robot cat

Their argument is that there no way in which contradictory combinations could be understood in a strictly extensional account, since they should have null extensions. They claim that these cases involve substituting properties which are atypical of the head N for typical ones. The particular properties to be modified or substituted are given by the MR. Cohen & Murphy define a **concept-forming operator**, |, which indicates how the MR role-modification is to work:

If X and Y are concepts, and if R is a role on X that Y fits,

then X | Y is a complex concept specializing X with modified value R for Y.

They argue that the interlocutors will generate an MR relevant to the current context. In their model, MR's are reduced to a mode of modification of the head by the modifier. That is, they argue that N meanings may be represented by a list of "roles", which can accept different specializations or values. So the MR is simply that which determines an appropriate role modification. For example, a complex concept *engine repair* can be formed as a specialization of *repair* by modifying a role taking *engine* as its value (for example, the role of "object of repair"). More complex combinations, such as *stone lion* or *virgin birth*, supply atypical values; for example, we might say that the "material" role of *lion* is usually filled by anything but the inanimate *stone*.

The idea, then, is that contextual demands activate roles according to their relevance. A matching process then attempts to modify the role with the other concepts in the NP. When different N roles are activated, different complex concepts are formed from the same lexical combination. Ultimately, then, the MR is viewed as the selection of a particular role of the head N to be modified by the A or N modifier.

III.2.2. What are Mediating Relations?

The MR directs the manner of modification of the N by the modifier. As such, it is unidirectional - specifying only the way in which the modifier affects the N. So the MR is in part the means through which a dependency relation between the head and the modifier operates. This means, then, that the MR determines both which properties of the A and N are relevant to the combination, and how the A properties modify those of the N.

As a result, the MR is the major clue to the sortal type of the combined concept. Depending upon the MR, a referent of NP type may or may not be of the head N type. This is most clearly seen in Warren's notions of the different functions of the modifier, which

derive from the role-allocation, which in turn is determined by the MR. In terms of property inheritance, the MR directs combination so as to account for both the defeat and the emergence of properties.

Where is the MR encoded? Is it part of the meaning of the head, the modifier, both, or neither? In general, the assumption has been that the MR is part of the meaning of the modifier. In analyzing the meaning of the modifier, it will be recalled, we uncover both the "referential content" and the MR. Only Cohen & Murphy deviate from this line: they are not entirely clear about whether the MR should be considered as a role on the N or as the modification of a role on the N. What is crucial is that it is part of the N and not the modifier. The precise "location" of the MR will have ramifications for the way in which combination occurs.

III.2.3. Difficulties with MR Theories

The theories of Warren, Levi and Aarts & Calbert are all predicated on some version of the Sense Selection assumption. If the sense generation view is correct, then the (f)act of combination itself places constraints on the ascription of content to the concepts. In contrast, these MR views aver that we can specify the conditions of application of a particular MR through examining the "referential content" of the head and the modifier. That is, we first access the concepts, and then derive the appropriate MR from the juxtaposition of these concepts. Only then do we combine them, through the MR. This scheme assumes both sense-invariance and post-access combination; as a result, it cannot allow for implicit semantic attachment. So the adherence to sense selection is complete. I take the considerations of Chapter II as a very strong argument against these MR theories.

Another problem for these MR views is the assumption of compositionality. Sense selection is a basic assumption of compositionality. In this connection, the MR is the "mode of combination" of the senses. However, as we have seen, there is an increasing body of evidence to suggest that a compositional account of concept combination will not go through.

If sense generation is accurate, it is unlikely that a finite list of MR's could do justice to its inherent creativity and flexibility. Another problem is raised by Clark (1983): the MR's express an arbitrary level of detail, and so can capture only an arbitrary amount of a combination's sense. I think that this stems from lack of constraint on what might count as an MR: none of the accounts offer principled grounding of their MR's. If the ontological status of the MR's is unclear, so is their epistemology and psychology: they offer no

mechanism for the semantic ascent from MR's to the senses of NP's.

So the views of Levi, Warren and Aarts & Calbert have similar flaws. The question is then whether Cohen & Murphy's account avoids these difficulties. Cohen & Murphy appear to adhere to the assumptions of post-access combination and explicit semantic attachment. This is one reason for thinking that their view may be problematic. Another is the way in which they adopt sense generation. Their assumption appears to be that the concepts associated with terms will be constant across contexts: the particular MR's, however, will not. However, the focus on each concept may vary: that is, it is as a result of a different contextual focus that different roles of the N are activated to a higher degree.

So the MR - as a role on the head N - is generated in context, whilst the concepts for the head and modifier are selected. Two issues suggest themselves. Firstly, since the MR is a role on the head, it must comprise part of the head's sense: this should place a constraint on the range of MR's which are possible. It is not clear that an MR on the linguistic picture could be reduced to an aspect of head N sense, nor that any role-modification is adequate to the types of tasks which MR's are used to effect, especially in Cohen & Murphy's arguments. Secondly, since the constituent concepts are not permitted any broad context-sensitivity, it is not clear that a sense-generation view is necessary. If there is a finite range of concept senses, composed of a denumerable list of role-value pairs, then there should be only a denumerable number of ways in which the head N can be modified by a particular modifier. Combination could then involve a selection from the range of combinatorial possibilities.

Let me go through the first point in a little more detail. On this view, the MR is reduced to a role on the N (or at least the modification of a role on the N), which latter is selected as a whole sense. So the MR must be a part of that sense: we should, in principle, be able to list the range of roles which might be modified to form the MR prior to the modification. The crucial problem is the MR-as-role-modification reduction. An example from Downing (Downing 1977) indicates the difficulty in stark terms:

apple-juice seat

This was used in the context of a friend being asked to sit at the apple-juice seat, meaning "the seat in front of which a glass of apple-juice had been placed". On the linguistic MR views, this would be handled by some specialization of Levi's IN, and parallels for Warren and Aarts & Calbert. It is unclear that the *seat* would have any of these MR's as a component role (as is required by Cohen & Murphy's view). But even if it did, this would be insufficient to convey the sense: *apple-juice* is unlikely as a permissible filler for the

"location" role. The only role which would be apt would be one which does not abstract away from the intended meaning - a role such as "in front of". But would this be part of our usual meaning for *seat*? I think not. Such an argument could equally apply, *mutatis mutandis*, to the linguistic views, in that they reduce the MR's to aspects of the modifier.

The outcome is that positing the MR as part of the sense of the A or the N has two possible results. Either the senses of the words are stretched to bizarre and unacceptable lengths in order to accommodate the MR's, or the MR's must be restricted to acceptable aspects of the words' senses, in which case they are inadequate to the task set them.

The second issue mentioned above follows naturally from the first. If senses are stable across contexts, and if the MR is reduced to an aspect of those senses, then there is both no requirement and no scope for a generational account of the MR. If the MR is part of a selected sense, then it will itself be selected.

There are then two options facing an approach akin to Cohen & Murphy's. If we are to retain the idea that the MR is a part of the sense of the modifier or the head, then we must abandon sense selection. We must allow that word senses can be flexible across contexts, thus providing the scope for them to contain an appropriate MR. By contrast, if we are to retain sense selection (and thus assume that senses are stable across contexts), then we must abandon the belief that MR's are encoded as a part of those senses. The MR would then be generated in the appropriate context, and is independent of both constituents. But sense selection has been discounted; so to preserve a MR view, we must switch to sense generation.

Would sense generation obviate the difficulties for MR views? I don't think it would. A sense generation mechanism of this sort would be massively circular. Consider again what the MR is meant to achieve: it directs combination in such a way as to determine the inheritance of properties from the constituents, and thus to determine the sortal type of the NP. In order to know what kind of MR would be appropriate, we would need to know what conceptual content to generate for the constituents; but in order to know what conceptual content to ascribe to the constituents (and thus to the NP), we would need to know either what the sortal type of the NP is, or what the MR is. And on this view, it is the MR which serves to determine the NP's sortal type. Clearly, then, we cannot use the NP's sortal type to determine the content ascription to the constituents, which then determines the generation of the MR, since we cannot know the NP type without first knowing the MR.

Such a view must therefore accept one of two conclusions. Either the mechanism for generating the MR is viciously circular, so the MR loses its explanatory value. Or we generate the MR on the basis of the sortal type of the NP, in which case the MR is superfluous, since its prime function was to determine just such a sortal type for the NP. In the latter case, the view must allow the agent access to information concerning the situation, including the NP referent's role within it. I argued in the last chapter that such access is both necessary and plausible, and that it occurs through instantiation. But once we grant ourselves such informations, the need for the MR dissolves.

Further problems for MR views stem from their restricted view of concept combination. MR's embody only a limited level of interaction between the two concepts. This has two facets. The first is that MR's embody a unidirectional modification relation. They indicate the effect of the modifier on the head N. They make no allowance for the variation of modifier sense according to the N being modified. Such interaction in modification forms part of the Manifestation Constraint.

The second aspect of the interactive nature of concept combination is the propagation of modification through correlated properties. MR theories take no account of this. For example, Cohen & Murphy claim that *robot cat* is comprehended by substituting the value of *robot* for that of *animal* in *cat*. They do not allow that denying the animacy of *cat* will have other effects (for example, it won't eat fish) nor that other properties of *cat* will be transformed, rather than defeated, by its being a robot. So modification will involve both the defeat and emergence of properties. Such nonmonotonic effects are not countenanced by current MR theories.

Hence, it seems that there are good reasons to doubt both the current theories, and the basic conception, of MR's as a major determinant of concept combination. We have seen that the theories fail to satisfy most of the aspects of the Manifestation Constraint. It also seems that, once the MR view is considered in some detail, it results either in a vicious circularity, or in the superfluity of the MR itself.

III.3. Combinations Theories

In this section, I discuss some approaches to concept combination or adjectival modification that have been advanced in recent years. Only some of the views have been advocated as cognitive theories: in their evaluation, I employ the Manifestation Constraint, and their approach to Privatives and Attributives. In the other cases, evaluation is based solely upon

the treatment of the problem classes. The views divide into set-theoretic and "intensional" accounts.

III.3.1. Set-theoretic Approaches

III.3.1.1. Classical Set Theory:

Hoepelman (1983) advocates an approach to adjectival modification based upon classical set theory. This provides an account of the problem classes, and issues in a typology of A + N combinations, as follows.

- (i). Privative Adjectives: Examples include *fake*, *former*, *alleged*. A fake trout is not in any sense a trout. The condition is that, if the N is a member of set X, then the NP (i.e., the N modified by the privative A) is a member of the relative complement of X with respect to the particular domain.
- (ii). Affirmative Adjectives: This type of adjective encompasses all but the privatives. They pick out a subset of the set of entities denoted by the N, and are represented by set intersection or subsection.
- (iii). Linear Adjectives: These are A's that induce a linear ordering on subsets of the universe of discourse (e.g., *tall*). The applicability of linear A's is mediated by the chosen "comparison class" (Klein 1980): for example, the *hot* in *hot bath* does not designate the same point on a linear measuring scale of heat as it does in *hot furnace*. So one cannot generalise from application to a subordinate to application to its superordinate.
- (iv). Strongly Predicative Adjectives: These behave like independent predicates, so their extensions are not affected by the N's with which they are combined, and they cannot be used in comparative constructions. Generalisation from subordinate to superordinate is possible: a *fourlegged animal* is a *fourlegged being*. They support binary membership conditions, so the universe of discourse is completely partitioned by the application of this type of adjective. Again, standard set intersection is the basic mechanism.
- (v). Weakly Predicative Adjectives: These include such examples as colour terms. This type of adjective occupies a position intermediate between types (iii) and (iv). Although they can be used in the comparative, they do not produce a linear ordering of the domain, and although they do allow a distinction between red and non-red objects, within the domain of red objects it is possible to discern degrees of red objects. However, as with linear adjectives, the question of selection of the appropriate reference class remains open.
- (vi). Strongly Extensional Adjectives: Where N and N' have the same extension, applying the same A will produce the same extension. For example, *fat*. If all and only cobblers are darts players, then the fat cobblers are the fat darts players. All strongly predicative

adjectives are strongly extensional.

(vii). Weakly Extensional Adjectives: These include *American*, since, if every buffalo is an animal, then every American buffalo is an American animal. It cannot include, for example, *skilful* or *tall*.

(viii). One Dimensional Adjectives: For these adjectives, the following obtains: if X is a more A N than Y, then X is a more A object than Y: if X is a bigger trout than Y, then X is a bigger object than Y.

Since Hoepelman did not aim to provide a psychological account, it is inappropriate to criticise his view for cognitive implausibility. Hence, I concentrate on his account of the problem classes. There are three difficulties with his account of Privatives. Firstly, he treats them as homogeneous: the same treatment is provided for *fake* as for *alleged*. This misses a crucial intuitive difference between Negating and Equivocating Privatives.

Secondly, if we consider only the Negating Privatives, is the view more tenable? I think not. The treatment determines that, whatever a fake gun is, it is just not, in any sense, a gun: it is a member of the relative complement of the head N. This omits the crucially perspectival nature of categorisation: that there just is some sense in which a fake gun is a gun, and that this is as real and coherent as that in which it is, say, a model. This stems directly from the adherence to classical identity in classical set theory. And this is linked to the classical view of doubt, as outlined in Chapter I: if an entity falls short of the defining properties, in however slight a way, this is cause for doubt. Treating semantic objects as total objects, then, results in a failure to capture the crucial perspectival aspects of categorisation.

Hoepelman assumes that privative behaviour is a result of adjective type, so it should occur always and only with a restricted set of adjectives. This does not appear to conform with ordinary language use. For example, combinations like *stone lion mirror fake gun*, and *pink grass mirrors alleged intruder*. Such effects are the result of the interaction of the properties of the head and the modifier, which latter can have affirmative effects in other noun contexts.

Hoepelman does not directly address the question of Attributivity, so I attempt my own interpretation. Clearly, Attributives are affirmative (his type (ii)), so we must assume that a basic intersective mechanism will operate: if someone is an attractive ballet-dancer then she is also a ballet-dancer. And there is no other type under which attributives might fall: they are, by definition, non-predicative (so types (iv) and (v) are inappropriate); they are not

necessarily either linear (iii) nor one-dimensional (viii); they cannot be weakly extensional (it is not the case that, if every ballet-dancer is human, then every attractive ballet-dancer is an attractive human), and hence they cannot be strongly extensional. The problem appears to stem from the fact that there is just no scope for a set-theoretic view not to be committed to explicit attachment: the properties necessary for the application of a concept derive from the set of extensions; and since set-theoretic views are extensional, they must adhere to explicit attachment.

I shall now note some reasons why a representational or "intensional" account based upon classical set theory is inadequate regarding the Manifestation Constraint.

Primarily, any set-theoretic account is committed to the exhaustive specification of content. In order to determine the set of entities which comes under a term, we must specify all of the properties which are relevant. This yields the "classical" view of concepts (Bruner, Goodnow and Austin 1956), in which a concept comprises individually necessary and jointly sufficient conditions on category membership. And there is a well-known difficulty in providing such definitions (Rosch 1975, Wittgenstein 1953): real-world objects, as conceived of in our lay theories, just do not seem to support such good definitions. Nor are our experimental methods suited to uncovering exhaustive specifications (Armstrong, Gleitman and Gleitman 1983).

Classical accounts also incorporate a problematic view of combination: they imply both post-access combination and sense selection. These two assumptions are denied by the Manifestation Constraint. A further implication is that combination is compositional, which has also been empirically challenged.

The inherent indefeasibility of classical representations is a further problem. A classical view implies that the relation between a concept and its properties is deductive. And such a view sits uneasily with a naturalised view of doubt. Defeat of any of the defining properties results in the defeat of the application of the concept, as a result of negative undermining. So there is no way to allow for combinations like *stone lion*, since these must render the use of the word *lion* unfounded, and produce either contradiction or anomaly.

A further question is: what are the primitive bases of the concepts (properties), and hence which concepts are built up out of others? Fodor (Fodor 1975) has argued that there cannot be, in principle, a learning theory for primitive concepts (qua properties); hence, they must be innate. So on this view all concepts may be built up from a finite set of such

innate primitives. Conceptual development is confined to the different ways in which the primitive concepts are combined. Fodor (Fodor 1981) has assumed*that, if there are such definitions in the language of thought (roughly, in concepts), there should be correspondingly good definitions in natural language. I have already noted that there appears to be no foundation for the claim that there are such definitions in language use. Those words which cannot be reduced must correspond to conceptual primitives, and so concepts such as *car* or *table* must be innate primitives. Without a rather extravagant evolutionary theory, the classical account thus could not account for many everyday concepts.

There is an important caveat. The argument is that a classical account is inappropriate as a general theory of concepts and their combination. This is not to rule out, for certain restricted (probably technical) domains, precise definitions, excising the innateness assumption. Or a particular theory might guide the extension of a lexical concept so as to generate a definitional sense. Further, that lexical concepts comprise central properties does not make them classical definitions. Centrality does not equal (classical) necessity, and a lexical concept is not a classical definition.

III.3.1.2. Fuzzy Set Theory:

Here, I will comment on the application of fuzzy set theory to the combination of prototypes. A set of combination rules derives from fuzzy set theory (Zadeh, 1965); it should be noted that this is an extensional method, mapping over sets of extensions rather than the properties of the concepts.

Fuzzy set theory replaces the bivalent characteristic functions of classical set theory with ones which assign values anywhere on a continuum between 0 and 1. Osherson & Smith (Osherson and Smith 1981) formalised the approach as regards categorisation. Their model represents the concept for *bird* by a quadruple

$$\langle B, TYP_{bird(x)}, b, d \rangle$$

with class **B** of birds, a characteristic function $TYP_{bird(x)}$ that measures degrees of "birdiness" by mapping the members of **B** into the continuum [0, 1], a prototypical bird **b**, and a distance metric **d** which operates between the members of **B**. A concept is primarily represented by assigning a typicality value $TYP_{bird(x)}$ to each entity **x**. There is no concern with the mental representations of the entities; as Cohen & Murphy (1984) comment, no mechanism is provided for calculating $TYP_{bird(x)}$ for any **x**; nor is there any information concerning the properties of birds or their relations to other concepts. And, crucially, although the quadruple includes the prototypical exemplar **b**, none of the combination rules utilises this exemplar, nor any domain knowledge.

The basic combination rule proposed for Affirmatives is the **min** rule of fuzzy intersection:

$$\text{TYP}_{a \cap b}(x) = \min [\text{TYP}_{a(x)}, \text{TYP}_{b(x)}]$$

If an object has a characteristic value (typicality) of 0.9 for *apple*, and 0.8 for *red*, its value for *red apple* will be the minimum of the two, 0.8. This rule implies:

$$\text{TYP}_{a \cap b}(x) \leq \text{TYP}_{a(x)}, \text{TYP}_{b(x)}$$

So the typicality of an entity in a conjunction is never greater than that in either of the constituents. And this is demonstrably false: for example, Samuel Beckett is a more typical Nobel Prize winning novelist than he is either a Nobel Prize winner or a novelist. Hence, as Osherson & Smith (Osherson & Smith 1982) point out, the "min" rule fails to respect our intuitions concerning concept combinations.

Clearly, a challenge to the **min** rule does not necessarily undermine fuzzy set theory per se: other rules are conceivable. Oden (Oden 1977) advocates a multiplicative rule:

$$\text{TYP}_{a \cap b}(x) = \text{TYP}_{a(x)} \cdot \text{TYP}_{b(x)}$$

However, this rule has precisely the drawback of the **min** rule: an entity cannot have a higher rating as a member of a combined concept than it does for both of the constituents.

Zadeh (Zadeh 1982) has responded to such difficulties by postulating the "normalisation" of combined concepts. This involves taking the entity with the highest typicality rating in the combined concept, and making its rating 1. Normalisation allows an entity to have a higher typicality in a combined concept than in either of the constituents. Zadeh claims that normalisation occurs through focussing on the intersection which the combined concept indicates. Since it is thus parasitic on the particular combination rule used, it still faces the difficulties for the proposed rules which Osherson & Smith note. Although Zadeh comments that the **min** rule applies only to "non-interactive" combinations, it is not clear precisely how we might view any combination as non-interactive. Osherson & Smith argue that the **min** rule yields counter-intuitive results because it takes account of only one of the combining concepts. For example, two red squares will be equally typical of *red square*, as long as their values for the poorer of the constituent concepts, *red* or *square*, are equal. So if they are equally poor shades of red, they will be equally good as *red squares*, even if one is a worse example of a square, so long as both are better examples of a square than of a red object. This counter-intuitive result cautions against the **min** rule once more. And the multiplicative rule is problematic in that it appears to over-compensate for the non-interactive behaviour of the **min** rule, by assuming that both constituents make an equal contribution to the combined concept. Hampton has demonstrated that there can, in fact, be unequal contributions - what he terms non-commutativity - in combinations (Hampton, 1987).

Although other combination rules have been proffered (e.g., (Jones 1982)), Osherson & Smith make a plausible case that no simple, knowledge-free function can properly map the typicality structures of constituent concepts into that of a combined concept. They adumbrate a thought experiment proof, as follows. Imagine a cube and a ball which each metamorphose so as to gradually approximate the other's shape. At the half-way stage, each entity should be equally typical of a ball as it is of a cube. And, if the typicality of an entity in a combined concept is a simple function of its typicality in the constituents, each of the entities should be able to be appropriately described as both a *round ball* and a *round cube*. This is because it should receive the same typicality rating as an instance of either *ball* or *cube*, and its rating as an instance of *round* should be constant. But this result is clearly in violation of very strong intuitions.

How would a fuzzy set approach characterise the problem classes? Consider Privatives. In any extensional analysis, "contradictory concepts" like *apple that is not an apple* cannot belong both to a classical set A and its relative complement not-A. This was the substance of Hoepelman's treatment of Privatives, and my criticism of that treatment. Osherson & Smith, however, take the classical set theoretic approach to be an adequate model of our intuitions concerning contradictions; hence, they argue, any account that allows an entity to belong (i.e., to have non-zero degrees of membership) in both *apple* and *not-apple* is in violation of these intuitions. An extension of their argument to Privatives is perspicuous. However, as Cohen & Murphy (1984) comment, Osherson & Smith's argument begs the question. If concepts are to be modelled like classical sets, then "contradictory concepts" do indeed denote the empty set, and Privatives denote the relative complement of the head noun set. But if concepts are not modelled by classical sets, then this objection disappears. And it is not that the fuzzy set account must allow that an entity can be an excellent example of both A and not-A; rather, being a member of both only to some degree is all that is required. And this is precisely the case made in Chapter II regarding the perspectival relativity of categorisation.

However, the cogency of Cohen & Murphy's point masks a further difficulty for fuzzy set theory: identifying a combination as Privative requires knowledge of the content of the constituents, in a way which is not permitted by fuzzy set theory. Prior to applying the combination rule, we would need to know the contents of the concepts, as is most clearly evidenced by functional privatives like *stone lion* and *pink grass*. This point presupposes that a distinction between Privatives and Affirmatives would be made through the operation of different interpretive rules. Why should this be the case? Consider again the basic min and multiplicative rules: both would imply that a statue of a lion could not be a better example

of *stone lion* than it is of either *stone* or *lion*; the same point holds for *fake gun*. But this is precisely our intuition: a statue of a lion just is a very good exemplar of *stone lion* and less so with respect to either of the constituents. So neither of the basic combination rules could handle Privatives. Now, this criticism also held for the affirmatives for which a non-intersective analysis was necessary. So the argument might then be that the same new, though unspecified rule, would operate for both types of combination. But this would mean that some crucial differences in combination behaviour were omitted - there would be no distinction between affirmative and privative, or predicative and attributive. What aspect of the fuzzy set quadruple would carry the burden of making this distinction? Only the characteristic function $TYP_{entity(x)}$ has any effect on typicality in the combined concept. And the privative/affirmative difference is most crucially marked with respect to the relation between the constituents and the combined concept. Since the very same terms can behave both privatively and affirmatively in different combination contexts, the fuzzy set theorist has two options. Either the typicality structure of a single concept is constant and the interpretive combination rule differs between contexts; or the typicality structure varies across contexts, and the interpretive rule remains constant. But on either count, the fuzzy set theorist fails to provide a formal, knowledge-free account of combination. Under the first option, choice of interpretive rule would have to be guided by the interaction of the content of the constituents. Under the second option, the flexibility of typicality will be at least partly a function of the very combination which it is being invoked to support. That is, the combination rule is vouchsafed knowledge-free because the typicality of the concepts has changed with context; but one very crucial aspect of that context is the combining concept itself - the variation in typicality is partly dependent upon precisely how the concepts are to combine. Which itself is dependent upon that variation in typicality. I think that this regress can only be broken by allowing for a more thoroughgoing role for knowledge. It is not, therefore, clear how a fuzzy set account could provide a knowledge-free mechanism for Privatives.

What, then, of Attributives? As with any extensional and set-theoretic account, fuzzy set theory is completely tied to explicit semantic attachment. Hence, there is just no possibility of arriving at an interpretation of *attractive ballet-dancer* which did not presuppose that the adjective directly modifies the head noun concept, so that the ballet dancer would be one who dances attractively.

So attempts to model concept combination in general, and the current problem classes in particular, in fuzzy set theory, would meet with little success.

III.3.2. "Intensional" Approaches

I will consider two distinct approaches, covered by the the rubric "intensional". The first is the linguistic-semantics approach involving some form of intensional logic; for example, the approach advocated by Montague (1974). Here, intensions are functions from indices (such as possible world-time pairs and contexts) to denotations. The second is the psychological approach in which an "intension" is more akin to the Fregean "sense" in that it is usually seen as some kind of representation of an entity denoted by a term or phrase.

III.3.2.1. Montague Grammar: Siegel (Siegel 1979):

Siegel's (1979) account utilises a variation of Montague Grammar. This postulates a fundamental syntactic distinction between attributive and predicative adjectives:

(i). Predicate adjective: one-place predicate type, $t//e$.

This combines syntactically with a semantically empty "be" to form a t/e (intransitive verb) phrase, and then with a term phrase (a noun, or noun phrase) to form a sentence.

(ii). Attributive adjective: type $t//e / t//e$ (CN/CN).

This type of adjective combines with common nouns to make more complex common nouns. Siegel's argument is that there is a semantic duality mirroring this syntactic distinction:

the property that an adjective represents may be bound to the meaning of a common noun it modifies in some way or it may be free, a simple predicate.
(Siegel, 1979: 226).

Those bound to the modified noun are "relative" interpretations, whilst the independent ones are "absolute". Although some adjectives are only absolute (e.g., *speckled, four-legged, naughty*), and others only relative (e.g., *actual, alleged, former*), others appear to be open to both readings (e.g., *delightful, good, clever*). Consider Siegel's example:

Marya is a beautiful dancer

On the relative reading, Marya is beautiful as a dancer: she dances beautifully, although in other ways she may not be beautiful. On the absolute reading, Marya is beautiful and a dancer: she is, then, beautiful in general.

Siegel argues that the two readings require different translations from syntax into the model theory. The absolute reading reduces the adjective to a simple extensional one-placed predicate applied to an individual; so the noun phrase is subject to a conjunctive paraphrase of two predicates:

beautiful'*(m) & dancer'*(m)

where *beautiful'* is the translation of the term *beautiful* into its semantic correlate, and * indicates that the prior string is to be understood extensionally. So *Marya*, the individual, is beautiful as well as being a ballet-dancer. In contrast, the relative reading translates into a function from properties of individual concepts to characteristic functions of sets of them. So the adjective combines with the property of being a dancer (the intension of the noun) to form an expression, *beautiful dancer*, that picks out a new subset of dancers. The adjective does not modify an individual, but applies to the intension of the head noun. The interpretation is then:

(beautiful'(^dancer'))*

Here, ^ indicates that the intension of *dancer* is being modified by the adjective. This does not support the inference that *Marya* is beautiful in general: the meaning of the adjective is relative to the meaning of the noun.

What, then, of those adjectives whose behaviour can be either relative or absolute? She argues that we can account for a relative meaning occurring in predicate position and an absolute occurring in prenominal position by the use of two semantically empty syntactic transformation rules. The first rule accounts for the fact that *dancer such that she is beautiful* and *beautiful dancer*, have absolute readings. More important for current concerns is the rule that accounts for the presence of some relative readings in predicate position. This rule deletes a dummy common noun. Semantically,

The dancer is good
will be something like

The dancer is a good Q

where Q ranges over common nouns, but is interpreted uniquely for any utterance in terms of the context. *The dancer is good* is most likely to be interpreted as *The dancer is good as a dancer*; that is, the Q will usually be interpreted as co-intensional with the head noun. A switch in Q away from the head noun is possible for a post-nominal but not for a prenominal adjective.

How might this view address the problem classes? Firstly, does the picture provide for sufficient flexibility in the interpretation of noun phrases? That is, does the relative/absolute distinction suffice? It may not: it does not appear to be the case that there are only two possible interpretations for the adjectives. Rather, there are just many different possibilities for a particular A + N combination. Superficially, the relative/absolute distinction parallels the explicit/implicit semantic attachment distinction. Indeed, the relative interpretation of an adjective carries the same kind of implication as the explicit semantic attachment of an adjective's meaning to the head noun. But the absolute reading is far stronger than implicit

semantic attachment: indeed, it is a denial of the role of *any* semantic attachment in the interpretation of the adjective. Both implicit and explicit attachment are qualities of attributives: neither can support a conjunctive paraphrase of a combination. And notice that in absolute readings, the adjective modifies the referent of the noun in a "deeper" way than in the relative case. Consider the example of *poor actor*: a relative reading would modify the acting behaviour, whereas an absolute reading would modify something more fundamental (or indeed, everything) about the actor himself. And this is just too strong: recalling the "ballet-dancer" example of Rudy and his friends, there is no obvious sense in which the meanings of *delightful* for Rudy and Ron are any more or less fundamental than each other, and certainly no more fundamental than Reg's. There are numerous ways of interpreting a combination which neither involve a simple dependence upon the head noun, nor upon the referent of that noun considered in a more "fundamental" way. Implicit attachment allows for the construal of the referent with respect to many other nouns, including one that involves a fundamental referent-modification. What Siegel calls an absolute modification is simply an interpretation of the adjective with respect to a noun type which is not explicitly given. This point was made by Platts: even apparently predicative uses of adjectives are themselves not open to a one-place predicate analysis. The apparently absolute reading of *beautiful* itself conceals an imprecise sense for *beautiful*: we cannot say in precisely what sense(s) *Marya is beautiful*.

The conception of an absolute modification is, then, a little too strong. The same might be said of relative modification. The implication of the CN/CN treatment is that we evaluate an adjective with respect to all aspects of the noun's intension: that is, with respect to the function from properties of individual concepts to characteristic functions of sets of them. But this requires that we take account of all of the content of the noun, which is modified by the adjective. This seems odd if we consider combinations like *intelligent man*: here we surely modify only one aspect of *man*. Since Siegel's concept of an intension is radically non-epistemic, it manifests none of the content which we would ascribe to a concept. So there is no clear way of partitioning such a formal "intension" into aspects which are and are not relevant to a particular case: how do we go about taking into account only certain aspects of a function (which is not the same as restricting its domain or range)?

Another question concerns Siegel's claim that relativity of the sort *A for a Q*, where *Q* is a count noun, is not open to attributives/prenominal *A*'s. This assumes that her relative/absolute distinction is exhaustive, an assumption that has been challenged. In fact, the current view of attributives is precisely that they express relative categorisation judgements or property ascriptions of the type *A for a CN*. Consider the transformational rule

that allows a relative reading in predicate position. Here, the locution *the dancer is good* is seen as elliptical for *the dancer is a good Q*, where *Q* might be seen as just the kind of covering sortal required by the relativity of categorisation. It might then account for attributivity, if the rule also applied to pronominal modifiers. However, Siegel states that it cannot. And note that it would require another kind of relativity: relativity of the interpretation of the adjective to a common noun which may or may not be co-intensional with the head noun. Such relativity allows for great intensional flexibility - far greater than that allowed by the relative/absolute dichotomy. And it is precisely the kind implied in implicit semantic attachment. Siegel advances no plausible reason why this kind of relativity should not be possible for pronominal modifiers, although it is clear that such a possibility would render the absolute/relative dichotomy non-exhaustive.

How would Siegel account for Privatives? She notes that *former*, *alleged* and *ostensible* are relative adjectives, so they would receive an intensional treatment. So *alleged thief* would emerge as:

(alleged'('thief'))*(x)

This has the crucial effect of placing the noun inside the scope of the intensional operator, which therefore means that we cannot infer that the alleged thief is a thief in any sense. This is simply because intensional contexts are opaque: they do not permit existential "quantifying in". It seems then that the account cannot countenance the perspectival relativity of privatives. Such a view parallels that of "intensional" A's in Montague Grammar generally. It does, however, have an advantage over set-theoretic approaches in making the effects of a Privative A relative to the N modified, which is fundamental to any interactive view of combination.

A second difficulty is that Siegel's account makes no distinction between Equivocating and Negating Privatives. That is, both *alleged* and *former* receive the same treatment; it seems reasonable to infer also that the same would hold for *fake*, *false*, *apparent* and *possible*. A further point is that, as a result of restricting the possible kinds of relativity for relative adjectives, the account cannot provide any view of what kind of entity an object referred to by a privative combination might be. That is, although we can say that an alleged thief is not a thief, or a fake gun not a gun (which has already been challenged), we cannot say what it is. A fake gun might be a toy or replica, but without allowing the meaning of the adjective to be relative to a noun which is non-co-intensional with the head noun, there is no obvious way of providing for this fact.

The argument, then, is that, although Siegel's account makes some interesting distinctions

between types of modification, her account has inappropriate consequences for attributives and privatives. The possibility of different "levels" of modification of a noun by an adjective, on different occasions of use, is a crucial contribution, and will be considered in III.4.

III.3.2.2. Psychological Accounts:

I have already commented on the failings of definitional accounts of concept combination. I will not discuss "prototype" views here; rather, I will focus on views which attempt some direct consideration of the content of concepts.

There appear to be two directions open to psychological intensional approach, regarding combination. The first is to consider combinations through the partitioning of the content involved in the interaction. Such a view has been advocated by Miller & Johnson-Laird (1976) and Smith & Osherson (1984). The second is to propose a major role for extra-lexical knowledge, as suggested by Medin & Shoben (1988), and Murphy (1988).

Miller & Johnson-Laird (1976) divided the properties of a noun concept into three types: "defining" properties, "whose values are fixed"; "characteristic properties", whose values operate like defaults; and "peripheral properties", whose values can be regarded as free variables. Depending upon which of these properties are to be modified by the adjective, different combination principles will operate. Consider *red ball*: the peripheral noun property of "colour" is modified. The basic mechanism for this is an intersective one: a red ball is therefore an entity in the intersection of all red things and balls. They argue that modification of the other types of properties are best understood as deviations from this basic rule. The crucial point is that the modifier of a peripheral property can take any value whatever, so long as it is of the right type.

Modification of characteristic properties is exemplified by the property of size, for which, they claim, most everyday objects have a characteristic range of values. A child's ball may vary in size but it cannot vary freely. So, adjectives concerned with size must be evaluated with respect to the particular characteristic range: *big* for a child's ball is not *big* for a football pitch. Hence, in modifying a characteristic property, an adjective is constrained by the concept it modifies; this has the implication that a simple intersective mechanism cannot suffice. Rather, the noun must have priority, and the adjective is then evaluated within the range permitted by the noun. However, where the adjective cannot specify a value within characteristic range of the noun, the adjective's value takes precedence, and replaces the default or characteristic value. So although there may be characteristic ranges for hair colour, a modification of the concept *hair* by *green* may occur where the uncharacteristic

value takes precedence.

The modification of definitional properties results in either redundancy or anomaly. Redundancy results when the adjective specifies a value already specified by the definition, and anomaly (or nonsense) when it specifies some value other than that given. Examples of the former include *unmarried bachelor*, and *canine dog*; and the latter, *dry water*, *noisy silence*, and so on.

Miller & Johnson-Laird's account is more of a sketch, and should not be expected to cover the range of issues attending adjectival modification. Nonetheless, the possibility of different types of property being modified in different ways and to different effect, is an important one, and forms the germ of the current approach.

How does the account fare with respect to Attributives and Privatives? It seems problematic on both counts. The modification of "peripheral" properties is by an intersective mechanism; this is, by definition, inappropriate for attributives like *attractive ballet-dancer*, and it is unlikely that being attractive is characteristic or defining of ballet-dancers. However, even if it were characteristic, any combination would nonetheless presuppose explicit semantic attachment: there is no mechanism to facilitate wide variation in the properties of the noun. There is then little prospect of the view's accounting for attributive behaviours. And any modification of defining properties would result in anomaly. This was Keil's prediction, which was denied in the last chapter. But note what this means: the approach disallows the possibility of a group of Functional Privatives like *stone lion* or *square basketball*, since these combinations would simply be anomalous. Nor could the view allow for the Functional Privatives which I have labelled "Equivocators", such as *pink grass* or *wooden frying-pan*. For Miller & Johnson-Laird, these are odd modifications of characteristic properties - the result would most definitely still be affirmative. But it seems clear that these combinations do result in an uncertainty concerning the inference from the NP to the head N; that is, that the modification renders doubtful the "defining" properties of the head concept. But Miller & Johnson-Laird cannot allow for this because they do not allow for the propagation of modification effects across the different types of property. So denying the characteristic material of a frying-pan cannot lead to doubt over its defining function. This ultimately non-interactive assumption casts a great deal of doubt on the adequacy of their picture.

Smith & Osherson (1984) have proffered a "Modification Model", which posits two basic mechanisms to derive the sense of a combination from its constituents. The adjective

directs the reorganisation of the constituents' content by restricting the range of acceptable values, and by increasing the importance of the corresponding dimension. So the A plays the central role in determining the properties of the combined concept. In essence, a N concept is represented as a set of dimensions, which differ in their diagnosticity, coupled with their possible values. The value of a dimension is akin to an expectation or subjective probability concerning a referent's having the property: if apples are typically red, then the "red" value will be given a higher "vote" than "green", in the concept for *apple*. Both the diagnosticity and the values of dimensions may alter as a result of combination. For example, if we combine *red* with *apple*, then the "vote" for the value of "red" will increase in the combined concept, and the diagnosticity of the dimension of colour will increase, with respect to the input noun.

The modification model was able to account for much of the data which confounded the fuzzy set approaches. However, it faces difficulties both in general and in its application to specific cases. Firstly, I will note some of the ways in which it falls foul of the Manifestation Constraint. Crucially, it seems to assume all three of the basic postulates of sense selection: sense-invariance, post-access combination and explicit semantic attachment. All three aspects have been denied. A further point is that the view assumes that the modifying effects of a particular adjective will be the same, or highly similar, for every noun concept it modifies. That is, although the "vote" assigned to an adjectival property might vary between different noun concepts, it nonetheless modifies the same aspect of the different nouns. This was denied by the findings of Medin & Shoben (1988), and Murphy (1988). Not only is the modifying effect constant across noun contexts, it is also restricted to the particular property or dimension which the adjective overtly labels; so the modification of *spoon* by *wooden* alters only the "material" property. This assumption of the independence of properties is shared by Miller & Johnson-Laird, and denied by both the findings of Medin & Shoben, and the possibility of there being a class of Functional Privative combinations (see above). Properties are correlated, and the modifying effect of an adjective is propagated through such correlated attributes. An implication of the Modification Model is that no extra-lexical knowledge is necessary to the understanding of a combination: the adjective itself specifies which property of the noun is to be modified. Such an assumption flies in the face of the evidence.

Regarding the problem classes, it is not completely clear what kinds of predictions the Modification Model would make. If we consider their experimental stimuli, we find that all of the combination types are predicative, or at least that the modifier can occupy predicative position with respect to the noun: for example, *red apple* can be expressed as *the apple is*

red, metal bicycle as *the bicycle is metal*, and so on for all of their stimuli (Smith & Osherson, 1984: 343, 348). Since they used no examples of attributive adjectives, their evidence cannot be said to apply to all adjective types; and the other evidence contrary to their approach appears to militate towards the conclusion that the Modification Model could not easily accommodate attributives. They are just too flexible to find a place in the rather inflexible framework of the Modification Model. They appear to demand a detailed and complex interaction of properties, linked to knowledge-rich processes; as a result, we would have to conclude that the Modification Model could not, as it stands, be a complete model of their behaviour. However, as Murphy comments, it could be subpart of a more complex and knowledge-driven model.

The picture is similar for Privatives. None of their stimuli are either Negating or Equivocating Privatives. Indeed, it is not clear that, with the range of tools in the theory, the possibility of Functional Privatives could be entertained. Why is this? The first reason is that, as I have commented, Privatives appear to involve either a denial or a doubting of the central properties of the head noun; but the Modification Model makes no allowance for centrality. No adequate account of Privatives could derive from a consideration of diagnosticity only, and this is their major construct. Nor could the Equivocating Privatives be accounted for through the correlation of a denial of diagnostic properties with the doubting of central ones. Even if the model did have a "centrality" construct, the independence of properties would disallow this prospect.

So it seems that, although the Modification Model represents a rare attempt to specify in detail what might happen during the combination of concepts, its assumptions make it non-generalisable to the current problem classes, and at best a partial account of its own problem set.

The second option for psychological intensional models - to allow for the influx of world knowledge - is seen in the model proposed by Murphy (1988). This is a variant of a Concept Specialisation Model, and an extension of the role modification model proposed by Cohen & Murphy. On this view, an adjective provides a restriction on, or specialisation of, the noun concept by filling an appropriate role or slot. For example, in the combination *apartment dog* the concept *apartment* would be a habitat, and would replace all other fillers in the "habitat" slot. Thus the concept for *dog* has been specialised, in that its habitat has been described more specifically. Now, Murphy's argument is that combination is knowledge-intensive, for two reasons. Firstly, outside knowledge must often be consulted in order to ascertain which is the appropriate slot to specialise. Secondly, knowledge is

crucial for a post-specialisation elaboration of the concept: making the combination more coherent and complete. Precisely which knowledge, and how it is accessed, is left open by Murphy. We have, however, already seen reasons to doubt that the role-modification model can avoid a vicious circularity in its knowledge-deployment and specialisation processes. Briefly, in order to know which role to modify, we need to know what kind of an entity we are describing with the combination; but to know that entity type, we are supposed to require the particular role or mediating relation. This problem is analogous to the "frame selection problem", which beset Minsky's frame theory (Minsky 1975). The other arguments adduced with respect to Cohen & Murphy's view are also applicable here. Hence its ability to account for the problem classes must be doubtful, even though the model is underspecified.

Somewhat more promising, though equally sketchy, is the account that Medin & Shoben give of their experimental findings. They propose a "hybrid model", utilising aspects of Exemplar theory. Since the assumption that combined concepts are derived exclusively from stored exemplars is untenable, they argue that new representations may be derived as required both from lists of exemplars and from other causal knowledge about the world. Both membership in categories and similarity relations may be affected by such causal knowledge. And such knowledge is important in determining which properties are more or less mutable or central than others. Although the process of combination is even less specified than in Murphy's account, this picture has the crucial advantage that it allows a role for two kinds of knowledge: the sense relations embodied in world knowledge, and the referent relations depicted by the relations between exemplars. This is close to the requirement noted in the Manifestation Constraint.

However, in order to see this, we need to consider the second major dichotomy in approaches to adjectival modification. This is the dichotomy between different "levels" and "types" of modification.

III.4. Levels of Modification

In the next two sections (III.4 and III.5), I consider two orthogonal ways in which the combination process in a Combinations theory might support flexibility. Firstly an adjective might modify a noun at different "levels" on different occasions. Secondly, the adjective's effect might vary not as a result of its modifying different properties of the noun, but as a result of the noun shifting in content. The conclusion will be that both types are desirable: the Manifestation Constraint requires that we allow for multi-level and multi-attachment

modification.

III.4.1. Levels of Modification

The crucial question here concerns the number of different levels of modification. Set-theoretic accounts assume single-level modification. When an adjective modifies a noun, we obtain some variant on the intersection of their denotations. The adjective, then, modifies all of the noun concept, since the intersection contains all and only those objects which satisfy both constituents.

III.4.1.1. Dual-Level Modifications:

Roughly equivalent types of modification have been proposed by Bolinger (Bolinger 1967) and Quirk et al (Quirk Greenbaum Leech & Svartik 1972). Bolinger distinguished "referent-modification" from "reference-modification": the former involves the ascription of a property to the noun entity considered in a way which is "deeper" than a description by the noun. The latter involves an ascription at a less deep level. There would be, then, two possible readings of

Poor actor

Reference-modification produces a description of an actor who, as an actor, is poor - he is unskilled at his profession. In contrast, referent-modification produces a description of an actor who, as a person, is poor, presumably in the pecuniary sense. Quirk et al's distinction between "inherent" and "non-inherent" modification fosters the same outcome: the former characterises the noun referent "directly", whilst the latter characterises some aspect of that referent. So referent- and inherent- modification appear to be co-extensive, as do reference- and non-inherent modification. A similar effect would be achieved if, following Siegel, *poor* has both a relative and an absolute meaning. On the relative reading, *poor actor* would mean poor in acting, whereas on the absolute reading, it would mean poor in the pecuniary sense. But in the latter case, the reference class against which poverty is measured might be either the class of actors or the class of people in general; such flexibility appears to be ruled out by the dichotomies of Bolinger and Quirk et al.

Another difficulty concerns precisely how to ascertain what is more or less inherent in a noun entity. The assumption appears to be that, if the meaning of the adjective is evaluated relative to the head noun, then the modification is non-inherent, or reference-modification. This assumption has a problematic implication: if a *poor actor* is *an actor who has poor acting skills*, this implies that the poor actor is indeed an actor. And this subsective implication is open to doubt: a poor actor may be someone who is not an actor, but is bad at

acting. The N "actor" is one of the *nomina agenta*, and as such modifiers could modify it at either of the levels noted in III.1. Indeed, the modification of such a N could also occur at the "deeper" level of an "inherent" modification. Without detailed alterations, the Bolinger and Quirk et al conception of non-inherent/reference-modification is unable to handle these phenomena of attributivity.

A similar conclusion emerges from a consideration of inherent or referent-modification. In the case of an *attractive ballet-dancer*, inherent modification would result in a ballet-dancer who, as a person, is attractive. But it is not clear that this reading concerns any more "fundamental" properties of the head noun than the non-inherent readings. For example, an attractive ballet-dancer might be someone who is attractive in facial appearance, in personality, or even in style of clothing. It is not obvious that these possibilities are any more inherent than his being able to dance ballet attractively. Indeed, dancing ballet attractively is an inherent modification of the person qua ballet-dancer, having an attractive face is inherent if the person is considered qua a physical person. It depends upon the criterion against which inherentness is evaluated. The implication once more is that we must take account of the perspectival relativity of such ascriptions; and this has no obvious link with the dichotomies proposed. It does not seem, then, that these phenomena are to be accounted for by an exhaustive dichotomy of modification levels.

III.2.1.2. Multiple-level Modifications:

Miller & Johnson-Laird's is an example of a three-level account of combination. Here, an adjective might modify a noun concept at the peripheral, the characteristic or the defining levels. I have already discussed this view, and will make only one comment here. They assumed that a particular adjective would have an invariant effect on a concept: it modifies the same noun in the same way in all contexts. There is no obvious way in which this view could account for attributive combinations, where the same adjective can have markedly differing effects on a particular noun concept. The implication, then, is that distinctions between types of content can yield only a partial account of attributivity and privativity.

A major difficulty with all of the views considered is that they assume a very small finite number of ways in which an adjective might modify a noun. That is, since the distinctions proposed are held to be exhaustive, modification only occurs with respect to a certain range of properties with a certain range of effects. This results in insufficient flexibility to account for attributives and, by extension, privatives. However, an account that avoids such rigidity has been put forward by Bartsch (1986, 1987).

Bartsch's approach utilises a version of Montague/Kaplan Grammar, distinguishing between the "character" and "content" of an expression. The former denotes a function from possible world/time pairs and contexts into denotations; and the latter is a function from world/time pairs into denotations. Characters are functions from contexts to contents, and contents are then functions from world/time pairs to denotations. So the character is the contextual determinant of the content. Both characters and contents can be either stable or unstable. Characters which determine the same content for all contexts are "stable": terms with such characters are interpretable in a context-independent way. Unstable characters yield different contents in different contexts. Stable contents are ones that output the same denotation in all world/time indices. Unstable contents yield different denotations at different indices.

Bartsch differentiates those adjectives which directly express properties (whose senses comprise a stable character and an unstable content), and those that require a particular contextual specification of the "thematic dimension" in combination with which they express properties (that is, they have an unstable character and content). The former are "Dimensionally Strongly Determined" expressions (and include *liberal*, *broadchested*, *humourless*), and the latter are "Dimensionally Weakly Determined" ones (which include *good*, *strong*, *satisfactory*). The argument is that a crucial aspect of the context for the interpretation of Dimensionally Weakly Determined expressions is the current thematic dimension. So, for example, in interpreting *John is a good swimmer*, we need to know exactly which aspect of swimming is under consideration: speed, style, or whatever. In contrast, the Dimensionally Strongly Determined expressions "presuppose" their particular dimension of interpretation: so *sick* presupposes "health", and *red*, "colour", as the thematic dimension.

A thematic dimension is understood as a set of properties: for example, one might contain all "swimming" properties; within this, we could specify sub-dimensions relating to speed, style, and any other aspects of swimming which might be evaluated independently. A property is a singleton subset of the thematic dimension, and an adjective indicates a specification within the dimension which may or may not correspond to a single property. Thematic dimensions are then domains of interpretation for adjectives. Contents are thus usually functions from indices into properties, and characters are functions from contexts (including the thematic dimension) into contents. Bartsch labels the character functions "pre-properties"; pre-properties can be understood as the set of all properties which the function assigns in all different dimensions. In combination with a thematic dimension, an adjective may or may not yield a single property. If it does not, further specifications may be possible to increase precision, and thus to assign a single property. For example, *John is*

a *good swimmer* can be taken in two ways: either as a general property ("good in swimming under at least one of the relevant aspects"), or as a pre-property. So *good swimmer* expresses a more precise pre-property than *good*, but still must be applied to another thematic dimension in order to yield a particular property ascribed to John. So a pre-property may be a function from a thematic dimension to a more specific pre-property, or to a single property.

Bartsch avers that thematic dimensions can be differentiated by the addition of perspectives and aspects: so a property may be located in more general or more specific dimensions. If we look at an individual from different perspectives, we project it differently onto the same or a different dimension. These projections determine the relevant properties an entity has, relative to a thematic interest. Dimensions can be of greater or lesser generality, and we may take into account properties of different types and levels of centrality. There is not, then, on Bartsch's position, a procrustean ordering of "levels" of modification, from which a particular modification must choose.

A pre-property indicates a (pre-)property at each dimension: this occurs through the intersection of the pre-property and the dimension. The function takes the intersection-point as the value for each dimension. Iteration of the pre-property, through its application to more precise dimensions, leads to a finer level of content; that is, to more specific properties.

Bartsch (1987) defines a schema of property construction, where two different thematic dimensions are applied to a single pre-property, thus composing a property from the pre-properties. If we consider the combination *excellent swimmer*, under the dimension of "style" then we have two dimensions: "swimming" (t_1), and "style" (t_2), and one pre-property (P^0) in "excellent". Then we can construct two pre-properties, $P^1 = P^0(t_1) =$ "excellent (swimming)", and $P^2 = P^0(t_2) =$ "excellent (style)". There is a combination & of the two pre-properties into a single pre-property, $P^3 = P^1 \& P^2$, which is the intersection of both. If this intersection is a singleton, i.e., $P^3 = \{P\}$, then the property P is said to be combined out of the pre-properties P^1 and P^2 . In this example, the property is "excellent (style and swimming)", which may be expressed by "excellent in swimming with respect to style". The schema of property construction is:

$$P^0(t_1) \& P^0(t_2) = P^0(t_1 + t_2) = P^0(t_{ij}) =$$

$$\{P \mid \text{there are } P^1, P^2 \text{ with } P^1 = P^0(t_1) \text{ and } P^2 = P^0(t_2) \text{ and } P \in P^2(t_1) \text{ and } P \in P^1(t_2)\}$$

In this case, the resulting pre-property is a set which contains P , where $\{P\} = P^2(t_1) =$

$P^1(t_2)$ = "excellent style (qua swimmer)" = "excellent swimmer (qua style)". Just such a mechanism will play an important role in the account of combination\$ to be given.

How might this apply to the problem classes? Bartsch (1986) argues that attributives express pre-properties, or characters, but not contents. Such adjectives cannot be interpreted as predicates independently of a dimension. So *attractive* maps only the dimension containing the properties of *ballet-dancer* on the property *attractive as a ballet-dancer*, which does not imply being a ballet-dancer. This directly captures one of the aspects of attributives, and also allows for the addition of a condition which specifies that the person who is attractive as a ballet-dancer also is a ballet-dancer. But note that the schema of property construction cannot allow for the third aspect, which is that the ballet-dancer is attractive in some respect other than manner of dancing. This is because the schema involves the intersection of *ballet-dancer* and *attractive* properties: so the constructed property must be evaluated with respect to the head noun. That is, the schema presupposes explicit semantic attachment.

Bartsch makes no direct comment about Privatives, but we might surmise that at least Proper Privatives could be seen as expressing pre-properties. So *fake* expresses the fact that the NP entity has some of the appearance properties of the head N: but exactly which ones are uncertain. A fake gun looks like a gun in some sense; sham rage resembles real rage in some manner; an apparent friend behaves like a real friend in some respects. In all of these cases, we might argue that we need some other source of information in order to be able to state in precisely which manner the NP entity "looks like" the head N entity. The parallels between relativising adjective meaning to a dimension, and relativising categorisation and identity to an implicit sortal, should be clear. We might then view the dimension as expressed in an implicitly attached noun concept. And this is exactly what I have already claimed.

Bartsch's approach provides for a great deal of flexibility in the interpretation of adjectives and nouns. The contribution of the adjective to a noun phrase is relative to both the combining noun and the thematic dimension; and it may involve properties of any or all levels of centrality. This marks a significant advance on the cognitive theories which advert to extra-lexical influence without providing any precise account of the way in which it might operate.

There are, however, problems in incorporating the ideas as they stand. The first is the commitment to realism in Montague-Kaplan Grammar: possible-worlds semantics is perhaps the

most extreme realist position available. However, the mechanism of the construction of properties from pre-properties can be seen to be independent of that framework, and independent of the character/content distinction. It is in this way that I shall utilise it. A second problem might seem to be in the intersective analysis of property-construction. There are two aspects to a reply to this problem. The first is that intersection takes place in a content-driven manner: it is not knowledge-blind set intersection. It is not, therefore, open to the criticisms I have set out. The second point is to concede that, as Bartsch outlines it, the view does assume sense selection and explicit semantic attachment. As it will be utilised in the current view, the mechanism will form part of the apparatus of sense generation, in which the thematic dimension may be provided by the implicitly attached N concept.

Bartsch's account allows for different degrees of specificity and different types of properties to form part of a combination. However, the adherence to explicit semantic attachment restricts the extent of flexibility permitted. In order to see how this might have a role, we must consider the possibility of different "types" of modification.

III.5. Types of Modification

The last section was concerned with flexibility arising from different ways of interpreting the adjective. One way in which such flexibility arises is in the relativity of adjective meaning to noun meaning. However, noun meaning itself is similarly flexible: so adjective meaning could be relative to a flexible noun meaning.

I have already mentioned the possibly dual nature of nomina agenta. We appear to require an extension of this to other nouns, in an approach which allows for as many interpretive levels and types as there are informational needs. If the adjective is influenced by the noun, and the noun by the adjective, then we must take account of the interaction of the adjective and noun.

What is the nature of semantic attachment? The source of attachment lies in the fact of partiality of lexical representation. Lexical concepts are partial objects; in the case of many adjectives, following Bartsch, this amounts to their expressing pre-properties. They require another source of information to indicate in exactly which respect they modify the noun. On Bartsch's view, this was the thematic dimension, expressed through the thematic interest of a discourse. Bartsch appeared to omit the possibility that the noun could play a crucial role in specifying a dimension; or that the interaction of the adjective and noun could constrain the selection of a dimension. I shall advocate both possibilities.

Platts (1979) distinguishes between two distinct types of attributives: explicit and implicit attributives. The first derives from explicit semantic attachment of the adjective to the head noun: so

Rudy is an attractive ballet-dancer

is read as

Rudy dances ballet attractively.

In implicit attributives the adjective is attached to some noun concept other than the head noun; the target then reads as:

Rudy dances ballet and is attractive in some other way

A specific example of this would be:

Rudy dances ballet and is attractive to look at

The implicit reading does not imply that, if the adjective does not ascribe a property directly to the referent considered as a member of the head noun type, it must therefore do so in some more fundamental manner. This is the gist of the dichotomies mooted by Bolinger, Quirk et al, and Siegel. And this is not what the explicit/implicit attachment distinction is about. Explicit attachment is more or less equivalent to Siegel's "relative" interpretation of an adjective: the explicitly given noun concept provides the domain of interpretation for the adjective. Unlike reference/non-inherent modification, this does not imply that a modification based upon explicit attachment is necessarily "shallow". It seems that this assumption derives from not separating out the effects of the adjective from those of the noun: the examples which motivate the Bolinger/Quirk et al distinction utilise nomina agenta, which just have the quality of being interpretable in two distinct ways; such examples do not necessarily tell us very much about the behaviour of adjectives in general. If the explicit reading of *attractive ballet-dancer* is an example of a non-inherent modification, then such a reading of *fake ballet-dancer* must be inherent, since it has the effect of altering the sortal type of the entity. So, explicit attachment leads to non-inherent modification only in particular circumstances: where the interaction of the adjective and noun supports it.

The attachment account is therefore able to provide a treatment of the two readings of attributives which are based upon explicit attachment: the person who is attractive at ballet-dancing (though not a professional ballet-dancer), and the person who is a ballet-dancer who is attractive at ballet-dancing. These just indicate the ascription of different ranges of properties to the combination sense from the head noun. For the first case, the ascription involves only the diagnostic properties of the noun, whereas for the second it involves more central properties.

Implicit semantic attachment diverges even further from the other constructs. An

inherent/referent-modification interpretation of *attractive ballet-dancer* is only one of the options for implicit attachment, and an imprecise one at that. An inherent modification would modify the noun referent not as a member of the category of ballet-dancers, but in some more direct and inherent manner. Presumably the interpretation would be that the ballet-dancer were attractive in some unspecified, though essential way. And this is just the general form of an interpretation based on implicit attachment, although the latter would not be committed to the inherentness of the modification. And note that we can envisage Bartsch's mechanism operating here. The general form of the implicit attachment interpretation might involve the ascription of an adjectival pre-property to the head noun: the general property of "attractive in some unspecified manner", or the pre-property of the equivalence class of "attractiveness" properties. With the application of a more precise thematic dimension, this could then yield a more specific property. Partitioning the range of properties that might be ascribed to a ballet-dancer as a person, we could differentiate between sub-dimensions relating to physical appearance, facial appearance, personality, and so forth. And the intersection of a sub-dimension and the pre-property then specifies a detailed property or set of properties. There could, then, be as many detailed readings of *attractive ballet-dancer* as there are ways of specifying the pre-property through the application of a dimension. From this vantage point there is no motivation for restricting ourselves to an exhaustive dichotomy of modification-types based upon whether that modification is more or less "inherent".

It seems, then, that the semantic attachment approach might have more success with attributives than the other views which have been canvassed. Before going on to discuss what kind of a relation semantic attachment might be, I will mention its possible application to privatives. For the attributives, the thematic dimension derived from construing the noun entity under a different sortal heading: a superordinate of the head noun. So a ballet-dancer was construed under the heading of, say, *person*, which was then subdivided along thematic lines. So there was a shift upwards in terms of a hierarchical organisation of general knowledge. In privatives, the shift involves an alteration of sortal type. I noted this in relation to Bartsch's approach, but it will stand reiteration. A fake gun can be a fake in many different ways - for example, as a toy, a replica, a model, a sculpture, or whatever. This is directly analogous to an attractive ballet-dancer's possibly being attractive in many different ways. We can, then, view the combination as inheriting, firstly, a pre-property of the equivalence class of all gun "appearance" properties (or the general property of "looking like" a gun). Only with the specification of a detailed thematic dimension, deriving from the implicitly attached N (as knowledge of what kind of entity the fake gun is), can we derive properties from the pre-properties. Such knowledge may be given through an

instantiation as, say, a replica. In such cases we construct a property from the intersection of the application of the pre-property to the head noun and the implicitly attached noun: the appearance properties which a gun and a replica, or a statue of a lion and a lion, have in common.

Now, to return to the question of semantic attachment. What kinds of objects stand in the relation of semantic attachment? I have employed the locution that the adjective is semantically attached to a noun concept. That is, the adjective's domain of interpretation is the particular thematic dimension which is represented in the noun concept. However, the selection of the implicitly attached noun is constrained by the interaction of the adjective and the noun: whatever a stone lion is, it must be something which looks like a lion, but which is made of stone. An entity which is a member of the implicitly attached noun category must have properties which indicate a specification within the pre-property; and the pre-property arises from the interaction of the adjective and noun. Hence, it is a complex relation, in which the adjective is interpreted in the light of the implicitly attached noun, and the noun phrase entity is, in effect, re-categorised as a member of the implicitly attached noun category. In a sense, then, the implicitly attached noun is a key into a part of the general knowledge base which would otherwise not be considered; it is thus a source for property inheritance.

It is important that semantic attachment cannot be extensional - it cannot be to a set. Platts argues the case as follows. Consider the sentence *Theo is a large flea*. If Theo is a large creature with a heart, and if all and only creatures with hearts are creatures with kidneys, then an extensional account requires that Theo is a large creature with kidneys. But now, Platts suggests, consider an abnormal situation as regards size development of creatures with kidneys, and a normal (or reverse abnormality) one for creatures with hearts. In such a case, the conclusion that *Theo is a large creature with kidneys* is false, even though the premises are true. So the semantic attachment of *large*, here, cannot be to the set of creatures with kidneys and hearts. Attachment is, on my view, to another concept, construed in terms of properties.

Another issue concerns whether the semantically attached noun is "given" prior to combination, or accessed on the basis of combination. My claim is that its accessing is constrained by the combination. For attributives, there appears to be a thematically-determined choice over whether to employ explicit or implicit attachment; for privatives, the combination appears to demand that an implicitly attached noun is employed (although the precise one chosen is left open).

What kind of a relation is semantic attachment? It is fundamentally a relation of property inheritance, or property specification: the semantically attached object forms the source of interpretation for the adjective (in explicit attachment) or for the combination as a whole (in implicit attachment). Hence, the relation is crucial to the process of concept extension: to the ascription of properties to a sense.

I claim, then, that an adequate account of the problem classes must wed multi-level to multi-attachment modification. This can be achieved by allying the flexibility of semantic attachment to the property-construction schema of Bartsch.

III.6. Summary

I Have argued that, in order to achieve an adequate description of the problem classes within the Manifestation Constraint, we must adopt a Combinations theory, in which the contents of the head and modifier interact. However, in order to achieve the kinds of flexibility and specificity required in understanding the problem classes, we need to add in two degrees of freedom. Firstly, the schema for the construction of properties from pre-properties and thematic dimensions, provided by Bartsch. Secondly, the possibility of implicit semantic attachment, whereby the thematic dimension for property-construction is provided by an implicitly attached N. However, in order to see how these aspects might fit into a Sense Generation view, we will need to be more precise concerning some of the basic notions I have been using. Hence, in the next Chapter, I discuss defeasibility, molecularity, perspectives, and a mechanism for property inheritance. I also make some comments on the distinction between lexical concepts, senses, and word-meanings. This will then allow the application of the current view to the problem classes.

CHAPTER IV

On Defeasibility, Perspectives and Property Inheritance

In this chapter, I will attempt to be more precise about some of the conceptions I have advocated. So far, we have a view of concept combination that takes lexical concepts as partial objects, with defeasible extensions in which they may inherit properties from an implicitly attached N. The argument has then been that a referent possessing the properties defined in such a generated sense can be categorised as a member of the NP and head N under appropriate perspectives.

In this Chapter, I shall be more precise concerning what kinds of properties are ascribed to a lexical concept (IV.1), and just how these ascriptions may be defeasible (IV.2). A particular type of defeasibility may attend upon C-relations, a type that will prove crucial in the treatment of Privatives. Regarding property inheritance, I note an intuitive distinction between "motivated" and "non-motivated" inheritance (IV.1), and advocate a motivated approach, based upon the construction and transfer of analogies (IV.4). An analogical mechanism receives support both from cognitive psychology (and hence the Manifestation Constraint), and from its interaction with partiality. I outline my notion of a perspective in IV.3, distinguishing between two different types, and countering a possible problem. And in IV.5, I attempt to synthesise the components of categorisation-relativity, perspectives, and partiality. This requires some alteration to Landman's definitions, and a more precise formulation of the categorisation schemas (AC and PC) given at the end of Chapter II. In IV.6., the foregoing discussion provides us with a means of articulating the distinctions and connections between lexical concepts, senses and word-meanings. And this then allows a brief summary of the view of concept combination advanced (IV.7).

IV.1. Introduction

Sense Generation carries a commitment to considering lexical concepts as partial objects. The process of generation is the process of extension of the partial object/lexical concept. We must, then, take into account the relations between a concept and its properties: what is the relation through which the use of a word supports the ascription of properties to the concept associated with the word? This is crucially related to the partitioning of content (see Chapter II). For example, if we assume that a concept has classically necessary properties, the relation between the concept and its properties is a deductive one, and must

therefore be indefeasible. Characteristic properties inhere in a concept through inductive relations: when we encounter a word in a text, we can ascribe such properties to the word's concept, providing the inductive motivation is not denied. The discussion of property centrality issued in the claim that central properties must be defeasible, with their defeat having consequences for categorisation. I argued that central properties might be those properties which, in virtue of the appropriate lay theory, are considered to be "essential" to the noun entity, as noted by Neisser (1986).

Now, in categorising something, we generally make (non-demonstrative) inferences from the diagnostic, observable aspects of the entity (the conditions of application of the concept) to the internal or non-visible aspects of that entity (the "essence" or function). Hence, we may distinguish between diagnostic and central properties, in a condition of categorisation or concept application. The inferences to both types of property are defeasible.

However, we are concerned, in Sense Generation, with a different though related condition. The condition of Sense Generation does not have the basic constraint of concept application. In Sense Generation, we engage in the construction of a representation of an entity for the purposes of the discourse. We must, then, make available to the sense generation process the central and diagnostic properties of the noun. Yet we must also respect their defeasibility under the operation of the theory of sense. My proposition is this. In the Sense Generation condition, the semantic contribution that the lexical concept makes is the ascription of central properties to the sense. However, here I take "central" to cover two distinct aspects of the entity: it includes the "essential" aspects of the entity, as well as the diagnostic properties. Henceforth, I will label these distinct sets the "Central-essence" and "Central-diagnostic" properties. My claim is that both sets manifest a C-relation to the use of a term; hence they are defeasible. So what I previously labelled "Criteria" (in the condition of concept application), are now Central-diagnostic properties. Although both central-essence and central-diagnostic properties are defeasible, the defeat of each has different ramifications for the sense which is generated. The generation of a sense will also involve the ascription of some non-diagnostic properties, which are connected to the term through S-relations. So the concept extension of the lexical concept (as a partial object) functions through the satisfaction of defeasible C-relations and S-relations. I should re-emphasise the point from previous chapters, that the C-relation is not intended as a strictly formal relation. It is not that particular concept extensions occur through satisfying some abstract C-relation. Rather there is just a multiplicity of contentful relations in the world that can support such an extension, as represented in our lay theories of a domain, or licensed by ad hoc explanations. My claim is that these relations can behave after the fashion of C-

relations, having the same qualities. There will be, then, many relations to be used. For example, a "lion" relation would relate the use of the word *lion* to the ascription of central properties to the sense.

How does the Sense Generation condition relate to the condition of concept application? We might think of the relation like this: after constructing a sense for noun phrase, we have a representation of an entity, to which the NP is applicable; so any entity having those properties can be categorised as a member of the NP category. Within the bounds of the current discourse, this sense functions as the conditions of application of the NP: as the conditions on NP-category membership. The important questions for our purposes are then: is the NP-type entity also a head N-type entity? Can the NP-type entity be described by the constituent modifier? And a crucial aspect of the response to these questions will be an appeal to the perspectival-relativity of categorisation.

The argument, then, is that the same kinds of relations - expressed as C-relations - operate both in word-meaning (i.e., concept application/categorisation of referents) and in concepts (i.e., sense generation/concept extension).

However, there is an important and oft-overlooked issue in property inheritance to be noted here. This is a contrast between what we might label "motivated" and "non-motivated" inheritance. Many accounts of inheritance appear to assume that inheritance is non-motivated; it arises, rather, from the relative positioning of the inheritor and inheritee in some structured knowledge base. The paradigm case of this is inheritance in taxonomic or hierarchical structures. Where two concepts stand in a relation of hyponymy, the subordinate inherits freely from the superordinate; in taxonomic structures, the superordinate is incorporated in the subordinate. An implication is that, where two concepts are not located within the same structure - where, for example, their superordinates are antonymic, or simply ontologically mis-matched - no such inheritance can take place. An example of this can be seen in the standard semantic network formalism (Collins and Quillian 1969, Woods 1975): here, a crucial problem was how to constrain the number of links which were to be traversed (and hence the number of concepts in an inheritance chain); this problem was at least part of the motivation for "spreading activation" models of memory. There are then two crucial aspects to non-motivated inheritance. The first is that inheritance results completely from the organisation of knowledge. The second is that there can be no principled limit on the number of links traversed, without adding to the mechanism. As a result, there is no obvious way of allowing for an alteration in the properties inherited as a result of a switch in relative placing as in implicit semantic attachment or goal-directed categories.

Further, the inheritance is always of what one might call "literal" properties. Since the kinds of knowledge structures are typically real-world knowledge structures, and since inheritance is grounded in relative positioning in that structure, then inheritance is of identical, untransformed properties. Exactly what I mean by this will be clearer after the discussion of analogy (Section IV.4). And since such structures are usually taken to define relations of category membership, there can be no obvious way of allowing for the perspectival-relativity of categorisation: if the inheritor does not inherit all of the defining properties of the inheritee, then the former cannot be categorised as a member of the latter.

The contrasting approach to inheritance - motivated inheritance - is one which I shall advance. In this view inheritance is motivated by informational requirements and the current thematic dimension; so it is less a passive inheritance and more an active construction of the relations that support ascriptions of properties. The mechanisms of constructing properties under thematic dimensions and implicit semantic attachment allow that discourse requirements can override general knowledge links. This is evidenced in ontological conflicts and other forms of defeasibility. The generation of a sense is not limited by the organisation of knowledge, and the ability of one concept to inherit properties from another is not defined by their prior relative placings in a structure. Rather, implicit attachment allows that we can shift placing in such a structure, even across ontological boundaries. Limiting the number of links traversed is once again the problem of constraining holism; and this is effected by the defeasible mechanisms for sense construction, which promote molecularity or local holisms. An interesting implication of this approach appears to be that it leaves much inheritance wanting a principle for inheritance. Although non-motivated inheritance gave no overt account of informational requirement, it seems that hyponymy is a reasonable basis for inheritance. If we deny that hyponymy is the only, or major source for inheritance, then what is the principle governing inheritance? My claim will be that a mechanism of analogical transfer is necessary and desirable to account for the kinds of inheritance that occur in privatives. Such a position allows that properties may be transformed in the inheritance process. For example, a stone lion inherits lion properties as represented in a statue, and not as they occur in real lions. Inheritance can be of analogues of the inheritee's properties, as well as identical properties. Further, the depth of inheritance is intrinsically constrained by the depth of analogy between the inheritor and the inheritee.

A role for analogy - a role which can find independent motivation through the idea of partiality - allows further flexibility in categorisation. Heretofore, I have been arguing that categorisation is always relative to a covering sortal (implicit or explicit), and a criterion of

indiscernability; an analogical inheritance mechanism requires a re-definition of indiscernability and indiscernable approximations, in order to allow that a partial object can approximate a less partial one not only through possessing a subset of the exact same properties of the latter, but also through possessing analogues of the properties of the latter. In such a way we allow for privative NP-type entities to be categorised as head N-type entities.

IV.2. On Defeasible Semantic Relations

In this section, I outline several different types of defeasible semantic relation, as candidates for the defeasibility that characterises concept extension. We will find, in Chapter V, that all three types have a role in the account of the problem classes.

IV.2.1. Types of Defeasibility

This typology has been modified after that of Pollock (Pollock 1986 Pollock) The use of a word or a combination results in an ascription of properties to the sense for the expression. The most important aspect of this is the effect of the combination on the central properties of the lexical concept - those which require no overt contextual motivation for their ascription to the sense, and which are crucial in determining the sortal type of any entity to which the concept applies. The following definitions of types of defeat will therefore be given in terms of the use of a noun term and the effects on the generated sense of a combining adjective. The types of defeasibility do, however, apply beyond this initial stage: for example, in accruing later content, there may be a modification of the content ascribed to a discourse referent.

P = the use, application or hearing of the N term

Q = the ascription of central properties to the lexical concept of N

R = the A lexical concept (central properties)

IV.2.1.1. Rebuttal:

if Q is an ascription of central properties to P, R is a **rebutting defeater** for this ascription iff:

R is consistent with P;

and (P & R) cannot be extended by Q; hence, $\neg Q$.

We may differentiate:

Type I Rebuttal: R is a reason for $\neg Q$ and $\neg P$;

Type II Rebuttal: R is a reason for $\neg Q$, but not $\neg P$;

Any defeasible deduction (defeat of necessary conditions) must behave after the fashion of Type I Rebuttal, since modus ponens supports modus tollens as its contrapositive. So Type I Rebuttal is the form of defeasibility which would hold for concept extensions based upon necessary properties. On the classical view, a defeasible concept is applicable if and only if we can show that *none* of the defeating conditions obtains. The (potentially open-ended) list of possible defeaters must be checked before an extension based upon a defeasible deduction can be supported. The logical possibility of doubt in another situation (i.e., the logical possibility of there being a context in which one of the defeating conditions cannot be shown to fail) is held to be equipollent with an actual, grounded doubt in this situation. Checking a list of potential defeaters is problematic, since the future ascription of new properties might include defeaters; not only would this be enormously difficult for a single concept or a straightforward concept combination, it would also be virtually impossible for any extension which did not depend solely upon an explicitly attached N type. So, if the practical possibility of defeasible deductions is problematic in the case of simple sense selection, it would lead to a computational explosion for sense generation.

Notice what Type I Rebuttal implies: that, if we attempt to combine a defeating adjective with a noun, then the use or application of the noun term is no longer permissible: P is denied. This mirrors the prediction of anomaly for apparently contradictory combinations which the classical view of concepts requires. In contrast, Type II Rebuttal, which is not open to the classical semanticist without doing great damage to the underlying epistemology, does not result in such an anomaly: the application of the noun term is applicable despite the defeat of the noun lexical concept's central-essence properties.

IV.2.1.2. Undercutting:

if Q is an ascription of central properties to P, RfR is an undercutting defeater for this ascription iff:

R is consistent with P;

and R is reason to deny that P would not be true unless Q were true;

So both P and Q may still be warranted, with independent semantic support. Undercutting attacks the *connection* between the use of a term and the ascription of properties to its concept, rather than the ascribed properties themselves. It denies that we can ascribe the central properties of the N concept to P, on the basis of the use of P in the current context. What is crucial is that there is no *a priori* reason why an undercutting should deny either P or Q outright. This contrasts with rebuttal, in which we can ascribe the negation of the defeated properties to the concept. Undercutting defeat does not deny the validity of the ascription of

the noun's central properties to the lexical concept tout court; rather, it renders them doubtful under the current state of information. If we were to gain more information, we might be able to make the ascription after all; undercutting simply puts the undercut properties "on ice", pending a later decision as to their fate. The later decision might involve a Type II Rebuttal, or an affirmation, of the undercut properties.

A standard example, deriving from the work of the later Wittgenstein, concerns pain. If we were to see a person grasping at her stomach, writhing around and screaming, we might infer that she was in pain. This would be the cause of her behaviour. But if we then noticed that she was involved in street theatre, we might revise our judgement. That is, we would undercut the inference from her behaviour to the ascription of pain. The behaviour still holds, but we do not suppose that it indicates pain. But we may then see someone running away with a bloody knife, and indeed notice that the other actors in the theatre have ceased acting and are screaming for help. In this case, we might then revise our previous revision, undercutting our inference from the behaviour in the context of the theatre to the ascription of acting/feigning pain. And so on. Pollock gives another example of undercutting. If we were to see a ball that looks red, under normal conditions we could ascribe the property of being red to the ball. But suppose that we know we are viewing the ball under red lighting conditions; then we cannot conclusively decide whether the ball is red or not, since all objects - including red ones - look red under such conditions. We must then undercut the ascription of redness, and await the outcome of further inquiry.

Whereas the classical realist notion of an entailment cannot allow for undercutting, this type of defeat is the very one that underpins C-relations: the kind of doubt inherent in undercutting is just that which attends upon anti-realist semantics (see Chapter I). My argument is that the major relation for property ascriptions to a concept, and hence for concept extension, may be characterised through C-relations (that is, the same kind of relation which holds between Criteria and the application of a concept). So an ascription made on the basis of C-relations carries certainty with respect to the ascription of properties, at least *ceteris paribus*. And this certainty holds in the full cognisance of the fact that the ascription might be defeated if some other concept combines with the extended concept as a defeater. A C-related extension, then, bestows certainty on the property-ascription, even though we may subsequently undercut that ascription.

IV.2.1.3. Defeat of Inductive Properties:

The defeat of an inductive property ascribed to a sense does not imply a defeat of the whole concept extension. These properties are either properties or values of properties

related to the concept by an inductive relation: they are "typical" or "characteristic" properties. In a concept combination, a N term's inductive properties may be defeated by the A's inductive properties, which have greater priority.

Symptoms are inductive properties supporting concept application. Hence, S-relations are relations which possess the same qualities as the relation between the Symptoms and the application of a concept. A lexical concept may be extended through C-relations and/or S-relations: the latter may be ascribed to a concept even where the former have been rebutted or undercut. This can be phrased as an extension through the ascription of non-central properties. In such a case, we might nonetheless be able to categorise an entity which has only the S-related properties as of the same type it would be if it also had the C-related properties. But there will be a crucial difference: Symptoms are not normally sufficient conditions on concept application; hence, an ascription based on S-relations, in contrast to one based on C-relations, will not support a fully generalisable categorisation of NP-type referents as head N-type referents. An S-related extension will facilitate such a categorisation only with respect to the limited range of S-related N properties ascribed to the sense: NP-type objects will be N-type objects only from the limited perspective of those S-relations.

IV.2.2. C-relations, S-relations, Defeat and Molecularity

If C-relations and S-relations are independent, then central and non-central properties are independent in sense generation. So the defeat of central properties does not necessarily imply defeat of non-central properties, and vice versa. The relation between the two types of central properties is a little more complex. If a lexical concept's central-essence properties are defeated this need not imply defeat of central-diagnostic properties. We know that objects can have the identifying or diagnostic properties of a certain type, even though they do not have the appropriate central-essence properties: robots and fakes. However, when central-diagnostic properties are defeated, this renders the central-essence properties at least doubtful. For example, if someone has told us an object is a lemon, but we know it is square and orange coloured, we might doubt that it is in fact a lemon, even though we could not state decisively that it was not. A rebuttal of central-diagnostic properties, then, may lead to an undercutting of central-essence properties.

The behaviour of defeat depends not only on the type of defeat, but also upon the type of network in which the concepts inhere. The interaction of defeat and knowledge structures will be influenced by the particularities of content. In order to reduce complexity, and

enable some general comments to be made, I will comment on hard and soft molecular structures only. And I will not distinguish between the defeat of the different types of central properties. Since the ascription of properties to senses may involve more primitive as well as complex properties, in this section I will use "concept" and "property" interchangeably.

Soft molecularity results in local holisms: the associated C-relations are same-level canonical links. This means that two concepts are interdefined, so that one could not be ascribed to a sense without the other. How a defeat affects such a network depends upon the type of local holism. We might distinguish two types.

If we note a C-relation between X and Y as $X \Rightarrow_c Y$, then the following is a *single-loop* local holism:

$$A \Rightarrow_c B \Rightarrow_c A$$

as is:

$$A \Rightarrow_c B \Rightarrow_c C \Rightarrow_c D \Rightarrow_c A$$

In both cases a lexical concept A has canonical links with one or more other concepts, whose own links involve A. In the case of *multiple-loop* local holisms, there is more than one loop of canonical links. For example:

$$A \Rightarrow_c B \Rightarrow_c C \Rightarrow_c D \Rightarrow_c A$$

and:

$$A \Rightarrow_c W \Rightarrow_c X \Rightarrow_c Y \Rightarrow_c A$$

Each of the loops is independent of the other: so the effects of ascription or defeat on each loop are independently assessed.

Consider, first, Type II Rebuttal. For a single-loop local holism, we might have:

$$A \text{ Rebut}(\Rightarrow_c) B \Rightarrow_c C, \dots, \Rightarrow_c A$$

Here, the rebuttal of B effectively "blocks" the links to the other concepts in the loop. In essence, the effect is to rebut all of the elements in the loop which occur after the defeated link. These properties cannot be ascribed to the concept extension of the lexical concept A. This will not defeat A, since it is not a Type I Rebuttal: the fate of A will depend upon the relation between A and the defeater of B. When rebuttal occurs in a multiple-loop local holism, only the particular loop which contains the defeated concept will be affected: the subsequent concepts will be "blocked", whilst the other loops will be assessed independently.

Undercutting in a local holism contrasts with rebuttal. Here, the C-related concepts are not

denied: their connection in the current context is severed. We might view this as a "temporary blocking" of the further links in the loop: with increased information, the ascription may be permissible.

The behaviour of defeat in hard molecularity resembles that in multiple-loop local holisms, with two important differences. Firstly, the C-relation links are all lower-level, and hence can obtain in the absence of A; secondly, the branches do not loop: they do not terminate in the starting concept (i.e., A). The first point is important with respect to rebuttal. Although a lower-level (central) B is rebutted, the ascription of this property is not necessary for the ascription of B's lower-level daughters. This contrasts with the local holism case, where the ability to establish the links with other same-level concepts appears to require that A's central properties be ascribed to the sense. In a hard molecularity, then, lower-level C-related properties might be ascribed after the defeat the central properties of A.

Undercutting in hard molecularity also makes the ability to reinstate a lower-level canonical link after undercutting less problematic. The S-relation links can motivate a re-establishment of the link under a perspective, since they can be ascribed without the central properties.

If the behaviour of defeat is sensitive to the type of knowledge framework in which the defeated relation occurs, how then do we discern the type of framework? All questions of this kind are complicated by the details and complexities of particular domains of content. The answer is reasonably straightforward for dialectical definitions, in which A is C-related to B and vice versa; here, we can examine the candidate contents, and ascertain their relationship. For a local holism, it is not possible for either of the concepts to hold without the other. In essence, we require a contentful biconditional relation. On this kind of test, then, we might consider *lion* as follows. The property "animate" does not induce a local holism, since, although *lion* is C-related to "animate", the reverse is not the case: something can be animate without being a lion. However, the position for other central-essence properties might induce a local holism: *lion* is C-related to "lion presumed essence" or "lion presumed ontogeny", and vice versa. There are at least two difficulties with this as a general method. The first is that it can work only for dialectical definitions: where the loop consists of more than two concepts, the position becomes somewhat more complicated. We would then need to ascertain whether the links did loop or not. For example, although concepts A and B might not be looped together, B could be linked to an arbitrarily long chain of concepts that eventually looped back into A. And this leads to the second problem; to ascertain

whether the links do loop, we would need an exhaustive check on all of the central properties. This is problematic enough for any sense generation approach; but it is inappropriate for the current view, which does not seek exhaustively specify conceptual content. This view aims to indicate some ways we might individuate different kinds of content, within and between concepts. A case for restricting the distance of connections to be examined might be made out, based upon the Accessibility and Contrastivity Constraints (see I.3.3.2), but would be inapt to the current study. The situation would be more problematic if we were dealing with subtle forms of combination; the problem classes do, however, appear to involve interactions on a rather coarse level. So we do not have to trace the intricacies of conceptual links. My heuristic, then, within this method, is to consider only the relations between the concept and those properties which are clearly defeated. In essence, certain central-essence properties will be in a local holism, whilst others, and the central-diagnostic properties, will be in a hard molecularity. This means that the method proposed above will suffice; and since neither of these structures pose problems for the operation of defeat and reinstatement, I will not comment on them in detail in my account.

The interaction of defeasibility and local holisms is particularly important for promoting constraints on holism. The defeat of links is an important constraint on property inheritance. As presented in this section, however, property inheritance appears to be non-motivated, depending upon the placing of a lexical concept in relation to others in a knowledge-structure. Through implicit semantic attachment and analogy this can be seen as only a part of a motivated inheritance viewpoint.

IV.4. On Perspectives

In this section, I outline what I take to be a perspective.

IV.4.1. What is a Perspective?

The conception of a sense defining a perspective on an entity can be traced to Frege: his "sense" was a "mode of presentation" of a referent. On his view, there could be more than one sense for a single referent, though never more than one referent for a single sense. For example, consider:

Samuel Beckett = The author of "Ill Seen Ill Said"

Both expressions have the same referent, but they have, on Frege's terms, different "cognitive values". This notion of "cognitive value" should not be misconstrued: it is an aspect of sense, which is wholly publicly specifiable. Similar views are found in recent work on the

psychology of concepts - especially by Barsalou and his co-workers (see Chapter II). Similarly, Heim's (1979) notion of objects in "guises", and Landman's idea that a single peg can play many different roles in a discourse, make use of the same intuition. My own contribution is aimed at adding further flexibility, in line with the Manifestation Constraint.

As I have already mentioned, I take adopting perspective on an entity to involve associating a description with that entity. And I have argued that we can isolate two components to such a perspective - the (often implicit) covering sortal type, and the criterion of indiscernability. Another way of construing this is that a perspective is the explanation, or explainability, of the presence or ascription of a set of properties by the inference to a more central set of properties which latter supports a categorisation of the entity. It is the location or explanation of a state of partial information with respect to a less partial state. That is, we explain the properties incorporated in the criterion of indiscernability by construing them as inhering in a particular type of referent, the knowledge associated with which expresses a less partial state of information. A perspective relates conditions of application of a concept to the conceptual content; it relates our information to a referent. It is thus crucially related to the interaction of an agent and her environment; it does not reside wholly in either.

Another way of viewing perspectives is as a further means for flexibility in interpretation. Recall, from Chapter II, the components of a discourse scenario. There, the thematic dimension and the perspective were said to be two of the ways in which to provide for flexibility; the former received some discussion in the last chapter. The idea here is that, even where two agents share a viewpoint on a situation, and have the same thematic interest, they may nonetheless adopt different perspectives on the same entity. This they do by associating different descriptions with the entity - by categorising it differently.

So the flexibility of perspectives comes down to the flexibility in the kinds of properties which can be included in a description, and the range of choice for a covering sortal. In this chapter, I have argued that the various types of conceptual content which we might distinguish may be independent. For example, S-relations are independent of C-relations, and the two groups of properties which come under C-relations are at least one-way independent. This means that our descriptions can contain properties from all or any of these types. For example, even if central properties have been defeated, we can still ascribe non-central ones to a sense. And similarly, the defeat of central-essence properties need not imply the denial of central-diagnostic ones. Leading on from this, and the relativity of categorisation, I should like to make a distinction between two extreme types of perspective. I shall do this by assuming that the covering sortal is constant, and focussing on the criterion of

indiscernability. The first is a

Type I Perspective. This is a categorisation based upon central-essence property ascriptions, plus a subset of central-diagnostic or non-central properties. Under this type of perspective, a NP-type referent can be said to be a head N-type referent in the general sense, even if we were to focus on different aspects of the NP. An example would be a categorisation of an affirmative NP object by its immediate superordinate N: a red apple is an apple in that the central properties of the latter subsume those of the former. Hence a Type I Perspective is implicit in the AC categorisation schema noted in Chapter II. And secondly we might have a

Type II Perspective. This is a restricted, limited-perspective categorisation, comprising a defeat of central-essence properties, and ascription of either central-diagnostic or non-central properties, and does not allow a generalisation beyond the current thematic concern indicated by those properties. For example, in Privatives, a fake gun is a gun with respect to shape or appearance only; if we changed our thematic interest, the categorisation might not hold. This kind of perspective is implicit in the PC categorisation schema.

The criterion of indiscernability can be said to limit the range of contexts or standpoints within which an entity categorised as NP-type can be categorised as head N-type. The two types are best viewed as being the polar points of a continuum of generalisability of categorisation judgements. A categorisation on a Type I Perspective should be generalisable to any other context, whilst the Type II should not. Between the two extremes there could be as many different criteria of indiscernability as there are ways of cutting up and combining conceptual content.

Generating a sense results in the construction of a description which we associate with an entity of an appropriate type. If we were then to employ this sense in the categorisation of other entities, or to categorise the same entity in different ways, then, combining the sense (the criterion of indiscernability) with the categorising sortal, we are adopting a perspective on the to-be-categorised entity. We might say that the sense itself defines a perspective on the entity in question. But the sense can also, as part of the discourse content, *support* many different perspectives on the NP-type entity. That is, if the NP sense defines the properties for NP type entities in this context, then this sense provides the space of possibilities within which other perspectives on the NP entity act as specifications. So the sense of the NP constrains permissible categorisation perspectives on the NP entity. Two questions then emerge. Firstly, if the sense does constrain such permissible perspectives, then how is such a sense arrived at? Clearly, this is the problematic of the current project, and an approach will be sketched in Chapter V. Secondly, once we have arrived at such a sense,

how exactly does it constrain the perspectives? That is, what is the constraint? We will be in a position to provide a tentative answer to this question at the end of this chapter.

The contrast between the two types of perspective was couched in terms of differences in "generalisability" of categorisation. What does this amount to? It is simply the question of whether we can maintain the same categorisation of the entity (i.e., the same overt categorising N, not the covering sortal) at other and deeper criteria of indiscernability (i.e., more central properties). Given the breakdown of the discourse scenario outlined in Chapter II, we can consider the generalisability (or relativity) of the categorisation with respect to different thematic dimensions, to different viewpoints, and to different orientations. A Type II Perspectival categorisation is one which is warranted only for the current criterion of indiscernability, under the current thematic dimension, and the current viewpoint. That is, the criterion comprises, minimally, a set of properties which, under the current thematic dimension, is just sufficient to categorise the entity as of the N type. We may or may not be able to support the same categorisation under a different thematic dimension. Although a statue of a lion can be categorised as a lion with respect to its shape, there is no reason why the same categorisation should hold on the basis of typical habitat, colour, function, behaviour, or dietary habits. In contrast, a Type I Perspectival categorisation is generalisable: it is based upon the central-essence properties of the categorising N type, plus a thematically-relevant selection of other properties. If the thematic dimension were altered, there is no reason why the central-essence properties should not hold. So a Type I Perspective can be maintained through all of the thematic dimensions which relate to the entity under the current viewpoint. And there is no reason why it should not be maintainable across different viewpoints also, as long as the entity is part of the focus in the scenario. Whether or not it can be maintained across other orientations will depend upon the content involved. I think that the kinds of perspectives envisaged by Barsalou and Frege, are Type I Perspectives. And it will clearly be the only kind of perspective open to adherents to a classical realist view of identity.

There are several implications of this view. The first is that adopting a perspective on an entity does not in any way imply that we believe that the categorisation could be generalised to broader perspectives. The criterion can be as slight as the flow of information in a discourse requires. And if the generated sense for a NP constrains permissible perspectives, the latter must be sensitive to the degree of flexibility and specificity occurring in the discourse and required by the agents. The second implication is that we might adopt two different perspectives defining apparently contradictory categorisations of the entity, and yet be warranted in both. By taking into account the criteria of indiscernability used to make

the categorisations, the apparent contradiction may evaporate: a stone lion is a lion under a very narrow (Type II) criterion, but a statue under a very broad (Type I) criterion. Conflict would only arise if the incompatible categorisations were both intended to be of Type I.

IV.4.2. A Problem and a Reply

An argument which might be levelled against this view embodies an attack on the possibility of Type II Perspectives. The argument runs like this. A stone lion is a lion to the extent that someone represents it as such. Thus, *any* entity could be labelled as a lion if we first decide that this is the case. For example, I may decide that a cup is a lion (in a circumstance such as playing a game with a child), and a tall tumbler is a giraffe. I can then say, for example, "let the cup handle be the lion's head", and so forth. I can then re-interpret the properties of the cup (the N) in terms of those of the lion (the NI). As long as I draw an analogy between the N and the NI, I can label the referent of the N as a member of the category NI.

What would this "representational" position imply? The first point is that no two people need agree on labelling the N referent as an NI. Suppose I decide that a cup is a lion; suppose, further, that a friend of mine calls and we decide to take some tea. When he first looks at the particular "lion" cup, there is no reason why he should categorise that cup as a lion. That is, the diagnostic properties for the cup are completely dissimilar to those for a lion, and there is nothing intrinsic to the cup, or the everyday range of interactions with it, to prompt the construction of an analogy between the cup and a lion. Until he enters the particular discourse or language-game surrounding the representation of the cup as a lion, there is no reason why he should make the same categorisation of it. Notice that this "representation" condition is a "goal-directed" categorisation: I categorise the cup as a lion with respect to the requirements of the particular language/imagination-game I am playing. It is a categorisation on the basis of an implicitly attached N, whose attachment stems solely from the current game or discourse goal. In a slightly different game, the cup might "be" a giraffe, and an egg-cup, a lion.

This indicates the central difference between the "representation"-based implicit NI categorisation, and the Type II Perspectival categorisation of, say, a stone lion. Under the Type II Perspective, the NP referent just does share *some* properties with the head N: a stone lion has the shape of a lion. So in any unprimed individuation condition, we would expect an agent to be able to individuate that referent in two different ways. Firstly, she might categorise it as a statue (a member of the NI under a Type I Perspective); secondly, she might assert that it is a statue *of* a lion: it is a lion with respect to its shape. And this is the

point: rather than having to *construct* the analogy between the statue and a lion, the source of the analogy is already evident. That is, the statue objectively has the properties that form an analogy with a lion - and these properties can be used in any unprimed individuation situation to categorise the referent of "stone lion" as a lion (a Type II Perspective).

A counter-argument might be made out as follows. A statue is itself a "representation": a statue of a lion involves the intention of the sculptor to make it represent a lion. And this seems identical to the "lion-cup" case mentioned above. I think not. As a representation, the statue should conform to a cultural norm for a representation, perhaps located in an appropriate lay theory. This has two aspects. Unlike the "lion-cup" case, we do not need to be primed in order to see the statue as a representation: the fact that statues, even the abstractions of Henry Moore, *do* represent something is a central-essence property of statues. An unprimed agent would not need to be told to examine the possibility of there being an analogy between the statue and another entity. The second aspect is that these very cultural norms of representation dictate that there be some analogy between the representation and represented. The analogy may become more tenuous or harder to locate the more abstract the representation (a continuum between Rodin and Barbara Hepworth, for example); but it is still there. Of course, this disregards other possibilities such as an artist's desire to flout these conventions; but even in this case, that it is a representation is not often in doubt. A third point is that we can envisage contexts in which a stone lion is categorised as a lion even where the former is not a statue or any other attempt at representing a lion. For example, a stone lion could just be a peculiar rock-formation; again, we should expect a Type II Perspectival categorisation of such a rock-formation in an unprimed situation.

The argument, then, is that the Type II Perspective is based upon properties which are publicly observable, information which is shareable without special priming. So the response to the "representation" argument is: yes, the categorisation of the NP as an N may depend upon a representation; but this representation must be from the appropriate, objective perspective.

Further, a representational context will set up a *belief-* or *opaque context*: and these are precisely those contexts into which we cannot "quantify in" under any circumstances. That is, there could be no circumstances under which we would be justified in asserting that "X is a lion", after being told that "X is a stone lion", since the latter construction sets up an opaque context. This is the effect of Jackendoff's assimilation of belief contexts to representation contexts (Jackendoff 1983). But this is the precisely the question at issue:

whether we can, in fact, quantify into a privative context under any circumstances at all. And my claim is that we can: we can do so under restricted Type II Perspectives. Here, the identity statement resulting from existential instantiation is qualified by the description. This conception of "limited quantification in", or of the combination's setting up a "quasi-opaque context" resembles Quine's "quasi-quotation" (Quine 1951). A Perspective, then, although mentally representable (and comprising manifestable knowledge), is not solely a mental representation.

IV.5. On Analogy

In this section, I sketch some central aspects of current conceptions of analogy. I then outline some of the motivations for proposing a role for analogy, including the interaction of partiality and analogy. I then comment on the possible role of analogy in concept combination.

IV.5.1. Central Aspects of Analogy

When an analogy between two entities or problem domains is constructed, knowledge concerning one (the *base*) is used to comprehend the second (the *target*). Given a problem situation (in the current study, the appropriate noun type categorisation of a referent of NP type) we identify the target (the NP) and access a *base* situation (in this case, the NI concept, deriving from the instantiation of the NP). The agent then constructs a *mapping* from the base to the target, which facilitates inferences being made about the target (thus, effecting a *transfer* of properties from the base to the target). Mapping thus involves a matching of the base and the target. Finally, she must then judge the soundness of the analogy, and the relevance, generality and validity with respect to the target, of the various possible inferences. The general principle is that, the deeper the analogy between the base and target, the greater is the potential for transfer, as empirically substantiated by Clement & Gentner (Clement and Gentner 1988).

Thagard & Holyoak identify three constraints on this process, each of which may play a different role in the different stages of the construction of analogies (Thagard and Holyoak 1988). The first concerns the *semantics* of the base and target: do they have the same properties, in the same relations to each other? Can we make the same predications of properties to the referents? The second concerns the *structural* parallels between base and target: these aspects focus on, for example, structural consistency. For example, Gentner (Gentner 1983) proposes that individuals observe the structural constraint of one-to-one

correspondence of objects, and consistency in over-all structure. Such causal-structural and systematic analogies contrast with the basic semantic ones deriving from the sharing of "mere appearance" properties and/or relations. The third constraint concerns the *pragmatics* of the mapping: for example, do the structures and properties have the same purpose in the context of their respective referents?

Crucial debates in the area have concerned the relative import of these constraints at the different stages of analogy construction. The access stage is strongly influenced by semantic factors - the general similarity, often only at a relatively superficial level, between the target and a base is sufficient for the access of that base. This has been supported by empirical evidence (Skorstadt, Falkenheimer and Gentner 1987). And although Burstein & Adelson (Burstein and Adelson 1987) and Thagard & Holyoak (op. cit.) argue that all three types of constraint have a role in access, they agree that the major influence is the semantic.

The main disagreement concerns the role of non-structural constraints in the mapping stage. Gentner claims that the only constraint operating at this stage is the structural one: all possible structural analogies between base and target are generated by the Structure-Mapping Engine (SME), which is a model of her structure-mapping theory of analogy. The validity of the inferences in the target, and the contextual relevance of those inferences are held to be evaluated by separate processes. Evidence to support the role of systematicity as a constraint on selection has been provided by Clement & Gentner (Clement and Gentner 1988). However, this does not show either that structural constraints are the only ones, nor that they might not be superseded by pragmatic constraints, such as goal-relevance. Thagard & Holyoak have argued that, although the structural constraint may be the strongest, it is not the only one: semantic correspondences relating predicates with similar meanings also play an important role. Burstein & Adelson (op. cit.) have shown that goal-relevance may constrain the selection of options from a range of structural correspondences, and Holyoak has argued that systematicity is usually correlated with goal relevance in real cases: what is construed as a systematic relation depends upon the current goal (Holyoak 1985) So the purpose of the analogy influences what is selected to map from the base to the target, and hence what constitutes a match in the current context.

The question of whether the selection and mapping is influenced by non-structural factors has another facet. This is whether the search for a mapping should be done by an essentially "stupid" or "intelligent" process. The assumption that only structural factors are of relevance is tantamount to requiring a "stupid" mapper, in which *all* potential

correspondences are mapped. In effect, then, the analogy to be used for the current purpose would then be a result of a post-mapping selection, as is the case for Gentner's SME. If we allow the mapper to be "intelligent", it can, as it were, look outside of the base and target to the goal, context, and informational requirements. The latter view is more consonant with the Sense Generation orientation.

The final aspect is the transfer of properties of the base to the target - facilitating inferences regarding the target which were not part of the mapping, but which derive from the base. This is, according to Thagard & Holyoak, influenced by all three constraints, although the critical one will be pragmatic. That is, only the kinds of content which are relevant to the solution of the current problem, are transferred.

So it seems that the goal of the analogy is crucial to all stages, and that structural correspondences are important in the mapping stage. This conclusion concurs with the argument of Thagard & Holyoak. An important possibility, evidenced by Burstein, is that analogies can compare entities at several different levels of description. Familiar entities or domains, he claims, can be used as the base for analogies at different levels of abstraction. For example, locating an analogy in terms of topology or appearance might facilitate a correspondence between behavioural systems or functional qualities, which might then be used to induce an appropriate causal structure in the target, by analogy with that of the base (Burstein 1988).

A further issue concerns whether the match should be a total or only a partial one. On all accounts, the match will be less than total. This has two aspects. Firstly, an analogy involves a selection of some properties from the base, upon which a mapping is constructed. This might involve emphasising structural or semantic properties. So the mapping is partial in the sense that a proper subset of both entities' qualities will be involved. But it is partial for a more important second reason. The mapping itself is partial: that is, within the range of attributes which are used to effect the mapping, there will be an imperfect correspondence between the base and the target. For example, there will be differences in some of the properties, their values, or relations.

IV.4.2. The Motivation for Analogy

A major motivation for proffering an analogical property inheritance mechanism is the similarity between the underpinnings of partiality and analogy. Following a consideration of this, I will mention some more general motivations.

IV.5.2.1. Analogy and Partiality:

The construction of an analogy between two entities, and the ensuing analogical transfer, embodies a partial informational link between the base and the target, and an extension of the latter based upon partial information. I mentioned the partiality of matching in the last section. Transfer is partial in two senses. Firstly, it is partial because it is selective: only properties which are relevant to the goal are carried over. And these properties are transformed according to the conditions of the analogy. Secondly, it is partial because it is wholly dependent upon the level of analogy: we may not have sufficient warrant, for example, for the transfer of central-essence properties, if we have a slight analogy. The stronger and deeper the analogy, the more central will be the properties that are transferred.

Another way of phrasing this is that, the stronger the analogy, the less partial is the target with respect to the base. That is, forging an analogy between the base and the target is an attempt to construe the target as an indiscernable approximation to the base. So, for example, with respect to the NP referent, the base (the NI concept) will be a less partial approximation than the target (the NP sense). And by effecting an analogical transfer, the aim is to render the target a better approximation to that referent. So, and crucially, a partial object may be partial in the sense that it bears only an analogy to a less partial object. This both broadens and makes more precise the conception of partiality with which most researchers have worked. In order to see this, I shall look at some notions espoused by Landman (op. cit.). In giving an account of the identity conditions of alecs (see Chapter I), Landman argues that part of the requirement for two alecs' being the same peg (so the second can be a "role" of the first) is that the first is an "indiscernable approximation" of the second. On my view, this is the requirement that all properties that can and cannot be predicated of the first alec must bear the same relation to the second alec; and all central-essence and central-diagnostic properties of the first, must be found in the second. If this condition fails, the second alec cannot be a role of the first. But this seems too restrictive: there is just no sense in which it could admit that an alec which is analogous to another could be a partial approximation. The mapping must be based only upon identity, both of properties and relations: it disallows a role for analogy, in which the properties and relations are similar or transformed. But on Landman's view, if we can assert that one entity is a partial approximation of the other, then we can claim that the two could be identical (or at least indiscernable approximations) as long as the growth of the properties of the former entity does not bring it into conflict with the properties of the latter. In essence, then, we can effect a "transfer" of the properties of the less partial entity to the more partial one. And this parallels the analogical transfer mechanism. The central difference is in the account of the matching of the properties of the target and the base.

Can we claim that the "analogy" approach improves upon the "bare" notion of partiality? I think we can, for several reasons. Landman's notion of "content" or "properties" is relatively undifferentiated. He equates an increase in information with an addition of properties: a target object is less partial with respect to a base if we add to the former a single property of the base which it previously did not have. So a target may become less partial with respect to a base only by accruing more of the exact properties of the base. It is partial with respect to the base simply because it has fewer such properties. Analogy can enrich this view substantially. Here, a target can be a partial approximation to a base even if it does not share any of the exact same properties: if it bears an analogy to that base. So the notion of an approximation is no longer simply a matter of the cardinality of properties. It is now expanded to allow for analogies: of course, the deeper the analogy, the more likely it is that the target will share identical properties with the base. And of course, we could elaborate upon the analogy up to a point where the target and base were effectively identical - indiscernable approximations in terms of properties and analogies of all levels. This seems to be a sharpening of the notion of an approximation to an entity based upon partial information, in line with the Manifestation Constraint. And with the possibility of different analogies comes the possibility of differentiating between types of content. We should, of course, preserve the distinctions already made: between central and non-central properties. But the addition of analogy allows a further distinction: between *analogies* which are of central and non-central level. That is, analogies of aspects of the target and base which are of central and non-central level in the entities. So we now have the possibility of approximation founded upon both analogical and literal similarity. And we have also the possibility of a role for analogy in categorisation, and thus in the definition of an indiscernable approximation. It is in this sense that an addition of considerations from analogy can broaden our conception of partiality.

IV.4.2.2. Other Motivations for Analogy:

There are several other motivations for a crucial role for analogy in concept combination. Firstly, consider again Type II Perspectives. A crucial question here concerns what kinds of properties can and must be brought under such perspectives. If the properties are to support unprimed categorisation, then they must not be completely peripheral, nor highly abstract. For example, to categorise something as a lion in virtue of its having a nose would be odd; as would to do so on the basis of its being a physical object. (In both cases, a bizarre scenario could be constructed so that these very properties could be individuating; but this need not detain us). The implication is that the criterion of indiscernability for a Type II Perspective must comprise properties in virtue of which the categorised entity (the NP-type entity) bears at least a superficial analogy to the categorising type (the N type). That is, the

NP-type entity could be categorised as being N-type, using central-diagnostic or non-central properties alone. So a fake gun can be a gun, feigned pain can be pain, a stone lion a lion.

A related point concerns the manifestation of head N properties in the NP-type entity. A statue of a lion and a toy gun both manifest the diagnostic and/or non-central properties of their head N, but in a manner which respects their sortal type. The former's *lion* properties are transformed after the fashion of a statue; and the latter's *gun* properties are transformed for a toy and not a replica. So they bear only an analogy to the head N, as a result of a transformation on the basis of the implicitly attached nouns of *toy* and *statue*, respectively.

Another point, connected to the above, concerns the nature of properties. They are themselves, by and large, not irreducibly simple: they reflect some prior categorisation. And this means that they reflect the indiscernability of certain distinctions under an appropriate criterion of indiscernability. For example, the property of having a "leg" might cover anything from human, to horse, to chair legs. And this simply means that the definition of a property will be such as to allow both literal and analogical parallels.

Another important point concerns the parallel between the access of a base, and the instantiation of a combination. An entity will be accessed for use as a base if it bears a superficial analogy to the target. As was argued in Chapter II, the instantiation of a NP is a more precise manifestation of the NP; it must therefore have, at minimum, a surface semantic similarity to the NP. And the concept associated with the instantiation is then a source for the inheritance of properties by the NP. So the target is the NP combination, and the base is the instantiation concept. Now, where the instantiation is implicitly attached, there may be a range of candidate instantiations: for example, the instantiation for *fake wine* might be anything from a de-alcoholised wine, through a "wine" made from peaches, to one made from entirely synthetic ingredients. Each of these possibilities will bear a surface similarity to the *fake wine* combination (a fake wine being something which can pass muster as wine whilst not actually being wine: having sufficient non-central-essence properties of wine). The crucial factors are then the thematic and functional aspects of the candidate in the discourse scenario. If the candidate matches the current thematic interest, and can fulfill the roles and functions of the described entity in the discourse scenario, then it will be selected in preference to the others. In essence, its similarity to the combination is correspondingly upgraded. And this is precisely the implication of the analogy studies: the semantic similarity between base and target is heavily influenced by pragmatic and goal-related factors.

Further, there are some interesting issues with regard to transfer. In a concept combination, the extent of transfer can be read as the extent to which the concept combination inherits properties from the instantiation. And a corollary of this is that the extent of transfer determines the criterion of indiscernability under which we can categorise a NP-type referent as of the base N-type. Clement & Gentner (op. cit.) found that the deeper the analogy, the greater the transfer. Hence, the wider the range of perspectives from which the categorisation could be made. And this is simply the obverse of the point noted earlier, that the fewer the properties which the NP referent shares with the N-type, the greater the extent to which the perspective of the categorisation must be qualified by contextual and restricting conditions. And by contrast, the more shallow the analogy, the less the transfer. On current terms, we can interpret the Clement & Gentner point as follows. The properties which are transferred from the base to the target will always be of either the same level as, or a more "shallow" level than, the analogy. So, for example, an analogy at central-essence level will support transfer at all levels, whereas one at non-central level will support transfer at that level only.

This has one final, important implication. It is that the inheritance of properties does not take place in a vacuum; it is, rather, motivated by the analogical mapping, which is itself constrained by the pragmatic-contextual aspects of the combination. And the choice of base (the inheritee) is not determined by the prior placing of the concepts in a knowledge structure. Transfer is of properties which do not form part of the mapping, and such properties may be transformed according to the terms of the analogy. Transformation may involve either a more detailed specification of properties or an alteration of the properties themselves.

IV.5. On Relativity, Perspectives and Partiality

The argument of the last section has two important implications for the proposal that lexical concepts are to be viewed as pegs. The first concerns the nature of concept extension; the second, the definitions of indiscernability and alocs. Concerning the former, the analogical property inheritance mechanism is a central part of concept extension; transfer extends the lexical concept, ensuring that it becomes a closer approximation to the instantiation. Concerning the latter, we will need to take a closer look at Landman's definitions, and propose some alterations.

If the lexical concept is a peg, so then is the extended concept, the sense of a word on the current occasion of use. Similarly for the senses of NP's. I have already mentioned the

possibility of categorising a NP-type entity as a member of different N types, which depends upon the sense as a constraint on categorisation criteria. So it is that, after having generated a sense for the NP, the sense description comprises the properties of a discourse referent. This discourse referent can then play the role of an NP-type entity in the discourse scenario. It can also, however, play different roles. A red apple can play the role of an apple, a fruit, a missile or a cricket ball. Each of these roles involve the NP-type discourse referent being categorised as a different type. The ability of the NP-type referent to play such different roles is constrained by the properties of that referent, in the same way that different categorisations are constrained. We can then say that different roles amount to different categorisations of the entity, so the roles can be considered in terms of the perspectives under which the categorisation takes place. In Landman's terms, pegs which play such different roles are "alecs", and the different perspectives will be the different "roles" of the Alec.

Let us look again at the definition of an Alec. Recall that different "branches" of information-growth are different ways of accruing the same information: they have the same starting-point and the same (total) end-point. Landman's definition of an Alec, then, is:

Alec(a, Y, s): peg a is an Alec with respect to a set of properties Y in s iff:

for every branch b and every peg d which at some stage of that branch has the properties in Y, there is an alternative b' for b where at some stage a has the properties in Y and a is an indiscernable approximation of d.

where

Y = the properties ascribed to a peg

s = the discourse situation

a is an **indiscernable approximation** of d in s iff for all properties P:

if s \Vdash P(a) then s \Vdash P(d), and if s $\dashv\vdash$ P(a) then s $\dashv\vdash$ P(d); and

if s \Vdash must P(a) then s \Vdash must P(d), and if s \Vdash may P(a) then s \Vdash may P(d).

We may read the turnstiles, \Vdash and $\dashv\vdash$, as "warrant for ascription" and "no warrant for ascription"; and the modals, **must**, as the ascription of **central-essence** properties, and **may** as the ascription of **central-diagnostic** and non-central properties. So alecs must subsume their roles: the latter must have all of the former's properties plus their own additional "role" properties. In this way, Landman ensures that an Alec can retain its identity, and play different roles. This is what we require for a discourse object to be able to be viewed in different ways in a discourse. However we choose to view the object after the construction of the NP, it will nonetheless still retain the NP sense properties (be a member of the NP type), as well as whatever other properties are ascribed to it. This reflects our ability to

categorise a stone lion as a lion in the full knowledge that its central-essence properties are those of a statue; to make a judgement on the basis of one perspective does not mean that we would make the same one from other perspectives. Although we choose not to focus on the central-essence properties of the entity, this need not mean that we are unaware of them. A shift in perspective might bring those properties into focus.

However, in the current project, we are interested in the ability to categorise the *peg*/NP object as a member of different N types. We need, then, to consider the relation between an *alec* and its roles, not as regards a role's retaining the identity of the *alec*, but as regards categorising the *alec* as a member of the role type. The standard form of categorisation is that the categorising type subsumes the categorised entity: the role N subsumes, or is an indiscernable approximation to, the NP, so that the required properties of the sortal N are embedded in the NP concept. We need, then, to define some conditions for categorisation to operate over *alecs* and their roles, as defined by Landman.

The first stage is to alter the definition of an indiscernable approximation, to take account of the role of analogy. This requires a second clause to the definition:

a is an indiscernable approximation of d in s iff for all properties P:

- (i). if $s \Vdash P(a)$ then $s \Vdash P(d)$, and if $s \dashv\vdash P(a)$ then $s \dashv\vdash P(d)$; and
if $s \Vdash \text{must } P(a)$ then $s \Vdash \text{must } P(d)$, and if $s \Vdash \text{may } P(a)$ then $s \Vdash \text{may } P(d)$;
and:
- (ii). for any or all of the properties in (i), d may possess analogues of the properties in a.

The second clause is intended to reflect the idea that many apparently "identical" properties are simply analogues when viewed at a finer level of detail. For example, no two lions' hearts are truly identical. It is not intended to imply that we can tolerate wide degrees of analogical difference in central-essence properties: a lion with a real heart and one with a mechanical heart might be deemed identical, but one with absolutely no organic inner mechanisms would be judged a non-lion. This implies that the permitted degree of analogical latitude may be greater for central-diagnostic and non-central properties than for central-essence ones. However, this is very much dependent upon the particularities of contents; the definition of an indiscernable approximation, then, though imprecise, will stand as it is. If we construe the set of properties, Y, in the above definition of an *alec* as the properties ascribed to the NP, then we can define the following categorisation schemas:

AC': *peg* a can be categorised as a member of the noun type of *peg* d iff

a is an *alec* with role d, such that d is an indiscernable approximation to a under

some criterion c , where $c \subseteq Y$.

PC': *peg a can be categorised as a member of the noun type of peg d^1 iff*

a is an alec with roles $d^1, d^2 \in \{NI\}$, such that d^1 is an indiscernable approximation to a under some criterion c^1 ,

where $c^1 \subseteq Y$ and c^1 is the intersection of the application of criterion-level pre-properties to d^1 and d^2 ,

and such that d^2 is an indiscernable approximation to a under some criterion c^2 , where $c^1 \subseteq c^2$, and $c^2 \subseteq Y$.

I am taking \subseteq in the intuitive informational sense of subsumption, rather than strictly as improper inclusion. Schema **AC'** and **PC'** are elaborations of the intuitive schemas **AC** and **PC**, given in Chapter II. I will now run through an example of the latter, to clarify its workings.

An example of a privative categorisation would be the categorisation of a stone lion as a lion. In this case, we would have:

a = *stone lion* discourse referent

d^1 = *lion* (head N)

d^2 = *statue* (implicitly attached N)

c^1 = central-diagnostic appearance properties

c^2 = central-essence and central-diagnostic properties

Y = "stone lion" sense (central-essence and central-diagnostic properties)

$\{NI\}$ = set of candidate implicitly attached nouns

The categorisation would therefore go through, and we could, again, paraphrase it with

The stone lion is a lion with respect to its appearance qua a statue

PC' defines the two different perspectives which we might adopt on an entity of privative NP-type: the NP entity can be categorised as a member of the head N type with respect to a narrow Type II criterion c^1 , which comprises properties from the intersection of the "appearance"-level pre-properties of the head and implicitly attached nouns; it can also be categorised as a member of the implicitly attached noun, with respect to a Type I criterion, c^2 , which is broader than c^1 .

PC' is, however, a little strong, since it presupposes that a particular implicitly attached NI is necessarily chosen in a privative combination, and hence that we generate a sense for a NP which includes detailed properties and not just pre-properties. We might doubt this, since a sense might not specify the pre-properties; so we will know that a stone lion is a lion regarding appearance, but not precisely what that appearance is. A modification of **PC'**, taking this into account, would be:

PC'': *peg a can be categorised as a member of the noun type of peg d^1 iff*
a is an alec with role d^1 and possible roles $\{NI \mid NI = d^2, d^3, d^4, \dots, d^n\}$, such that
 d^1 is an indiscernable approximation to a under some criterion c^1 ; and where
 c^1 subsumes Y and c^1 comprises criterion-level pre-properties of d^1 whose intersec-
tion with the application of criterion-level pre-properties to any $d^i \in \{NI\}$ would
preserve d^1 as an indiscernable approximation to a under the same depth of cri-
terion.

The locution "same depth of criterion" is intended to carry the implication that the relation of information inclusion, when applied to the criterion applicable for *a* with respect to *d*, and *a* with respect to any $d^i \in \{NI\}$, will be undefined. And an implication of this is that each of the members of NI will bear the same level of indiscernable approximation to the NP. This definition does not depend upon the selection of a particular implicitly attached noun: it simply states that the criterion of indiscernability is consistent with any of the members of $\{NI\}$. So a choice of $d^i \in \{NI\}$ would not restrict further the generalisability of the categorisation. Notice that it defines only the Type II Perspectival categorisation of the NP-type entity as a member of the head N category.

The argument, then, is that these schemas are constraints on categorisation judgements: for a categorisation of the appropriate type to be warranted, it must be an instantiation of the relevant schema.

Having been more precise about the perspectival relativity of categorisation, and how this relates to partiality, I should now comment on the constraints on a permissible categorisation perspective. Earlier, I mentioned that we could view a sense as either supporting or defining perspectives on the entity in question. The notion of "supporting" a perspective is exemplified in the categorisation schemas: the criterion of indiscernability must subsume the range of properties which we know are true of the discourse referent; that is, it must subsume the sense. If the criterion is more extensive than the sense - for example, in being more detailed or inclusive concerning the central-essence properties, then the sense could not support a categorisation using such a criterion. Before such a criterion could be used, we would have to expand the sense. But notice that this just could not be done for certain ways of expanding the criterion: for example, we just could not expand the criterion under which a stone lion can be viewed as a lion beyond the range of *lion* properties which a stone object could have. A stone lion cannot be a lion in the general sense. So although we may often expand the criterion under which an affirmative categorisation holds, we can rarely do so for privative combinations. The constraint, then, is that the intended criterion of indiscernability should subsume the generated criterion: the intended generality of the

categorisation should not outstrip its actual generality.

What of the constraint on the covering sortal? Two different situations arise: one in which an implicitly attached noun is chosen, and one in which it is not. In the first case, we have the kind of categorisation paraphrasable with

The stone lion is a lion with respect to shape qua a statue

Here, we could envisage categorising the object with cognates of *statue*, like *sculpture*, or some other locution like *representation in stone*. The precise properties of the criterion are constrained by the sortal, so if we are to effect a categorisation with a criterion of exactly the same level of specificity, we need a sortal which is a synonym of the implicitly attached noun. Otherwise, we could utilise a criterion of only pre-property level, in which case, we could use various sortals which might be superordinates of the implicitly attached noun: *stone object*, *inanimate object*, and so on. So it is then that a change in sortal carries with it an attendant change in the criterion. In the case where an implicitly attached noun is not utilised, there is a similar constraint. We could categorise the NP entity with nominal terms which contain a noun which is a superordinate of the head N: for example, *stone object*. Again, then, there is little scope for flexibility here. It seems that a concept combination, in defining the properties of a NP discourse referent, strongly constrains permissible categorisation perspectives.

IV.6. On Concepts, Senses and Word Meanings

Having discussed some of the central components of the approach, we are now in a more promising position to consider the crucial relations between lexical concepts, senses, and word-meanings. My argument will be that all three notions are independent but related.

The distinction between senses and lexical concepts has been noted. The latter is the starting point for the generation of the sense: it is the constant, iterable semantic contribution that a lexical entry makes to the generation process. This involves the ascription of central properties - central-essence and central-diagnostic - to the sense for the term being understood or applied. These properties are feature-descriptions that correspond to the criterial-diagnostic and lay-theoretic "essential" properties of the referents to which the term usually applies. By Positive Undermining, we need some strong reason for defeating this ascription. However, the process of generation very often involves such reasons - the background-conditions relative to which the ascription is made can fail in either of two ways, resulting in two different kinds of defeat. Where we have direct evidence that the ascription is inappropriate (for example, where a combining modifier is in conflict), then we might Rebut the

ascription. In cases where the background-conditions simply fail to meet the ordinary requirement, although we have no determinate evidence against the ascription, it may be Undercut. In Chapter V, we will find that such defeat as regards the central-essence properties is a hallmark of Privatives.

So it is, then, that the sense generated for a term may bear varying degrees of resemblance to the lexical concept. The mechanisms for flexibility allow it to alter even to the extent of inheriting central properties of an entirely distinct type of entity; and there is the possibility for many different specifications and degrees of specification. Such a sense, then, is a description of a referent, either a real one or one set up for purpose of a discourse.

Both concepts and senses are therefore (mentally represented) descriptions of entities. Senses are thereby candidates for the label of "word-meaning". On a psychologistic account, this would indeed be the case. However, on a strictly realist account, the denotations of kind terms would be the meanings - that is, the entities of which senses are descriptions. There are, however, good reasons for wanting to make semantic distinctions at neither "level", as noted in I.3.4. Consider, first, making a distinction between meanings at the level of referents. Then we would find that, regardless of our perspective on an object, we would still be using the same meaning. For example, if we view a statue of a lion as a lion, a statue, an obstacle, or a tourist attraction, we would still be using a single meaning. And the well-known propositional attitudes problems would also arise: I can believe that Cicero was a great orator, whilst not believing that Tully was also - simply because I do not know that the same referent can be referred to in these two ways.

If we consider locating meanings as senses, difficulties also arise. In addition to problems with psychologism generally, there is the simple problem of making too many distinctions. This comes down to the question of the individuation of senses. If I am considering a piano as a piece of furniture, and my sense for this piano comprises properties like "colour", "size", "shape", and so on, then does this sense differ from one that also has the property "varnished"? That is, should incidental/inductive properties suffice to differentiate word-meanings? The consideration so far has been that such properties, as non-central, should not count in the individuation of senses. In I.3.3, I noted Peacocke's criterion for the identity of senses, that two senses are evidentially equivalent if and only if they have the same canonical links. So an identity function would take account only of the canonical links of a sense, and not of any links that were above the target sense in a network. That is, two senses are identical if and only if they share exactly the same canonical links (central property ascriptions); non-central properties are irrelevant. Two difficulties emerge, however.

Firstly, the Manifestation Constraint is not concerned with the exhaustive specification of content; in order to map the identity function noted above, we would need to know all of the canonical links of a sense, which may not be possible. Whilst this may be a reasonable conclusion for a theory of (cognitive) senses to arrive at, it is not appropriate for a theory of word-meanings to be so limited. A basic requirement of such a theory is to provide just such a means of discriminating meanings. A second problem is that it again appears to make too many distinctions between meanings as senses. For example, if sense A has one central property more or less than B, they must be judged non-identical. For example, a sense for *dog* with and without the property "dog-collar". Again, such fine discriminations are important to a theory of senses, but appear to burden a theory of word-meaning with too much specificity.

So it is that the individuation of senses and of referents makes too many and too few discriminations, respectively, for a theory of word-meaning. We appear to need, as I have stressed throughout this study, a view that locates word-meanings in the relations between senses and referents. Having discussed perspectives we are now in a position to see what this might mean. Perspectives associate descriptions with referents. Perspectival descriptions are the *intended generalisability* of the overt categorisation statement. That is, categorising a statue of a lion as a lion is relative to the description based upon central-diagnostic properties arrived at through the intersection of the application of the pre-property "shape" to the diagnostic properties of the concepts for *statue* and *lion*. There is no reason to evaluate the categorisation with respect to a description incorporating central-essence properties. But the perspectival description must itself be evaluated with respect to the facts of the matter, the evidence. And this is what is meant by the claim that meaning is relational: it is the relation between the intended categorisation and the facts of the matter, between the perspectival description and the (evidence about the) referent. If the perspectival description subsumes the evidence about the referent, then the categorisation statement is warranted; if it does not, then it is unwarranted. Since, as I have argued, perspectival descriptions must be publicly specifiable (though mentally represented), and since the (evidence concerning) the referent is equally so, we avoid any residual charge of psychologism. Word-meanings are still concerned with the relations between words and the world, but this relation is complicated by the mediation of a perspective, which is crucial to the evaluation of the words/categorisation statements. This issues in the claim that, although word-meanings and mental representations are separable, one needs the latter to properly construe the former. Indeed, I shall now set out why one needs to understand the sense in order to understand the word-meaning.

Senses are descriptions of entities. In the case upon which I have focussed, there is the generation of a sense for a NP followed by the categorisation of a discourse referent whose identity conditions are defined by that sense. In evaluating categorisations of that referent, then, the sense is the evidence against which the perspectival description is measured. It is with respect to this case that the categorisation schemas AC', PC', and PC'' are defined. This stems from the interest in discussing the categorisation implications of generating a sense for a concept combination.

There is, however, a second condition to which I have alluded. This concerns the categorisation of referents without a prior generation of a sense. It is the everyday case, in which we are presented with referents, in which we categorise an entity on the basis of the available evidence. In the case in which we can make a generalisable categorisation, we infer the type's central-essence properties from the presence of the central-diagnostic ones. That is, in applying a term to a referent, we use the lexical concept for that term: depending upon the range of diagnostic properties available, we make inferences to different central-essence properties (if any). In doing so, we may need to generate a sense for a concept combination. Again, then, the sense defines the evidence that we have about the referent, and thus is that with respect to which an intended categorisation must be evaluated.

So word-meanings are relations between perspectival descriptions and the facts of the matter. The perspective is that relative to which the overt categorisation statement must be evaluated; or rather, the evaluation of the perspectival description with respect to the evidence is that relative to which the overt categorisation must be judged. And this evidence very often is defined by the sense generated for the N term used in the overt categorisation statement. Thus it is that warrant for categorisation statements occurs. Where the perspectival description subsumes the description of the evidence, the categorisation judgement is warranted; where it fails to subsume it, the judgement is unwarranted.

We can then say that two word-meanings are identical if and only if they relate descriptions of the same canonical levels to the same referents. We might then note two ways of individuating word-meanings, a coarse and a fine level. At the coarse level, word-meanings are the same if they associate descriptions of the same level (from a particular lexical concept) with the same referent. At the fine level, they are the same if they associate the same particular property-descriptions with the same referent. The coarse level is indeed rather coarse, since the current view allows for only the diagnostic and central levels. This is, though, clearly an empirical matter, and the range of types of contents might expand with investigation. Which is to say, that two word-meanings are identical if and only if the

categorisations are warranted under the same types of perspectives. It is with the coarse level that I will be concerned in the treatment of the problem classes.

So it is that lexical concepts, senses, and word-meanings are inextricably related. And it is through perspectives that they are so intertwined. A perspective associates a description with a referent, and a word-meaning is the evaluation of the applicability of a categorisation based upon that description, resulting in warrant or its lack. Senses can be seen to form the evidence against which perspectival descriptions are evaluated, and are themselves derived from lexical concepts.

IV.7. Summary: the Approach in Outline

I have argued that the process of generating a sense can be viewed as one of extending a lexical concept: as the ascription of properties to that sense. I have also argued that the generation of a sense is part and parcel of the process of constructing a discourse referent. Sense Generation is then the ascription of properties to the discourse referent. Clearly, other factors (such as focussing, determiners and so forth) will have a crucial role in deciding what kind of discourse referent is constructed; but it is the sense that is generated that provides the properties which are ascribed to the discourse referent. In this way we construe the generation of a sense as the construction of a peg for a linguistic expression. The sense provides the content for the identity conditions of the peg.

Having constructed such a peg, what roles can that discourse referent play in the subsequent discourse? Consider a peg for a NP. We can examine how that peg might be categorised as a member of the head N type, where such a categorisation can be viewed as a role of the NP peg considered as an alec. If we have ascribed the N's central-essence properties to the NP sense/peg, then categorisation of the peg as a member of the head N type (i.e., the role type) is permissible with respect to a criterion of indiscernability of central-essence level. The categorisation will thus obtain under a Type I Perspective. If the peg has only central-diagnostic properties of the head N, the categorisation will carry a more clearly defeasible level of certainty (after the criterial manner). It will obtain under a Type I Perspective where central-essence properties are undefeated, and a Type II Perspective where the central-essence properties are defeated. If the peg has only non-central properties of the head N, then the categorisation will hold only under a restricted, Type II Perspective. In the Privative case, the latter kinds of categorisation (based upon a defeat of the head N's central-essence properties, and an ascription of central-diagnostic and/or non-central properties) will require some specification of the NI, and hence some variant of the PC schema.

The combination process parallels the views proposed by Hampton (1988), Medin & Shoben (1988), and Murphy (1988). It can be divided into two basic "stages" or aspects of the combination process. The first is the Combination Stage, and the second is an Extension Stage. The second stage may have one or more substages; since the Extension Stage is optional, the process of sense generation for a combination can involve one or more stages. The Combination Stage involves the determination of the concept extension of the N by the modifier. Given the current thematic dimension, the N lexical concept would be extended in a particular way. The modifier acts in two ways. Firstly, it determines which of the properties of the N's lexical concept are ascribed to the sense (for example, by acting as a defeater for some of the N's properties). This includes acting as a defeater for head N properties, as in Privatives. In Privative combinations the Modifier stipulates only the pre-properties of the head N to be inherited by the NP. Secondly, the modifier prompts an ascription of properties from its own lexical concept to the NP sense. This ascription may be the source of the first aspect of modifier influence, as I will claim for Functional Privatives. If there is no requirement for further information concerning the entity, then the generation of the sense might cease at this point. At this point, the NP sense is consistent with many different instantiations, since it underspecifies the referent of the NP in the discourse scenario. All of the different instantiations will possess some specification of the pre-properties of the N. For example, at the end of the first stage, the sense for *fake gun* might include the pre-properties of "appearance" of a real gun. Each possible instantiation indicates a different way of faking, and thus a different specification of the pre-property. But since we have not selected the NI as yet, the categorisation schema appropriate to this stage would be PC''.

Now, if there *is* a requirement for more specific information concerning the NP entity, the sense may be extended through selecting a NI (through the instantiation of the NP). The Instantiation of the NP is a discourse referent where the head N is a modifier of the N type of the referent. For example, an instantiation of *fake gun* might be to a referent which was a "replica (of a gun)". This directs the inheritance of properties from the head N to the NI *replica*; the NP can now inherit properties from a different (NI-based) domain of world knowledge. The property inheritance mechanism is an analogical mapping. This mapping is effected to the deepest possible level under the current state of information. For Negating Privatives, the outcome of this stage is a sense whose categorisation possibilities are expressed in PC'; a Type I Perspective on the NP as a member of the NI and a Type II as a member of the head N are thus supported. Hence the former categorisation can be expressed through AC'. The position for Equivocating Privatives is a little more complicated, but follows the same essential pattern.

This brief sketch indicates that the parallel with the approaches of Medin & Shoben and Murphy, are clear. A relatively knowledge-free combination is followed by a knowledge-intensive extension. The mechanism to be proposed here is, however, somewhat more precise concerning these processes and the interactions of the different kinds of knowledge involved.

IV.8. Summary

I have attempted to be more precise about some of the conceptions of my approach. I have noted what kinds of properties (central-essence and central-diagnostic) the lexical concept contributes, and the precise ways in which they might be defeated by a modifier, with what effects. **Type I Rebuttal** denies both the applicability of the N term, and the ascription of the central properties to the NP sense; **Type II Rebuttal** denies only the latter; **Undercutting**, a form of defeat not available to the classical realist, denies neither outright, but casts the latter into doubt so that the ascription cannot be warranted on current information.

In a Privative combination, then, the modifier acts as a defeater for the head N's central-essence properties. The remaining properties are then employed to determine the instantiation, a concept related to which forms the base for **analogical transfer** to the NP sense as target. I argued that an analogical mechanism for property inheritance was motivated on both specific grounds relating to partiality and concept combination, and on the general ground of being a motivated inheritance approach.

The sense thus generated defines a peg for the NP. This peg can then be categorised as a member of the N type of an alec (or role that the peg can play during a discourse), under an appropriate perspective. In a Negating Privative combination, the NP entity can be categorised as a member of the head N under a restricted, **Type II Perspective** (utilising a criterion of indiscernability that does not include the N's central-essence properties); and as a member of an implicitly attached N type under a generalisable **Type I Perspective** (where the criterion does include the NI's central-essence properties).

Utilising an analogical mechanism for property inheritance required an alteration to Landman's definition of an indiscernable approximation (to allow for analogue properties' being used in a categorisation criterion). I also set out three conditions on categorising a peg as a member of an alec's type; these are alterations and more precise formulations of the AC and PC schemas given at the end of Chapter II. These conditions are constraints on the warrant of the relevant categorisations. As a result, they are concerned with the

meanings of words, and not simply with mental representations. In the final section of this chapter, I considered the relations between concepts, senses and word-meanings, and argued that the latter two notions are inextricably linked through perspectives. These ideas, and the categorisation schemas, will play a crucial role in the treatment of the problem classes in Chapter V.

CHAPTER V

On Privatives and Attributives: the Current View

In this chapter, I provide a treatment of the problem classes within the current framework. I then go on to consider Privatives in detail, distinguishing in detail between Functional and Privative and Negating and Equivocating Privatives. I also give a briefer account of Modal and Pure Privatives. These accounts give rise to a tentative typology of Privative Combinations. I then provide an account of the kinds of Attributives that have figured in the project. A final implication will be that we can range combinations along a continuum leading from extreme Privative to the Predicative cases.

V.1. On Negating and Equivocating Privatives

V.1.1. Qualities and Types of Privatives

My argument has been that received views of Privatives cannot properly account for some very strong intuitions as a result of their non-observance of the Manifestation Constraint. We can isolate three assumptions which have been challenged. Firstly, there is an assumption that all privative A's can be given the same treatment, since their behaviour is uniform. Hence, A's as disparate as *fake*, *apparent* and *former* have been given the same treatment (e.g., Hoepelman, Montague). Secondly, the characteristic behaviour of privatives is their failure to support an inference from being privative NP-type to being head N-type: whatever a fake gun is, it just is not, in any sense, a gun. Thirdly, the privative behaviour is exclusively a function of A type: if the A is not semantically of privative type then it will, by definition, behave affirmatively.

The main thrust of the arguments I have presented has been to challenge the second assumption. I have argued that a fake gun can be categorised as a gun under a limited perspective, and that this is an instance of the perspectival-relativity of categorisation judgments. We can adopt many different perspectives on a particular entity, under which different categorisations may be accurate or not. Understood within this framework, Privatives give rise to senses which support two extreme type of categorisation perspectives. However, I will note, in V.3, two small subgroups of Privatives - the Modal and Pure Privatives - that do appear to fulfill the second assumption (e.g., *former* or *imaginary*). As a result, these A's will receive a slightly different treatment from the main types.

I shall now challenge the third and first assumptions in turn. In challenging the first, I will make a fundamental distinction between types of privative behaviour, deriving from Negating and Equivocating Privatives. In challenging the third assumption, I will extend the idea of privativity to cover not simply A types, but also the results of concept combinations; that is, some combinations are Functional Privatives.

V.1.1.1. Negating and Equivocating Privatives:

The paradigm case of a Negating Privative is *fake*, whereas that for Equivocating Privatives is *apparent*. Negators involve a direct denial of the central-essence properties of the head N: a fake gun cannot perform the essential function of a gun. In my treatment, a Negator prompts a Type II Rebuttal of the head N's central-essence properties. In contrast, Equivocators involve an equivocation over the ascription of the head N's central-essence properties to the NP sense. If someone is described thus:

He is an apparent friend

then, on the basis of current information, all we can state is that he in some way resembles a friend. But we have insufficient information to decide whether he is a real friend or not. With the addition of further information, we might derive a privative or an affirmative interpretation of the combination. In my treatment, Equivocators involve an Undercutting of the head N's central-essence properties.

This, then, is the most important difference between Negators and Equivocators: they require different kinds of defeat of the central-essence properties of the head N. So, regarding the NP sense, they contrast in terms of coarse, central level content.

Another important difference is that describing something as an *apparent N* appears to involve a particular epistemic attitude towards the entity on the part of the describer, in a way that describing something as a *fake N* does not. In contrast, describing something as a *fake N* involves the assumption of some epistemic attitude on the part of the creator (or manifester) of the entity or behaviour, whereas describing it as an *apparent N* does not. Let me elaborate. Consider an entity described as a *fake gun*: that it is a fake gun is a matter of objective fact, in that it bears the appropriate relation to real guns. Under a state of total information, it would still be a fake gun. In contrast, someone described as an *apparent friend* remains an apparent friend only insofar as we adopt the appropriate attitude to him; that is, only while we have a perspective on him based upon the current limited state of information. Once we have expanded our information, he ceases to be an apparent friend, and becomes either a real friend or a fake friend. So the crucial aspect of Equivocators is that they involve a suspension of judgement as a result of having strictly partial

information, whereas Negators involve a direct judgement of the central-essence properties of the head N. In addition to this, the epistemic attitude embodied in Equivocators is evidenced in the particular A employed. The A used indicates not only that we have strictly partial information about the entity, but also what we expect the result of an extension of that information to be. That is, whether we expect the outcome to be privative or affirmative. For example, describing someone as an *apparent friend*, appears to presume that he is not, in fact, a real friend; whereas to say that someone is a *probable friend* carries the contrary assumption. In my treatment, I will label the two possibilities as Default Privatives and Default Affirmatives, respectively. However, this partitioning may not be exhaustive. Consider the following:

The man in the corner is the alleged criminal

In this case, it is not the speaker's attitude which is evidenced in the description, but her report of the attitude of someone else. For example, the attitude of the prosecution in a court case. And there is no reason why the speaker's and the other's attitudes need coincide. Indeed, if they did, or did not, then the speaker might be more likely to use a Default Affirmative or Privative accordingly. It seems, then, that *alleged* represents a "pure" form of equivocation, in which the speaker does not know whether an alleged criminal really is a criminal, and has no clear presumption either way.

The epistemic attitude related to Negators is closer to that involved in *alleged* than in other Equivocators. To describe an entity as a *fake N* is often to assume that there has been some deliberate attempt at deceit on the part of the person who either created the entity or manifested the behaviour. A fake gun is intended to be mistaken for a real gun; this means, also, that there will be some (limited) perspective from which we might mistake the NP entity as a member of the head N type. If we do not look closely, we might categorise a fake gun as a real gun. In contrast, although we might initially categorise an apparent N as a real N, we do this without presuming that there was necessarily any intent to deceive. So whereas Negators appear to typically involve some presumption concerning the creator's intent, Equivocators carry such a presumption only contingently. In the treatment, I hope to capture the difference between interpretations which do, and those which do not impute an intention to deceive to the creator, by allowing for the selection of different implicitly attached nouns through different instantiations. For example, if we know that there has been such a defrauding intention, then the instantiation for *fake gun* might be *replica*, whereas it might be *toy* if there is no such intention.

A final difference between Negators and Equivocators might be in the range of N's which can coherently be modified by each. That is, certain N's which make reasonable sense

when combined with a Negator make little sense when combined with an Equivocator. For example, consider the following pairs of combinations:

This is a fake/false/sham man

This is an apparent/ostensible/alleged man *?

This is a fake/false/sham lion

This is an apparent/ostensible/alleged lion *?

This is a fake/false/sham chair

This is an apparent/ostensible/alleged chair *?

The combinations using Negators are at least interpretable, whereas it seems that those using Equivocators are problematic. And there are many areas where both types might be used:

This is a fake/false/sham friend

This is an apparent/ostensible/alleged friend

This is a fake/false/sham compromise

This is an apparent/ostensible/alleged compromise

This is a fake/false/sham ally

This is an apparent/ostensible/alleged ally.

I am not claiming that all of the Negators are equally applicable in all of the cases; it suffices for my argument that at least one of the Negators makes clear sense in cases where none of the Equivocators can operate.

It seems, then, that the domain of application of Equivocators is narrower than that of Negators; it may consist of a proper subset of the latter. But what is the semantic source of this difference? It seems that Equivocators cohere most with N's which involve some "evaluative" dimension; their referents have a very clear rating on some dimension of socio-cultural value. For example, *friend*, *ally*, *victory*, *defeat*, and *murderer*. And those cases with which they do not appear to cohere so well include the basic kind or artefact terms like *lion*, *dog*, *man*, *chair*, and *table*. The latter restriction may derive from the above point concerning the non-obligatory assumption of an attitude of the creator. If something is possibly not a "real" N, where the N is a kind or artefact term, then it seems that we would need to be committed to the NP entity's being created so as to allow for mistaking or deceit. We don't just stumble across fake men everyday. But this is not the case for compromises, victories, friends and solutions; these are all variable by their nature,

and the fact that an apparent solution may turn out not to be a solution does not in any way imply that the "maker" of the solution intended to deceive. •

These comments concerning the different domains of application are necessarily tentative. The possibility is not integral to the distinction between Negators and Equivocators. If the distinction between Negators and Equivocators is well-founded, then there appears to be no real reason to suppose that a uniform treatment is possible or desirable. Thus the first assumption of standard views of Privatives is challenged.

V.1.1.2. Functional and Proper Privatives:

The distinction between Negators and Equivocators is orthogonal to another distinction, concerning the origin of the privative behaviour. The examples used in the last section are all of Proper Privatives; that is, they result from the semantic type of the A's involved.

The assumption that privative behaviour is confined to a restricted set of A's results from a non-interactive view of concept combination/ adjectival modification. On this view, which is exemplified by Hoepelman, the behaviour of the NP is completely conditioned by the A: the A alone determines whether the properties of the A are to be interpreted attributively, predicatively, or privatively. On the current, interactive view, we would have to argue that such evaluations are a result of the interaction of the modifier and the noun. So although the modifiers *stone*, *plastic*, *square* or *waterless* can behave affirmatively in certain N contexts, they can also have privative effects in other N contexts. Consider these contrasts:

stone lion/stone bridge
plastic flower/plastic container
square basketball/square picture
waterless ocean/waterless paint

In the first of each of these combination pairs, we can discern the same kind of privative behaviour as in Negators (such as *fake gun*, *false hope*). That is, the inference from

This is a stone lion

to

This is a lion

will go through only in a restricted Type II Perspective. It also seems that the modifier effects a Type II Rebuttal of the N's central-essence properties. But the source of the defeat differs from Proper Privatives. For Functional Privatives, it is the interaction of the properties of the modifier and noun which is crucial. In the context of modifying *lion*, *stone* creates a privative NP; in modifying *bridge*, it creates an affirmative NP. Hence the defeat of the N's central-essence properties is indirect, in that it is a result of the propagation of

interactive effects through the properties of the N. There is nothing in the lexical concept of *stone* which determines that it must defeat the central-essence properties of a N with which it combines; nor is there any specification that *stone* should be incompatible with the presumed internal essence and other central-essence properties of *lion* (for if this were in the lexical concept, then every other conceivable contrasting property would also have to be in it). So *stone* does not directly defeat "internal essence"; rather, attempting to ascribe *stone* to *lion* as a "material" or "substance" leads to a defeat of an animacy property, which then defeats the internal essence assumption. So although Negating Proper Privatives and Negating Functional Privatives embody the same kind of defeat of the head N's central-essence properties, the origin of that defeat differs.

The analogues of *apparent friend*, *ostensible honesty*, *probable murderer* and *presumed integrity*, in the Functional Privatives, are *blue orange*, *wooden frying-pan*, *square lemon* and *soft knife*. In each of these cases, the central-essence properties of the head N are Undercut, so that we are uncertain as to whether an entity of NP type could be categorised as a member of head N type from a Type I or a Type II Perspective. This defeat again occurs through the propagation of the interactive effects. For example, modifying *frying-pan* with *wooden* modifies a central-diagnostic property of material; this may or may not be linked to a rebuttal of the central-essence properties regarding function. We might imagine a case in which the object has a heat-resistant coating and so can fulfill the function of a frying-pan; so the Undercutting of the central-essence properties does not lead to their rebuttal. Here, the Undercutting of central-essence properties results from the rebuttal of the central-diagnostic properties (as noted in IV.2.2).

Now, we might also be able to note analogues of Default Privatives and Affirmatives in the Functional domain. For example, it may be that *wooden frying-pan* and *blue orange* are Default Privatives, since the rebutted central-diagnostic properties appear to have a relatively close link to the central-essence properties. In contrast, *square lemon* and *soft knife* appear to involve a rebuttal of central-diagnostic properties which are only contingently related to the central-essence properties; hence they are Default Affirmatives.

The crucial difference, then, between Proper and Functional Privatives lies in the source of the defeat of the head N's central-essence properties. This distinction challenges the third assumption, noted above.

V.1.1.3. Summary: the Meaning of Privatives:

The discussion indicates two aspects to the meaning of a Privative A. Rather, there are two

different aspects of the contribution which such a modifier makes to the NP sense. Exactly the same contribution is made by otherwise affirmative modifiers in a Privative interaction.

Firstly, the modifier acts as a defeater for the central-essence properties of the head N. The particular type of defeat depends upon the the type of Privative (or the nature of the interaction): Negators involve a Type II Rebuttal, and Equivocators involve an Undercutting.

Secondly, the modifier prompts an ascription of a pre-property of the head N to the NP sense. This pre-property is of central-diagnostic or non-central level, and involves the "appearance" of the head N or some cognate. This functions through the application of the pre-property to the head N to yield a more precise pre-property.

In Section V.3, we will see that Pure Privatives concur with the main types in prompting a form of defeat of the head N's central-essence properties; however, the type of defeat differs, as does the second aspect of the A's meaning.

V.1.2. An Account of Negating Privatives

In this section, I give an account of each of the main types of combination in the Negating and Equivocating groups. I begin with Negators, and follow with Equivocators. In each case, I will indicate some properties which might be included in the central-essence, central-diagnostic and non-central properties of each of the concepts. Following the nature of the Manifestation Constraint, these are not intended to be exhaustive content specifications.

V.1.2.1. Negating Proper Privatives:

The example which I shall consider here is *fake gun*.

fake:

Type II Rebuttal of N's central-essence properties

Ascription of the result the of application of "Appearance" pre-property to N, to NP sense.

gun:

Central-essence properties: internal mechanism for propelling bullets (usually small, controlled explosion); barrell for directing bullets; essential function of

being able to kill, maim or otherwise damage a target.

Central-diagnostic properties: trigger to fire with; metallic grey colour; handle to the rear of trigger; barrell mounted over trigger, to the fore of handle; hand-sized; magazine holding bullets in handle.

Non-central properties: silencer; sights; kinds of noise made; ability to fire repeated shots.

The lexical concept for *gun* comprises an ascription of the central properties to the sense.

Stage I: Combination

This has two components. The first is the defeat of the N's central-essence properties by the A. In the schema for Type II Rebuttal,

P = use of term "gun";

Q = ascription of central-essence and central-diagnostic properties to the N concept (and thus to the NP sense);

R = the application of the adjective "fake" to "gun";

Here, then, the application of *fake* to *gun* results in *fake* acting as a Type II Rebutter of the central-essence properties of *gun*; this means that we can infer $\neg Q$ (as regards central-essence properties), but that we cannot infer $\neg P$ (so the term *gun* can still be coherently applied). Hence, we cannot ascribe any of the central-essence properties of *gun* to the sense/peg for *fake gun*. The independence of central-diagnostic from central-essence properties prevents a defeat of the latter defeating the former.

The second aspect is the application of the pre-property "appearance" to the concept for *gun*. The pre-property selects the pertinent properties of the N concept. It selects those central-diagnostic properties that relate to the appearance of a gun (leaving out, for example, having a magazine to hold bullets); and those non-central properties that come under the current thematic interest and which relate to the appearance of a gun (leaving out the ability to fire repeated shots). Applying "appearance" to "gun" yields, then, a pre-property which can be labelled "appearance of a gun". This pre-property is then ascribed to the NP sense. We cannot, at this stage, say in precisely which way the fake gun manifests its gun appearance, in which way the pre-property is specified.

At the end of this stage, then, we have constructed a peg for *fake gun*, to which has been ascribed the "appearance of a gun" pre-property. On the basis of these properties, the peg (as an alec) can play various different roles, and can be categorised as a member of the N

types of those roles under appropriate perspectives. A number of different N's will be indiscernable approximations to the NP, with respect to the criterion of the pre-property: a gun, toy, model or replica, will all be indiscernable approximations to the "fake gun" peg with respect to the criterion of "appearance of a gun". For each of these N's, the NP peg is an indiscernable approximation to the N (thus ensuring that we can view the roles as being played by the same peg: it is a toy with the "appearance of a gun" pre-property, as well as its own properties). Each of these possible N's potentially provide specifications within the pre-property. In categorising the NP peg as a member of the various role N's, the latter must subsume, or be an indiscernable approximation to, the former, with respect to the operant criterion. Hence, we can see that the following categorisations would be warranted:

The fake gun is a gun with respect to appearance

The fake gun is a toy (gun) with respect to appearance

The fake gun is a model (of a gun) with respect to appearance

The fake gun is a replica (of a gun) with respect to appearance

These are all categorisations under Type II Perspectives, in which the criterion of indiscernability does not incorporate the central-essence properties of the categorising N. The first categorisation statement satisfies the condition of categorisation schema PC'', where

peg a = *fake gun* discourse referent;

peg/role d^1 = *gun* N type;

NI possible roles d^2, d^3, d^4 = *toy, gun, replica* N types;

Y = *fake gun* sense;

c = "appearance of a gun" criterion;

Here, the *gun* sense is an indiscernable approximation to the *fake gun* sense with respect to appearance; appearance subsumes the properties ascribed to the NP sense; the intersection of the application of the pre-property of "appearance" to "gun" and any of "toy", "model" or "replica" concepts would result in a peg which is an indiscernable approximation to "fake gun" under the same "appearance" criterion as before.

And the categorisation judgements regarding the NI's can be seen to satisfy schema AC': taking a = *fake gun* discourse referent, and d^i 's = each of *toy, replica* or *model*, we can say again that each d^i is an indiscernable approximation to the peg with respect to c = "appearance of a gun".

Hence, if we were to cease the construction of the NP sense at this point, it would define a Type II Perspective on the fake gun as a gun, and support various Type II Perspectives on it as a toy, a replica, and a model. If, however, there was sufficient requirement, we might

seek to extend the sense through Stage II.

Stage II: Extension by Implicit Attachment

In order to provide a specification within the pre-property, the "fake gun" must be instantiated as a particular type of entity within the discourse situation. The instantiation is a more precise version of the NP peg defined at the end of Stage I. There are two minimal constraints on the selection of an instantiation. The first is that the instantiation and the peg must be superficially similar: selection of the instantiation is based upon those of its properties that subsume the pre-property ascribed to the NP peg. Each of the candidate instantiations possesses a specification of the pre-property of "appearance of a gun". The second constraint is that the instantiation must be able to play the role of the peg in the discourse scenario. That is, whatever roles, causes and effects the peg entity has, the instantiation must also have. For example, consider the following scenario:

Marlowe entered his office. He saw someone in the corner, coming towards him. Quickly brandishing his fake gun, Marlowe frightened the intruder into the corner. Then he slugged the guy on the temple with it, reducing him to a crumpled pile of bad suit, hat and shoes.

The range of options for an instantiation would be a model, a toy, or a replica. There are two evident requirements on the instantiation: first, the fake gun must look sufficiently like a real gun to be able to force a hoodlum into a corner. A toy would therefore be unlikely. Second, the fake gun was used to club the intruder, and hence must be heavy and solid enough to do so; a model would therefore also be unlikely. We are left with a replica, which could support both functions. Hence the instantiation is as a replica of a gun.

The instantiation of the peg determines what the implicitly attached concept will be. In this case, it will be the that of the immediate superordinate of the description of the instantiation, *replica*. The head N of the NP is a modifier of the NI. The NI is then utilised as a source for the NP's inheritance of properties.

The inheritance of properties occurs through the analogical transfer from the NI to the NP. Hence, for the transfer,

the base = NI concept ("replica")

the target = NP peg ("fake gun")

where the latter is the peg as construed through the instantiation (that is, it has the "appearance of a gun" pre-property, and we know that it can perform the roles and functions of a replica of a gun).

Having accessed the base, the next stage is to effect a matching of the base and target, with the goal of pursuing the matching of "replica" and "fake gun" as deeply as possible. In this case, since we know that a replica can play the role of the fake gun entity in the current scenario, we can match the base and target down to central-essence level. For an artefact, central-essence properties are crucially concerned with function; in this case, the function of a replica might be said to be as close as possible to a facsimile of a real gun, in terms of appearance, weight, and so on, such that one could only tell that it were not a real gun by attempting to fire it. The peg entity matches these central-essence properties of "replica". The transfer of properties from "replica" to the "fake gun" peg can then take place at central-essence and any other level.

Transfer begins with the detailed specification of properties which come under the pre-property "appearance of a gun". Hence, colour, weight, shape and material properties of the NP are given the values of "replica" rather than those of "gun". Hence the appropriate properties are transformed, so that the "appearance of a gun" is really an "appearance of a gun as a replica". This is achieved through the mechanism Bartsch suggests for the construction of properties through the iteration of pre-properties. If we apply the pre-property of "appearance" to "replica", then we construct the pre-property "appearance of a replica", which stands alongside "appearance of a gun". The crucial point about the fake gun in this case, though, is that it has gun properties as they are manifest in a replica, and replica properties as they are manifest in a gun. Hence, we take the intersection of "appearance of a gun" and "appearance of a replica", to form the properties which might be labelled "appearance of a gun as a replica". In terms of the schema noted earlier (III.4.1.2),

t_1 = "gun" concept

t_2 = "replica" concept

P^0 = "appearance"

$P^0(t_1) = P^1$ = "appearance of a gun"

$P^0(t_2) = P^2$ = "appearance of a replica"

$P^1 \ \& \ P^2 = P^3$ = "appearance of a gun as a replica" = "appearance of a replica as a gun"

Transfer also involves properties from other levels. At the central-essence level, causal relations which support the replica's looking like a gun will be transferred. For example, it is often the case that, in order to provide for sufficient weight, a plastic replica will have weights placed inside the handle. In addition, a central-essence property of a replica might be concerned with the assumed attitude of the creator, as mentioned in V.2.1.1: we might, then, transfer the belief that the maker intended the replica to deceive.

At the end of Stage II, the peg has been extended through the process of concept extension. This has involved both the addition of properties of central-essence level, and the specification of the pre-property. So the NP *fake gun* sense has been ascribed the central-essence properties of the NI *replica*, and the central-diagnostic and non-central properties of the head N transformed through the analogical transfer from *replica*.

Having constructed the peg, it may then play the role of other N types, through being categorised as a member of those types. These types are indiscernable approximations to the *fake gun* peg with respect to different criteria of indiscernability. We can isolate two different categorisations:

The fake gun is a gun with respect to appearance, as a replica

The fake gun is a replica with respect to central-essence properties

The first is a Type II Perspective on the the NP entity as a member of the head N; the second is a Type I Perspective on it as a member of the NI. The first is warranted in virtue of its being an example of the schema PC'. Here, then,

peg a = *fake gun* peg

peg d¹ = *gun* N type

peg d² = *replica* N type

c¹ = "appearance properties"

Y = "fake gun" sense

c² = central-essence and appearance properties

The fake gun can be categorised as a gun, since the latter is an indiscernable approximation to a fake gun with respect to appearance properties (which latter subsume the properties in the "fake gun" sense); the appearance properties result from the intersection of the application of the appearance pre-property to "gun" and "replica"; the replica is an indiscernable approximation to the fake gun with respect to both central-essence and appearance properties. Hence the criterion with respect to which the gun is an indiscernable approximation to the fake gun subsumes that criterion with respect to which the replica is such an approximation. Implicit in the last point is the characterisation of the second categorisation (of the fake gun as a replica) in terms of the AC' schema.

Hence, the NP sense defines two contrasting, but not contradictory, perspective-relative categorisations of the fake gun as a gun and as a replica. It could also support other, less extreme, perspectives on the fake gun as a replica.

The NP sense can be seen to support only a limited set of perspectives. If one were to attempt to categorise the fake gun as a gun in any sense other than its appearance (for

example, if we were to try to include it among items of ballistic weaponry, along with canons and anti-tank guns), the categorisation would fail. This is because it would be attempting to apply a criterion of indiscernability which outstrips that which actually obtains; that is, it would be an attempt to over-generalise the categorisation.

V.1.2.2. An Account of Negating Functional Privatives:

The example which I shall consider here is *stone lion*.

stone:

Central-essence properties: solidity; mass; inanimacy; inorganicness;

Central-diagnostic properties: texture; colour; particular shape; hardness;

Non-central properties: whether sculpted; location; size;

lion:

Central-essence properties: animacy; ontogenesis; presumed "internal essence";

Central-diagnostic properties: shape (legs, head, tail, etc.); proportion; colour; carnivorous;

Non-central properties: typical habitat; tendency to roar, etc.

Stage I: Combination

Combining the lexical concepts of *stone* and *lion* results in a Type II Rebuttal of the central-essence properties of "lion", and an ascription of the result the of application of the "appearance" pre-property to the N, to the NP sense. However, this comes about as a result of the interaction of the properties of the head and modifier, rather than the semantic type of the modifier.

In combination, the central-essence properties of the modifier take precedence; this accords with many views on the head-modifier relationship. Attempting to ascribe the properties of inanimacy and inorganicness to the NP results in a Type II Rebuttal of the "lion" central-essence properties of being animate and organic, which then propagates to a defeat of the crucial central-essence properties of the lay-theoretic presumption of ontogenesis and internal "essence". So we can ascribe the central-essence properties of the modifier to the peg. The application or use of the N term *lion* remains open as a result of the type of defeat, as does the ascription of the non-central-essence properties of "lion" to the peg.

However, there is an incompatibility in attempting to combine central-diagnostic properties. On the one hand, *lion* prompts an ascription of the shape and proportion of a lion; on the other hand, although the lexical concept for *stone* does not impose any particular shape,

whatever shape is ascribed must be shape as a stone object. And this is at odds with an ascription of central-diagnostic properties based simply upon "lion". So, we first apply a single pre-property of "shape" to the two concepts, resulting in two other pre-properties: "shape of a stone thing" and "shape of a lion", where the former rules out certain possibilities of shape, such as flexibility and detail. Then we form the more precise pre-property from the two, resulting in the pre-property, "lion shape as a stone thing". This corresponds with two intuitions. The first is that we can know that the NP entity is lion-shaped, but only as regards stone things. So it may miss out much of the detail of a lion which is not usually employed in diagnostic circumstances. The second is the contrast with Negating Proper Privatives at this stage. Whereas the "fake gun" was known only to have the "appearance" of a gun (the lack of specificity reflecting the manner of defeat of the head N's central-essence properties), a stone lion is known to be lion-shaped in a particular way. Because of the source of defeat we have a more specific pre-property. However, in common with Negating Proper Privatives, we cannot, at this stage, say precisely how the NP entity manifests this pre-property.

At the end of the combination stage, we have generated a sense through the ascription of the pre-property of "lion shape as a stone thing" to the NP peg. A number of different N's will be indiscernable approximations to the NP, with respect to the criterion of the pre-property. For example, a statue, ornament, or rock-formation, will all be indiscernable approximations to the "stone lion" peg under the criterion of the pre-property. And again, each of these possible N's could provide specifications within the pre-property. In categorising the NP peg as a member of the various role N's, the latter must subsume, or be an indiscernable approximation to, the former, with respect to the operant criterion. Hence, we can see that the following categorisations would be warranted:

The stone lion is a lion with respect to shape

The stone lion is a statue (of a lion) with respect to shape

The stone lion is an ornament (of a lion) with respect to shape

The stone lion is a rock-formation (of a lion) with respect to shape

These are all categorisations under Type II Perspectives. A minor difference between the Negating Proper Privative case and the current one concerns the criterion of indiscernability. In that case, the criterion which applied for the categorisation from *fake gun* to *gun* did not differ appreciably from the criteria which applied to the other categorisations. In this case, however, the criterion must be narrower for the categorisation of the stone lion as a lion, than for those as statue, ornament, or whatever. Whereas the pre-property of "lion shape as a stone thing" applies as the criterion for the other categorisations, for the categorisation as a lion, the criterion must be a subpart of that pre-property.

The first categorisation statement satisfies categorisation schema PC'': where

peg a = *stone lion* discourse referent;

peg/role d^1 = *lion* N type;

NI possible roles d^2, d^3, d^4 = *statue, ornament, rock-formation* N types;

Y = "stone lion" sense;

c = "shape of a lion" criterion;

So, the sense, "lion", is an indiscernable approximation to "stone lion" with respect to appearance; appearance subsumes the properties ascribed to the NP sense; the intersection of the application of the pre-property of "shape" to "lion" and any of "statue", "ornament" or "rock-formation" would result in a peg which is an indiscernable approximation to "stone lion" under the same "shape" criterion as before.

And the categorisations regarding the NI's satisfy schema AC': in this case, taking a = *stone lion* discourse referent, and d^i 's = each of "statue", "ornament" or "rock-formation", we can say again that each d is an indiscernable approximation to the peg with respect to the "shape" criterion.

Hence, if construction of the NP sense ceased at this point, it would define a Type II Perspective on the stone lion as a lion, and support Type II Perspectives on it as a statue, an ornament, and a rock-formation. If, however, there was an informational requirement, we might seek to extend the sense through Stage II.

Stage II: Extension by Implicit Attachment

In order to be more specific concerning the way in which the stone object manifests the shape of a lion, we may instantiate the "stone lion" peg. Once again, the two constraints are: Firstly, the instantiation must bear a superficial similarity to the peg; the properties of the former are a specification of the latter. Selection is thus based upon the properties of the instantiation that subsume the pre-property "lion shape as a stone thing". Hence, we might have the three candidates already noted: the concepts for *statue, ornament* and *rock-formation*. Notice that this constraint is stronger for Negating Functional than for Negating Proper Privatives: the latter only provides a general pre-property, whereas the former provides such a pre-property qualified by some "material" or other property restriction. Hence the range of candidates for Negating Functional Privatives is more restricted, and there is greater similarity between the members of the candidate set. Secondly, the choice from the candidate set must be of an instantiation that can play the role of the peg in the discourse scenario. In the case of Fred in the park in London, the only viable instantiation is "statue of a lion".

The instantiation provides the NI type, whose concept, "statue", is used as the base in the process of analogical property transfer. The matching of the base and target (the "stone lion" peg) is then constructed. Since a statue can play the role of the stone lion entity in the current scenario, we can match the base and target down to central-essence level. The function of a statue might be to represent whatever is the subject. This carries with it the permissibility of violations of various of the usual properties of the subject. In this case, the peg entity can be said to match those central-essence properties of "statue". As a result of the match, the transfer of properties from "statue" to the "stone lion" peg can take place at central-essence and other levels.

The first aspect of transfer is the specification of properties which come under the pre-property "lion shape as a stone thing". So the size, proportion and shape constraints of a real lion are relaxed: a statue can be of odd proportion, huge size, and be shaped in a more or less realistic manner. These properties are then given the values of "statue" rather than those of "lion". The process of analogical transfer, then, transforms the properties so that the "shape of a lion" is really a "shape of a lion as a "statue". Again, there is a construction of properties from the iteration of pre-properties. Applying the pre-property of "shape" to "statue" constructs the pre-property "shape of a statue". The crucial point is that it has lion properties as they are manifest in a statue, and statue properties as they are manifest in a statue of a lion. Hence, we take the intersection of "shape of a statue" and "shape of a lion as a stone thing", to form the properties which might be labelled by "shape of a lion as a statue". Notice that this specifies the "shape of a statue" pre-property in two ways: it indicates what the statue is made of (not bronze), and it indicates what the shape of a statue is a shape of (a lion, not a unicorn). And it specifies the "shape of a lion as a stone thing" by making more precise what kind of "stone thing" is relevant.

Transfer at the central-essence level includes causal relations which support the statue's looking like a lion. This will particularly reflect the interpretive license accorded to representations or works of art: as a statue, we know that the object need not closely mirror the properties of the represented object. Again, the assumed attitude of the creator might be transferred: with the goal not to deceive the public into believing that the statue is a real lion, but to represent a lion for certain purposes (memorial or ceremonial).

At the end of Stage II, then, the process of concept extension has ascribed to the sense for *stone lion* the central-essence properties of the NI concept "statue", and the central-diagnostic and non-central properties of the head N transformed through the analogical transfer from "statue".

Once again, the peg may then play the role of other N types, through being categorised as a member of those types. These types are indiscernable approximations to the *stone lion* peg with respect to different criteria of indiscernability. There are two extreme categorisations:

The stone lion is a lion with respect to shape, as a statue

The stone lion is a statue with respect to central-essence properties

The first is a Type II Perspective on the the NP entity as a member of the head N; the second a Type I Perspective on it as a member of the NI. The first is warranted since it is an example of schema PC'. Here,

peg a = *stone lion* peg

peg d¹ = *lion* N type

peg d² = *statue* N type

c¹ = "shape" properties

Y = "stone lion" sense

c² = central-essence and shape properties

The stone lion can be categorised as a lion, since the latter is an indiscernable approximation to a fake gun with respect to shape properties, and the latter subsume the properties in the "stone lion" sense; the shape properties result from the intersection of the application of the shape pre-property to "lion" and "statue"; the statue is an indiscernable approximation to the stone lion with respect to both central-essence and shape properties. Hence the criterion with respect to which the lion is an indiscernable approximation to the stone lion subsumes that criterion with respect to which the statue is such an approximation. And, as with Negating Proper Privatives, implicit in the last point is the characterisation of the second categorisation (of the stone lion as a statue) in terms of the AC' schema.

Hence, the NP sense defines perspective-relative categorisations of the stone lion: as a lion and a statue, respectively. It could also support other, less extreme, perspectives on the stone lion as a statue (i.e., ones comprising non-central-essence properties of "statue").

Once again, the NP sense can support only a limited set of perspectives. If one were to attempt to categorise the stone lion as a lion in any sense other than its shape (for example, if one were counting the number of mammals in the park), the categorisation would fail. For this reason, it is inappropriate for Fred to reply "yes" to the zoo-keeper. He would be attempting to apply a criterion of indiscernability which outstrips that which actually obtains, thus over-generalising the categorisation, beyond the perspectives under which it can hold.

V.1.3. An Account of Equivocating Privatives

For Equivocating Proper Privatives, I will consider a Default Affirmative and a Default Privative concurrently (since they differ only in the A used); for Equivocating Functional Privatives, however, I will consider the two types in succession.

V.1.3.1. Equivocating Proper Privatives:

The examples I shall present are *apparent friend* and *probable friend*, as Default Privative and Default Affirmative, respectively. Both types of Equivocator behave in the same way during the first two stages of Sense Generation.

apparent:

Undercutting of N's central-essence properties; default assumption that further information would result in their Type II Rebuttal;

Ascription of the result the of application of "Behaviour" pre-property to N, to NP sense.

probable:

Undercutting of N's central-essence properties; default assumption that further information would result in a reversal of the Undercutting and an ascription of the properties to the NP sense;

Ascription of the result the of application of "Behaviour" pre-property to N, to NP sense.

friend:

Central-essence properties: loyalty; real affection; no dubious motivations;

Central-diagnostic properties: solicitous behaviour; tolerance, willingness to take time; various vague kindnesses;

Non-central properties: gifts; blandishments; overt displays of affection;

Stage I: Combination

This has two components. The first is the defeat of the N's central-essence properties by the A. If we consider the schema for Undercutting,

P = use of term *friend*;

Q = ascription of central-essence and central-diagnostic properties to the N concept (and thus to the NP sense);

R = the application of *apparent* or *probable* to *friend*.

The application of either A to *friend* results in the A Undercutting the central-essence properties of "friend", severing the connection between P and Q. Hence, we cannot infer either $\neg Q$, nor $\neg P$. That is, we cannot deny outright the ascription of the central-essence properties of "friend" to the sense: rather, we just cannot make the ascription on the basis of current information; nor can we deny the coherence of the application of the term *friend* to the entity in question. The central-essence properties of the head N are thus placed in doubt, pending the availability of further information. Once again, the independence of central-diagnostic from central-essence properties means that a defeat of the latter does not defeat the former.

The second component is the application of the pre-property "behaviour" to the concept for *friend*. The pre-property selects those central-diagnostic properties that are related to the behaviour of a friend (leaving out, for example, having a mental attitude of tolerance); and those non-central properties that come under the current thematic interest and which relate to the behaviour of a friend. Applying "behaviour" to "friend" produces a more precise pre-property which can be labelled "behaviour of a friend", which is ascribed to the NP sense. We cannot, at this stage, say in precisely which way the apparent or probable friend manifests his friendly behaviour.

In constructing a peg for either *apparent friend* or *probable friend*, we have ascribed to it the "behaviour of a friend" pre-property. As a result, the peg (as an alec) might play various roles, and can be categorised as a member of the N types of those roles under appropriate perspectives. Various N's will be indiscernable approximations to the NP, with respect to the criterion of the pre-property. For example, a friend, careerist, social incompetent, or a psychopath, are all indiscernable approximations to "apparent friend" and "probable friend" peg with respect to "behaviour of a friend". For each of these N's, the NP peg is an indiscernable approximation to the concept (thus ensuring that we can view the roles as being played by the same peg: he is a careerist with the "behaviour of a friend" pre-property, as well as other personality characteristics). Each of these N's could provide specifications within the pre-property. In categorising the NP peg as a member of the various role N's, the latter must subsume the former, with respect to the operant criterion. The following categorisations are warranted:

The apparent/probable friend is a friend with respect to behaviour

The apparent/probable friend is a careerist (friend) with respect to behaviour

The apparent/probable friend is a social incompetent (as a friend) with respect to behaviour

The apparent/probable friend is a psychopath (regarding friendship) with respect to

behaviour

These are all categorisations under Type II Perspectives. Again, the first categorisation satisfies categorisation schema PC'', where

peg a = *apparent/probable friend* discourse referent;

peg/role d¹ = *friend* N type;

NI possible roles d², d³, d⁴ = *careerist, social incompetent, psychopath* N types;

Y = "apparent/probable friend" sense;

c = "behaviour of a friend" criterion;

"Friend" is an indiscernable approximation to "apparent/probable friend" with respect to behaviour; behaviour subsumes the properties ascribed to the NP sense; the intersection of the application of the pre-property of "behaviour" to "friend" and any of "careerist", "social incompetent" or "psychopath" would result in a peg which is an indiscernable approximation to "apparent/probable friend" under the same "behaviour" criterion as before.

The categorisations regarding the NI's satisfy schema AC': where a = "apparent/probable friend" discourse referent, and d¹'s = each of "careerist", "social incompetent" or "psychopath", we can say again that each d¹ is an indiscernable approximation to the peg with respect to the "behaviour" criterion.

Hence, if the construction of the NP sense terminated at this point, the sense would define a Type II Perspective on the apparent/probable friend as a friend, and support various Type II Perspectives on him as a careerist, a psychopath, and a social incompetent. If, however, there was sufficient requirement, we might extend the sense through Stage II.

Stage II: Extension by Implicit Attachment:

The extension of the sense here is constrained by our not having determinate information concerning the central-essence properties of the head N, and hence of the NP. In order to provide a specification within the pre-property, the peg must be instantiated.

The first constraint on the instantiation is that it must have properties which subsume the "behaviour of a friend" pre-property. Candidate instantiations include as a "careerist friend": a deliberate deceiver, acting friendly in order to further his own career; secondly, as a "social incompetent (as a friend)": the person, through lack of social skills, is ill-equipped to present a consistent manner of behaviour: friendly behaviour may not indicate friendly feeling; and as a "psychopath (regarding friendship)": the person is psychopathologically unable to behave honestly, and simply behaves in a friendly way towards anyone. Notice that each of the instantiation candidates for Equivocators are specifications of the

general Negating Privative which could be expressed as *false friend*.

The second constraint is that the instantiation must be able to fulfill the roles of the peg in the discourse scenario. Suppose the scenario being described in the discourse concerns someone who is a highly ambitious employee of a company, for whom it would be advantageous to cultivate the support of an immediate superior in the hierarchy. In this case, the most plausible instantiation would be as a "careerist friend".

The NI is the concept which is superordinate to the category of the instantiation: in this case, it will be *careerist*, whose lexical concept is the base for analogical transfer, and the NP sense (construed in the light of the instantiation: the apparent/probable friend has the "behaviour of a friend" pre-property in the context of his role in the company) will be the target.

The matching of the base and target is constrained by our lack of information concerning the central-essence properties of the head N and NP. Since they are still Undercut, they cannot be employed in the analogical mapping. So the match is effected with respect to properties of the same level as the pre-property (i.e., central-diagnostic), or ones more shallow (non-central properties). The central-diagnostic properties of the NP and NI exhibit a good match: the behaviour of an apparent/probable friend (who is career-oriented), and that of a careerist. Hence, we can transfer central-diagnostic properties from the concept "careerist" to the NP sense. Once again, we can view this as a transformation of the head N properties through the analogy with the NI: the apparent/probable friend manifests the behaviour of a friend only as would a careerist. For example, the vague acts of kindness might be too extreme, or behaviour might be so solicitous as to be cloying: he manifests friendly behaviour taken to extremes. Transfer, then, results in a specification of the pre-property "behaviour of a friend", by taking the intersection of this pre-property and the result of the application of the "behaviour" pre-property to "careerist" ("behaviour of a careerist"). This produces the more precise properties we might label "behaviour of a friend as a careerist" or "behaviour of a careerist as a friend". Transfer might also include non-central properties of "careerist".

At the end of Stage II we have the same situation for Default Affirmatives and Default Privatives. The NP peg has been ascribed a specification of the pre-property ascribed in Stage I, in addition to other non-central properties; the central-essence properties remain Undercut. The peg can then play the role of two distinct N types, under which it may be categorised relative to perspectives. As in Stage I, these are Type II Perspectives, since

they do not incorporate the central-essence properties of the categorising N. Each N is an indiscernable approximation to the NP peg with respect to the criterion of "behaviour of a friend as a careerist". These may be paraphrased as:

The apparent/probable friend is a friend with respect to behaviour as a careerist

The apparent/probable friend is a careerist with respect to behaviour as a careerist

Although the criterion of indiscernability is of the same coarse level as in the categorisations at the end of Stage I (i.e., central-diagnostic level), it is a more detailed specification of that criterion. As with the same categorisation at the end of Stage I, the first categorisation satisfies the condition of categorisation schema PC'' (here, only one d^1 , *careerist*, is relevant). And the second categorisation satisfies schema AC'.

The NP sense at this stage allows us to be quite specific concerning those aspects of the entity about which we do have information; this is compatible with the incompleteness of our information regarding the entity as a whole. If there is an appropriate informational requirement, we might inquire further about the person, thus extending the sense on the basis of more contextual indicators. However, because the A's are marked for their default tendency, we might also extend the senses according to these tendencies, even where our inquiry cannot satisfy the informational requirement. So although an accurate extension does require further information about the NP entity, this is not necessary for the process of extension to take place. It is at Stage III that the difference between Default Affirmatives and Default Privatives becomes marked.

Stage III: Further Extension by Context

This stage is crucial to whether the NP is understood privatively or affirmatively. The Type II Perspective on the apparent/probable friend as a careerist (at the end of Stage II) is consistent with a Type I Perspective on him as a friend. We may find either that the affirmative categorisation (as a friend) can be generalised to a Type I Perspective, or that the privative categorisation (as a careerist) can be so generalised. But we cannot coherently generalise both under a single state of information. An information-inquiry driven extension of the sense, then, may be either consistent or inconsistent with the default tendency of the A.

Default Privatives: "Apparent friend":

Consistent (Privative) Extension:

Suppose we find that the excessively friendly behaviour has been accompanied by distinctly non-friendly behaviour, such as disloyalty. We are then able to evaluate the central-essence properties of the head N; the Undercutting can be consolidated as a Type II Rebuttal of the

head N's central-essence properties. How does this come about?

The new information concerning the apparent friend is not consistent with the central-essence properties of the head N. Since those central-essence properties were already in doubt, this conflict is resolved through the primacy of the extension knowledge. Disloyalty is in conflict with the central-essence loyalty, and thus acts as a Type II Rebutter: we can then ascribe the denial of the central-essence properties to the NP sense.

This ascription allows us to then attempt to pursue the matching of the NI concept "careerist" (as the base) to the NP sense for "apparent friend" further than at Stage II. The non-friendly behaviour of the person can be seen as a manifestation of careerism in the context of friendship, so the matching can extend to central-essence level. As a result of the depth of the match, transfer of properties can occur at all levels. So the central-diagnostic properties of "friendly behaviour as a careerist" can be linked to the causal and motivational aspects of a careerist. Insincerity, lack of loyalty, and self-seeking aspects of careerism will be transferred.

The categorisation possibilities for the peg are the same as for the end of the second stage for Negating Privatives. The head N type is an indiscernable approximation to the peg with respect to the same criterion "friendly behaviour as a careerist" which forms part of a Type II Perspective. Under such a perspective, the apparent friend may be categorised as a friend:

The apparent friend is a friend with respect to behaviour, as a careerist

Again, this categorisation is permissible since it fulfills the requirements of schema PC'. The second possible categorisation is a Type I Perspective on the apparent friend as a careerist, satisfying AC':

The apparent friend is a careerist with respect to central-essence properties

Inconsistent (Affirmative) Extension:

Suppose we find that the person is, in fact, genuinely loyal and affectionate. We may then reconsider the Undercut central-essence properties of "friend", and reverse the Undercutting. The central-essence properties may then be ascribed to the NP sense.

This occurs, firstly, by noting that at least one of the central-essence properties of "friend" might hold for the apparent friend. This then allows us to forge a new analogy between the head N as the base and the NP sense as the target. Finding that the analogy can extend to central-essence properties supports a transfer of properties of this and other levels. Hence we can transfer properties of absence of dubious motivations, and so on.

Linking the central-essence properties of "friend" to the central-diagnostic properties of "friendly behaviour as a careerist" suggests two things. Firstly, we might generate some ad hoc explanation of why this person has been acting as if he might not be a real friend. Secondly, even though the sense can now support a Type I Perspective on the apparent friend as a friend, it seems that this odd behaviour would lead us to judge that he was, in fact, a somewhat atypical friend.

This affirmative outcome then results in the head N being an indiscernable approximation to the apparent friend with respect to central-essence properties:

The apparent friend is a friend with respect to central-essence properties

This satisfies the condition of schema AC'. And a Type II Perspective on the apparent friend as a careerist might be possible:

The apparent friend is a careerist with respect to behaviour

Default Affirmatives: "Probable friend":

Consistent (Affirmative) Extension:

This will behave precisely like the Inconsistent Extension of Default Privatives, resulting in a sense which can support a Type I Perspective on the NP entity as a member of the head N and a Type II Perspective as a member of the implicitly attached N.

Inconsistent (Privative) Extension:

An Inconsistent Extension of a Default Affirmative will behave exactly like the Consistent Extension of the Default Privatives.

V.1.3.2. Equivocating Functional Privatives:

The difference between the two sub-types of Equivocators is more marked in the functional case. The example of a Default Privative Equivocating Functional Privative will be *wooden frying-pan*; the example of a Default Affirmative will be *green lemon*.

Default Privatives: "Wooden frying-pan":

wooden:

Central-essence properties: solidity; mass; inanimacy; inorganicness; flammability;

Central-diagnostic properties: texture; colour; particular shape; hardness;

Non-central properties: whether sculpted; location; size;

frying-pan:

Central-essence properties: heat-resistant; non-permeable when heated;

inorganic; supporting function of frying.

Central-diagnostic properties: shallow circular; flat bottom; long thin handle; metal pan;

Non-central properties: size; whether "non-stick"; material for handle;

Stage I: Combination

Attempting to combine the lexical concepts for *wooden* and *frying-pan* Undercuts the central-essence properties of "frying-pan"; there is also an ascription of the result the of application of the "shape" pre-property to the N, to the NP sense.

Attempting to ascribe the properties of wood to the NP results in a Type II Rebuttal of the central-diagnostic property of material for "frying-pan": "wooden" defeats "metal". The link between the central-diagnostic and central-essence properties is contingent, so this defeat does not rebut the central-essence function of "frying-pan". The propagation of defeat is uncertain: we have insufficient information to determine whether the fact that the frying-pan is made of wood prevents the fulfillment of its basic function. Although there is a default tendency to assume that it will not be able to fulfill this function, there is nonetheless the possibility that, for example, the wood is specially treated so as to be heat-resistant. In which case, it might support the function. As a result, the Type II Rebuttal of the central-diagnostic property of material prompts an Undercutting of the central-essence properties of the head N.

The second aspect of the Combination Stage is the application of the pre-property of "shape" to the head N concept and to the modifier. This results in two more specific pre-properties: "shape of a frying-pan" and "shape of a wooden thing". Taking the intersection of these pre-properties, we arrive at a more specific pre-property: "shape of a frying-pan as a wooden thing". This means that the NP entity has frying-pan shape only as it is manifest by a wooden object; and it has the shape of a wooden object only as the shape of a frying-pan. This exactly parallels the situation for Negating Functional Privatives: the relatively detailed pre-property stems from the source of the defeat in the semantic interaction.

At the end of this stage, then, the peg for *wooden frying-pan* has been ascribed the "shape of a frying-pan as a wooden thing" pre-property, and can be categorised as a member of the N types of its possible roles under appropriate perspectives. Several N types will be indiscernable approximations to the NP, with respect to the criterion of the pre-property: for example, *frying-pan*, *sculpture*, or *drinking-bowl*. Each of these N's can provide specifications within the pre-property, indicating in precisely which way the "frying-pan"

pre-property is manifest in the wooden frying-pan discourse referent. Hence, the following categorisations are warranted:

The wooden frying-pan is a frying-pan with respect to shape as a wooden thing

The wooden frying-pan is a sculpture with respect to shape as a wooden thing

The wooden frying-pan is a drinking-bowl with respect to shape as a wooden thing

These are all categorisations under Type II Perspectives. The first categorisation statement complies with categorisation schema PC'. "Frying-pan" is an indiscernable approximation to "wooden frying-pan" with respect to shape, which subsumes the properties ascribed to the NP sense; the intersection of the application of the pre-property of "shape" to "frying-pan" and either "drinking-bowl" or "sculpture" results in a peg which is an indiscernable approximation to "wooden frying-pan" under the same "shape" criterion as before. Once again, the categorisation judgements regarding the NI's satisfy schema AC': $a =$ *wooden frying-pan* discourse referent, and $d^1 =$ either *sculpture*, or *drinking-bowl*. Each d^1 concept is an indiscernable approximation to the peg with respect to a $c =$ "shape".

Stage II: Extension by Implicit Attachment

Both "sculpture (that resembles a frying-pan)" and "drinking-bowl (that resembles a frying-pan)" will be potential instantiations for the NP. Both have properties which subsume the indiscernability criterion operant at the end of Stage I.

Suppose, then, that we have a discourse context in which the scenario takes place in an art gallery. Someone might say, "there's a wooden frying-pan in the next room". In such a context, the instantiation would be a sculpture that resembles a frying-pan. This is a more specific instance of the pre-properties ascribed to the NP sense. Hence the NI concept chosen will be "sculpture", as the immediate superordinate of the instantiation, and both of the constraints on selection of an instantiation are respected.

The NI, as the base, is then matched to the NP sense. As with Equivocating Proper Privatives, this match can only be effected at the central-diagnostic/non-central levels, since the central-essence properties of the head N (and hence of the NP) have been Undercut. The match holds at central-diagnostic level: the pre-property of "shape of a frying-pan as a wooden thing" (of the NP) will match with the central-diagnostic properties of "sculpture" as having a shape which is (partly) determined by its material. Hence, we can transfer central-diagnostic properties from the concept "sculpture" to the NP sense. Once more, this is a transformation of the head N properties through the analogy with the NI: the wooden frying-pan has the shape of a frying-pan only in the way that a sculpture would have that

shape. For example, the proportions might be inapt to a real frying-pan, or the finish need not be so smooth. The transfer specifies the pre-property "shape of a frying-pan as a wooden thing". Taking the intersection of this pre-property and the result of the application of the "shape" pre-property to "sculpture" ("shape of a sculpture"), produces the more precise properties, "shape of a frying-pan as a sculpture", or "shape of a sculpture as a frying-pan".

So the NP peg has been ascribed a specification of the pre-property ascribed in Stage I, in addition to other non-central properties. The peg can play the role of two distinct N types, as a member of which it may be warrantably categorised relative to appropriate perspectives. As in Stage I, these are Type II Perspectives. Each N is an indiscernible approximation to the NP peg with respect to the criterion of "shape of a frying-pan as a sculpture". These may be paraphrased as:

The wooden frying-pan is a frying-pan with respect to shape, as a sculpture

The wooden frying-pan is a sculpture with respect to shape

Although this criterion is of the same depth or level as in the categorisations at the end of Stage I (central-diagnostic), it is a specification of that criterion. As with the same categorisation at the end of Stage I, the first categorisation statement satisfies the condition of categorisation schema PC'' (again, only one d^1 , "sculpture", is relevant). And the second categorisation satisfies schema AC'.

If there is an informational requirement, we might attempt to find out more about the entity. Two possible outcomes are envisaged: Privative (extension consistent with default tendency) and Affirmative (extension inconsistent with default tendency).

Stage III: Further Extension by Context

Consistent (Privative) Extension:

Suppose further inquiry reveals that the entity is an unvarnished piece of wood in the shape of a frying-pan. This information concerning the wooden frying-pan is not consistent with the central-essence properties of the head N. Since those central-essence properties were already in doubt, this conflict is resolved through the primacy of the extension knowledge. The latter thus acts as a Type II Rebutter: we can then ascribe the denial of the central-essence properties to the NP sense.

This ascription allows us to then pursue the matching of the concept for NI *sculpture* (as the base) to the NP sense for *wooden frying-pan* further than at Stage II. The fact that the entity has the shape of a frying-pan is due to the central-essence properties of a sculpture

as a representation, so the match, and hence transfer, can extend to central-essence level. So representational, artistic and other properties of a sculpture will be transferred.

The categorisation possibilities for the peg are the same as for the end of the second stage for Negating Privatives. The head N type is an indiscernable approximation to the peg with respect to the criterion "shape of a frying-pan as a sculpture" which forms part of a Type II Perspective. Under such a perspective, the wooden frying-pan can be categorised as a frying-pan:

The wooden frying-pan is a frying-pan with respect to shape, as a sculpture the permissibility of this categorisation of the peg is evidenced in its fulfillment of schema PC'. The second possible categorisation is a Type I Perspective on the wooden frying-pan as a sculpture, satisfying AC':

The wooden frying-pan is a sculpture with respect to central-essence properties

Inconsistent (Affirmative) Extension:

Suppose we find that the object has been given a special, heat-resistant coating. This allows us to reconsider the Undercut central-essence properties of "frying-pan", and to reverse that Undercutting. So we can ascribe the central-essence properties to the NP sense.

By noting that at least one of the central-essence properties of "frying-pan" might hold for the wooden frying-pan, we can then forge a new analogy between the head N as the base and the NP sense as the target. Finding that the analogy can extend to central-essence properties supports an appropriate transfer of properties. So we can transfer properties of being non-flammable, impermeable, and being able to support the function of a frying-pan.

One implication of a Default Privative's emerging as an Affirmative is that the NP entity is a highly atypical member of the head N category. For it was just as a result of the NP entity's not having a crucial central-diagnostic property (material) that the central-essence properties were undercut. And this defeat occurred through an ascription of a property highly atypical of the head N, to the NP.

The affirmative outcome, then, results in the head N being an indiscernable approximation to the wooden frying-pan with respect to central-essence properties:

The wooden frying-pan is a frying-pan with respect to central-essence properties This Type I Perspective satisfies the condition of schema AC'. And a Type II Perspective on the apparent friend as a careerist may be possible:

The wooden frying-pan is a sculpture with respect to shape

The last categorisation perspective is somewhat odd and non-natural, and may be an unlikely occurrence.

Default Affirmatives: "Green lemon":

green:

Central-essence properties: entity must be at least two-dimensional; must be visible;

Central-diagnostic properties: hue range; saturation; brightness; focal colour;

Non-central properties: connotations of the colour; particular shape of object; whether transparent or not;

lemon:

Central-essence properties: presumed internal "essence"; ontogenesis; organicness;

Central-diagnostic properties: oval shape; yellow colour; sour-sweet taste; size and proportion;

Non-central properties: particular size; country of origin; grain and texture of skin.

Stage I: Combination

Combining the lexical concepts for *lemon* and *green* results in an Undercutting of the central-essence properties of "lemon", and an ascription of the result the of application of the "appearance" pre-property to N, to the NP sense.

Ascribing the property "green" to the NP results in a Type II Rebuttal of the central-diagnostic property of colour of "lemon". Since the connection between the central-diagnostic and central-essence properties is contingent, this does not rebut the central-essence properties of "lemon". We have insufficient information to determine whether the colour green prohibits the entity from having the internal essence of a lemon. Although there does appear to be a default tendency to assume that it is really a lemon, it is possible that it is green because it is a different kind of fruit (a lime), or not a fruit at all (a plastic ball). As a result the Type II Rebuttal of the central-diagnostic property of colour results in an Undercutting of the central-essence properties of the head N.

Next, there is the application of the pre-property of "appearance" to the head N concept and to the modifier. This produces two more specific pre-properties: "appearance of a lemon" and "appearance of a green thing". Taking the intersection of these pre-properties, we arrive at a more specific pre-property: "appearance of a lemon as a green thing". Hence, the NP

entity has the appearance of a lemon only as it is manifest by a green object; and it has the appearance of a green object only as manifest in the appearance of a lemon.

To the peg for *green lemon*, then, has been ascribed the "appearance of a lemon as a green thing" pre-property. Several categorising N types will be indiscernable approximations to the NP, with respect to the criterion of the pre-property: for example, a lemon, a lime, or a plastic ball. Hence, the following categorisations go through:

The green lemon is a lemon with respect to appearance as a green thing

The green lemon is a lime with respect to appearance as a green thing

The green lemon is a plastic ball with respect to appearance as a green thing

These are all categorisations under Type II Perspectives. The first categorisation conforms with categorisation schema PC'', and the others with AC'.

Stage II: Extension by Implicit Attachment

Both "lime" and "plastic ball (shaped like a lemon)" will be potential instantiations for the NP. Both have properties which subsume the indiscernability criterion operant at the end of Stage I.

Suppose, then, that we have a discourse context in which two children are playing a game. One might say, "pass me the green lemon". It seems likely that the instantiation will be as a plastic ball that resembles a lemon. Hence the NI concept chosen will be *plastic ball*, as the immediate superordinate of the instantiation. Thus, both constraints on the selection of an instantiation are respected.

The NI is then matched to the NP sense. As with other Equivocators, this can only be effected at the central-diagnostic/non-central level, since the central-essence vproperties have been Undercut. Hence, we may transfer central-diagnostic properties from the concept "plastic ball" to the NP sense. There is a transformation of the head N properties through the analogy with the NI: the green lemon has the appearance of a lemon as it would be manifest in a plastic ball. For example, the proportions might differ from a real lemon, the texture and grain of the surface might be different, and certain details might be omitted (no leaf or stalk). The pre-property "appearance of a lemon as a green thing" is thus specified. Taking the intersection of this pre-property and the result of the application of the "appearance" pre-property to "plastic ball" ("appearance of a plastic ball") results in the properties, "appearance of a lemon as a plastic ball", or "appearance of a plastic ball as a lemon".

So we have ascribed to the NP peg a specification of the pre-property ascribed in Stage I, in addition to other non-central properties; and the central-essence properties remain Undercut. The peg can then play the role of two distinct N types, under which it may be categorised. As in Stage I, these must be Type II Perspectives. Each N is an indiscernable approximation to the NP peg with respect to the criterion of "appearance of a lemon as a plastic ball". These may be expressed as:

The green lemon is a lemon with respect to appearance, as a plastic ball

The green lemon is a plastic ball with respect to appearance

The criteria of indiscernability for each of these categorisations are specifications of their analogues from the end of Stage I. As with the same categorisation at the end of Stage I, the first categorisation satisfies the condition of schema PC'' (in this case, only one d^1 , *plastic ball*, is relevant). And the second categorisation satisfies schema AC'.

Two extreme outcomes of further inquiry are possible: Privative (extension consistent with default tendency) and Affirmative (extension inconsistent with default tendency).

Stage III: Further Extension by Context

Consistent (Affirmative) Extension:

Suppose we find that the object is not, in fact, made of plastic: that the children are playing a game with fruit; and that, having cut one open, the children find that it has a sour-sweet taste. This allows us to reconsider the Undercut central-essence properties of "lemon", and to reverse that Undercutting. So we can ascribe the central-essence properties to the NP sense.

We find that the "taste" central-diagnostic property of "lemon" might hold for the green lemon. This then supports an attempt to extend the analogy between the head N and the NP sense. Finding that the analogy can extend to central-essence properties supports an appropriate transfer of properties. So we can transfer properties of having a particular essence, ontogenesis, and being organic. In addition to this, some ad hoc hypothesis regarding why the lemon does not conform to the usual diagnostic property of colour will be constructed: for example, it might simply be unripe; or it might be a newly created strain of lemons, whose colour and texture are assumed to be more desirable (*designer lemons*, as part of a range of *lifestyle fruit*, perhaps).

Notice that, when the Default Privative emerged as an Affirmative, the NP entity was a very atypical member of the head N category. A wooden frying-pan is a very odd frying-pan and there is no obvious way of explaining this in terms of permissible operations on

ordinary frying-pans. In contrast, although the Default Affirmative "green lemon" is a somewhat odd lemon, this can be explained through operations on lemons or by appeal to our ordinary lay-theoretical knowledge regarding them and other fruit. So the other central-diagnostic and non-central properties of the head N can coherently be applied to the the Default Affirmative NP sense, whereas this would be more difficult for the Default Privative. Such an entity, then, is likely to be a more typical member of the head N category than is a Default Privative entity.

The affirmative outcome results in the head N being an indiscernable approximation to the green lemon with respect to central-essence properties:

The green lemon is a lemon with respect to central-essence properties

This Type I Perspective satisfies schema AC'. It does not appear to be the case that a Type II Perspective on the green lemon as a plastic ball would be coherent:

The green lemon is a plastic ball with respect to appearance

Inconsistent (Privative) Extension:

Suppose further inquiry reveals that the entity is made of plastic: throwing it against the wall makes it bounce back, unsquashed and unbruised. Chewing it produces an unpleasant taste, and results in being admonished by parents. This information is not consistent with the central-essence properties of the head N. Since those central-essence properties were already in doubt, this conflict is resolved through the primacy of the extension knowledge. The latter thus acts as a Type II Rebutter: we can then ascribe the denial of the central-essence properties to the NP sense.

This ascription then supports the attempt to pursue the matching of the NI concept "plastic ball" (as the base) to the NP sense for "green lemon" further than at Stage II. The fact that the entity has the appearance of a lemon is due to the central-diagnostic property that a child's plastic ball is often a representation, so the match, and hence transfer, can extend to central-essence level. So the representational, manufacture, entertainment and other properties of a plastic ball will be transferred.

The categorisation possibilities for the peg are again the same as for the end of the second stage for Negating Privatives. The head N type is an indiscernable approximation to the peg with respect to the same criterion "appearance of a lemon as a plastic ball" (a Type II Perspective criterion). Under such a perspective, then, the green lemon can be categorised as a lemon:

The green lemon is a lemon with respect to appearance, as a plastic ball

the permissibility of this categorisation of the peg is vouchsafed by its adherence to schema PC'. The second possible categorisation is a Type I Perspective on the green lemon as a member of the category of plastic balls, satisfying AC':

The green lemon is a plastic ball with respect to central-essence properties

V.2. On Modal and Pure Privatives

V.2.1. The Qualities of Modal and Pure Privatives

In this section, I examine "Modal" and "Pure" Privatives. If Equivocators are "less privative" than Negators, then Pure and Modal Privatives are both more privative than Negators, and Pure Privatives are the most privative. The paradigm case for Modal Privatives is "former"; for Pure Privatives, it is "imaginary". They represent effects originating in the modifier's semantic nature, and hence I do not think that there are any analogue Functional Privatives. However, as with Negating Privatives, it seems to be the case that, when combined with an appropriate N, Modal and Pure Privatives might have affirmative effects: for example, *imaginary fiction*.

I have characterised Negators and Equivocators in terms of two different components of their contribution to the NP sense. The first is the type of defeat of the head N's central-essence properties. The second is the ascription of a pre-property of the head N to the sense (so the central-diagnostic properties of the head N are not all defeated).

The chief contrast between Modal and Pure Privatives, on the one hand, and Negators and Equivocators on the other, regards the second component. Modal and Pure Privatives involve a defeat of the central-diagnostic properties, and they ascribe a relation between the NP entity and the N's central-essence properties, to the NP sense. Consider the combinations:

He is a former senator

He considered an imaginary fish

It is not obvious that either object should have properties or pre-properties of the head N. A former senator is a person who once was a senator; since *senator* is a member of the class of nomina agenta, there is no reason why the central properties of the role should be retained once one ceases to perform the function. In fact, it does seem that a former senator just is constituted by the fact that he or she no longer has those central (functional) properties. And notice that there do appear to be constraints on which kinds of N terms can coherently occur with Modal Privatives:

There is a former dog *?

There is a former puppy

There is a former man *?

There is a former child

There is a former poet

There is a former table

There is a former frying-pan

Where the combination is coherent, the NP object does not have any of the central-essence properties of the head N. And there is no reason why they should have any of the N's central-diagnostic properties; nor is there any clear reason to deny the central-diagnostic properties. A former senator may have some of the central-diagnostic properties of a senator, but this will be a result of peculiarities of content regarding the individual. And the independence of central-essence from central-diagnostic properties means that there is no reason why someone who is not, in the general case, a senator, should not look or behave a lot like one. So we might say that the central-diagnostic properties of the head N are Undercut in combination with a Modal Privative A. And it seems that the central-essence properties of the head N are defeated by a Type II Rebuttal. That is, although at the present time the properties are negated, this should not disallow the application of the head N to the NP entity. For it is part of the meaning of *would-be*, *potential*, *possible* and so forth, that the NP entity just may well have these properties in future; and where *former* is applied to a member of the class of nomina agenta, then there is again no reason why the NP entity should not have the head N's central-essence properties in future. For example, a *would-be king* is someone who may well have the central-essence properties of a king in future, though at the moment he definitely does not have them; and a *former poet* might become a poet again, whereas a *former child* cannot become a child again. It seems, then, that there are two different cases for Modal Privatives, according to the type of relation between the NP entity and the central-essence properties of the head N. For all future-oriented Modal Privatives, the relation must express the possibility that the head N's central-essence properties may be ascribed to the NP sense; this is also the case for past-oriented A's when modifying nomina agenta. For the past-oriented A's when modifying other kinds of N's (especially *child*, *puppy*, and other indicators of positions or states on a non-reversible dimension), it seems that the relation must express the impossibility of making the ascription. So once again we find that the content of the N can have an effect on the behaviour of the A. Whereas the contrast between Functional and Proper Privatives was that the privativity of the latter was independent of content, and that of the former was not, in Modal Privatives the fact of privativity is content-independent but the way that it behaves is not.

A final point about future-oriented Modal Privatives is the contrast with Equivocators. For Equivocators, we might be able to ascribe the central-essence properties of the head N to the NP sense, but the time-scale is relatively short; and the ascription is based not upon the NP entity's possessing them or not, but rather upon the describer's state of information about the NP entity's properties. An apparent friend may well have all of the central-essence properties of a friend at this present moment, but my describing him thus indicates that I have insufficient information to make the ascription. Future-oriented Modal Privatives, in contrast, definitely do not have such properties at the moment, and may or may not have them in the future. The latter, then, appear to carry an implication of a possibly long-term time-scale, unlike the former.

Pure Privatives embody a more extreme situation. Consider the following combinations:

There was an imaginary fish

There was an invisible fish

There was a fictional fish

All of these combinations involve entities which are mental representations, or at least ones to which no perceptible object in the world corresponds. These kinds of combination just seem to carry the implication that the categorisation of the NP entity as a member of the head N cannot be supported in a way which is external to the particular representation scheme. That is, whatever properties these representations do have, they are properties of representations only, and not of the entities represented. So neither the central-essence nor the central-diagnostic properties of the head N can be ascribed to the NP sense. For Pure Privatives, the appropriate form of defeat appears to be Type I Rebuttal: not only can we deny the central-essence properties, the application of the N term to the NP entity must also be denied. Again, the second component of the meaning of Equivocators and Negators fails here. However, we should note one exception to this: where Pure Privatives are employed in a scenario that involves the production of some artefact, they might behave like Negating Privatives; for example,

He drew the imaginary fish

where the appropriate NI type might be *drawing*.

It seems, then, that rather than directly ascribing the properties of the head to the NP, these A's ascribe some relation between the entity and the head N's properties. Modal Privatives ascribe a temporal relation between the NP entity as it is now, and the way it was in the past or will or might be in the future. The NP entity incorporates a Type II Rebuttal of central-essence properties, and an Undercutting of central-diagnostic properties, both of which it once had, or may or must have in the future. This kind of condition appears to

hold for A's like *former*, *would-be*, and the temporal readings of *possible*, *probable* and *potential*. For Pure Privatives, the relation is one of representation: the NP entity has properties which represent the central-essence and central-diagnostic properties of the head N, both of which have been defeated by the A. So the Pure Privative NP entity neither has, has had, nor will have, these properties of the head N.

An implication of the defeat of the central-diagnostic properties of the head N is that the Combination Stage imposes little constraint on the selection of an instantiation for the NP. One of the constraints on selection of the instantiation for Negators and Equivocators was that the central-diagnostic properties of the instantiation should subsume the pre-properties ascribed to the NP sense by the combination. Since this ascription fails here, so then does the constraint.

V.2.2. An Account of Modal Privatives

The example I will take is *former senator*; the account should also cover future-oriented Modal Privatives, and past-oriented ones; in the latter case, the only difference will be in the relations.

former:

Type II Rebuttal of N's central-essence properties;

Undercutting of N's central-diagnostic properties;

Ascription of relations between NP entity and head N central-essence properties, to NP sense: the fact of having had them in the past, and the possibility of their ascription at some unspecified point in the future.

senator:

Central-essence properties: elected to public role and duties; professional fulfillment of public service;

Central-diagnostic properties: highly developed public relations skills; high profile at public functions; influence in public and governmental affairs; referred to by title of "senator";

Non-central properties: smokes cigars; talks loudly; overbearing manner; apparent sincerity.

Stage I: Combination

The Type II Rebuttal of the head N's central-essence properties means that we can ascribe the negation of those properties to the NP sense, but this does not invalidate a coherent

application of the head N term to the NP entity. And the application of the head N to the NP entity (the categorisation of the NP entity as head N type) is then contingent upon the possibility of ascribing the central-diagnostic properties of the N to the peg. However, the central-diagnostic properties of "senator" are Undercut by the A: so we can neither deny them nor ascribe them to the peg under the current state of information. This corresponds to the intuition that, once having lost office, a former senator may or may not carry on in the public eye. Whether or not the central-diagnostic properties are ascribed to the sense will depend upon the particularities of the instantiation, in Stage II.

The final aspect of the Combination Stage is the ascription of the relation to the NP. In this case, we ascribe the possibility that, at some unspecified time in the future, the former senator may again acquire the central-essence properties of a senator. Which is to say, he may be re-elected.

So the peg has been ascribed the relation to the central-essence properties of the head N, as well as their denial. There are thus no clear categorisation possibilities for the NP peg: a Type I Perspective as a member of the head N is ruled out by the Rebuttal; and a Type II is ruled out by the Undercutting of the central-diagnostic properties. In order to add more content to the peg, the NP must be instantiated.

Stage II: Extension by Implicit Attachment

I have already commented that the Combination Stage places no constraint upon the selection of the instantiation of the NP. Hence, the constraints will stem from our lay theories concerning what kinds of qualities senators usually have, and their life after holding office; and the major constraint will derive from the discourse scenario. So there is a possibility that the strength of the constraint will vary according to the indicators in the context. Specifically, whether the NP entity can be ascribed the central-diagnostic properties of the head N will depend upon the instantiation. For example, if the context were:

The former senator tended his rose-beds regularly

then there would be little indication of what kind of an instantiation is appropriate; hence, there would be no obvious grounds for reversing the Undercutting of the central-diagnostic properties. If, however, the context were:

The former senator retired to his villa in the hills

there may be reason to assume that the central-diagnostic properties of the head N can be rebutted. Finally, if the context were:

The former senator gave his fifth speech of the week to the business community, to much acclaim

we might have reason to reverse the Undercutting of the central-diagnostic properties.

Suppose, then, that the context is the last one, and the instantiation is then as a highly-placed business-man involved in politics. The NI type would then be *highly-placed businessman*. So the base for analogical transfer will be the NI concept, and the target will be the NP sense: one who is a highly-placed businessman and who once held office and may again do so. Matching between the two may extend to central-essence level, as may transfer. At the central-diagnostic level, if the thematic interest were in his behaviour, then we could construct pre-properties of "behaviour of a highly-placed business-man" and "behaviour of a senator", whose intersection would be "behaviour of a business-man as a senator" or "behaviour of a senator as a businessman". So there is a transformation of the central-diagnostic properties of the senator through the analogy with a businessman; perhaps his public influence occurs solely through addressing the business community, and not the public directly. These properties are then transferred to the NP sense.

We now have a NP sense which comprises the central-essence properties of the NI, the intersection of the application of a "behaviour" central-diagnostic pre-property to the NI and the head N, the fact of having held office in the past, and the possibility of holding it again in future.

The peg as thus defined, then, may be categorised in the usual ways for a Negating Privative outcome. Since the NI "highly-placed businessman" is an indiscernable approximation to the NP peg, with respect to central-essence properties, the following categorisation, satisfying AC', obtains under a Type I Perspective:

The former senator is a highly-placed businessman with respect to central-essence properties.

And since the head N concept "senator" is an indiscernable approximation to the NP peg with respect to the "behaviour" central-diagnostic property, the following categorisation, which satisfies PC', holds under a Type II Perspective:

The former senator is a senator with respect to behaviour, as a highly-placed businessman

The important point is that, unlike Negators and Equivocators, support of the second categorisation is not a necessary part of the meaning of the A. Rather, as we have seen, it is entirely contingent upon the peculiarities of the content and information available in determining the instantiation. The instantiation as a businessman allowed the Undercutting of the head N's central-essence properties to be reversed, but had the context been the first possibility noted, then the NI (*gardener?*) would simply be irrelevant to the Undercutting,

and the categorisation correspondingly inappropriate.

V.2.3. An Account of Pure Privatives

Since my primary concern in this project is to consider some of the ways in which objects which are ostensibly not of a particular type can nonetheless be categorised as members of that category, through the aegis of publicly specifiable conditions which support perspectives, I will have little detail to offer concerning Pure Privatives. Pure Privatives do not support the kinds of perspectives which I have argued for: rather, the NP entities are in the domain of (mental) representations, whose irrelevance to the current enterprise has been commented on (Section IV.4.2).

Briefly, the situation is as follows. The example I shall take is *fictional fish*.

fictional:

Type I Rebuttal of N's central-essence properties;

Rebuttal of N's central-diagnostic properties;

Ascription of relation between NP entity and head N central-essence properties, to NP sense: the impossibility of the NP entity (as a representation) having the central-essence properties of the head N (as the represented object);

Ascription of relation between NP entity and head N central-diagnostic properties, to NP sense: the entity represents these properties in some way (for example, being a two-dimensional representation of a three-dimensional object); it cannot have these properties, however;

fish:

central-essence properties: presumed internal essence and ontogenesis; organicness;

central-diagnostic properties: shape and proportion; water-dwelling; scales; able to swim;

non-central properties: colour; sea- or fresh-water dwelling.

Stage I: Combination

The A prompts a Type I Rebuttal of the head N's central-essence properties. Not only can we deny those properties, we can also deny the coherent application of the head N to the NP entity. So we cannot categorise a fictional fish as a member of the category of fish. Similarly, there is a rebuttal of the central-diagnostic properties of the head N. So the categorisation of the NP entity as a member of the head N fails on two fronts: firstly, the Type I Rebuttal denies the application of the term; secondly, none of the central-diagnostic

properties can be ascribed to the sense. The final aspect of the A meaning is the relation of representation ascribed to the NP sense: the actual properties of the head N themselves cannot be ascribed to the NP entity. Rather, some representation, requiring some representational code, can be ascribed. Such a code, with an appropriate translation, must also be ascribed. This code will be derived from the appropriate instantiation (Stage II).

At the end of this stage, there are no clear possibilities for a categorisation of the entity which would be more informative than the NP itself. We might categorise the entity as a *mental representation of a fish*, but this is less helpful than *fictional fish*.

Stage II: Extension by Implicit Attachment

Once again, the constraint upon the selection of the instantiation derives from our world knowledge about fictional things, and especially from the information available in the discourse context. If we know that

He sat on the river-bank with his fishing-rod imagining a marvellous fish. The fictional fish weighed one hundred pounds.

the instantiation might be a huge, fantasy fish. The NI might then be *fantasy*. The match between this concept as the base and the NP sense as the target can extend to central-essence level. So the central-essence properties of the NI will be transferred; this will include the particular representation scheme/code. Similarly, we might apply the pre-property of "representation" to both the NI and the head N, resulting in pre-properties "representation as a fantasy" and "representation as a fish". The intersection results in a fish-representation as transformed through a fantasy: "representation of a fish as a fantasy". So these properties are ascribed to the NP sense. Notice that it does not have the properties of a fish as transformed through the analogy with "fantasy"; rather, it has the properties of a representation of a fish thus transformed.

At the end of this stage, then, "fantasy" will be an indiscernable approximation to the NP sense with respect to central-essence properties, thus supporting a Type I Perspectival categorisation of the fictional fish as a fantasy:

The fictional fish is a fantasy with respect to central-essence properties

However, since the fictional fish has none of the properties of a real fish, we cannot categorise it as a fish under a Type II Perspective. What we could do, however, is categorise it as a *representation* of a fish with respect to central-diagnostic properties, and hence under a Type II Perspective.

V.3. A Typology of Privative Combinations

The treatments advanced in this Chapter allow us to postulate a typology of Privative Combinations, according to the effects on the properties of the head N; and the kinds of perspective permissible. In noting the latter, I shall take account only of the perspectives on the head N and NI. Regarding the former, I will note only central properties.

Type A. Pure Privatives:

(i). Property ascriptions from head N:

Central-essence: Type I Rebuttal

Central-diagnostic: Rebuttal

(ii). Permissible Perspectives:

Stage I:

Type I: none

Type II: none

Stage II:

Type I: re. NI

Type II: none

Examples:

Imaginary, fictional, invisible

Type B. Modal Privatives:

(i). Property ascriptions from head N:

Central-essence: Type II Rebuttal

Central-diagnostic: Undercutting

(ii). Permissible Perspectives:

Stage I:

Type I: none

Type II: none

Stage II:

Type I: re. NI

Type II: (possible) re. N

Examples:

Former, would-be, potential, possible, probable

Type C. Negating Privatives:

(i). Property ascriptions from head N:

Central-essence: Type II Rebuttal

Central-diagnostic: Pre-property

(ii). Permissible Perspectives:

Stage I:

Type I: none

Type II: re. N and all $N_1, N_2, \dots, N_n \in \{NI\}$

Stage II:

Type I: re. NI

Type II: re. N

Examples:

C.1. Proper: *false, fake, sham, bogus, counterfeit, pseudo, ersatz, feigned*

C.2. Functional: *stone lion, plastic flower, ice rose, elastic ladder, square basketball*

Type D. Equivocating Privatives:

(i). Property ascriptions from head N:

Central-essence: Undercutting

Default Affirmative: central-essence will be ascribed

Default Privative: central-essence will be Type II Rebutted

Pure Equivocator: no assumption

Central-diagnostic: Pre-property

(ii). Permissible Perspectives:

Stage I:

Type I: none

Type II: re. N and all $N_1, N_2, \dots, N_n \in \{NI\}$

Stage II:

Privative outcome:

Type I: re. NI

Type II: re. N

Affirmative outcome:

Type I: re. N

Type II: (possible) re. NI

Examples:

Default Affirmative:

D.1. Proper: *presumed, probable, assumed, standard, expected, presupposed*

D.2. Functional: *green lemon, squashed lemon, soft knife, straight banana*

Default Privative:

D.3. Proper: *apparent, ostensible, seeming, manifest, reputed*

D.4. Functional: *blue orange, wooden frying-pan, square basketball, straight*

boomerang

Pure Equivocator:

D.5. Proper: *alleged*

The differences between the types can be ranged along a dimension of privacy, which I will outline at the end of the chapter. The important implication is that the account concurs with the standard treatment only regarding Type A. For all of the rest, at least some categorisation of the NP entity as a member of the head N is permissible.

V.4. On Attributives

The approach which has been used to describe the behaviour of Privatives can also be applied to Attributives, the second problem class. As we have noted, there are two basic readings of Attributives to be accounted for. The two readings for *attractive dog*, are:

- (i). The entity is a dog and is attractive for a dog
- (ii). The entity is a dog and is attractive in some other way

However, when such an Attributive is combined with one of the nomina agenta, a further possible reading arises. Consider *attractive ballet-dancer*; this has two readings parallel to the two above, as well as one in which *ballet-dancer* is construed not as regards the profession of the person, but as an activity:

- (iii). The person is attractive at ballet-dancing

This reading is potentially the most privative of the three, the first is the least. It seems that the last reading is consistent with either a denial or an ascription of the head N's central-essence properties. That is, there is a reading which is simply indeterminate on this issue; and it may be extended to allow for the denial of the central-essence properties. This expresses the intuition that one can be an attractive ballet-dancer, a good teacher, a successful writer, without being at the same time a professional ballet-dancer, teacher or writer. The A's then modify the performance of the actions, but not the profession, which latter may or may not be applicable to the person in question. This is best construed, then, as a form of Equivocator.

attractive:

central-essence property: ascription of pre-property to NP: picks out a subset of the properties of the head N;

ballet-dancer:

central-essence properties: trained and paid to dance ballet in public performance;

central-diagnostic properties: ability to dance ballet; style; skill; physical characteristics; stamina;

non-central properties: modern or classical dance; solo or group performer.

V.4.1. An Account of the First Reading:

Stage I: Combination

The combination begins with the ascription of the properties of the lexical concepts to the sense: the central-essence properties of the head N are ascribed, as is the pre-property of attractiveness (as the equivalence class of all attractiveness properties: attractiveness in style, skill, physical characteristics, stamina).

The peg incorporates the central-essence properties of the head N, and the equivalence class of central-diagnostic "attractiveness" properties. At this point, the specific way in which the ballet-dancer is attractive is not given: the NP sense is consistent with a number of different instantiations of this pre-property. The combination is an Affirmative combination, so the categorisation of the NP entity as a member of the head N can be effected with respect to a Type I Perspective criterion of central-essence properties:

The attractive ballet-dancer is a ballet-dancer with respect to central-essence properties.

This satisfies the condition of AC'. In order to know exactly how the ballet-dancer is attractive, we need to proceed to Stage II.

Stage II: Extension by Thematic Dimension

Following on from Bartsch, we can see that the pre-property of "attractive as a ballet-dancer" can be made more specific by the incorporation of some thematic dimension, which indicates to which aspects of the ballet-dancer's central-diagnostic properties the A applies. Bartsch argues that such dimensions are often given, explicitly or implicitly, in the discourse. On the current view, they may be provided by the instantiation of the NP. As with the Equivocators and Negators, the constraints on the selection of the instantiation is that it must subsume the pre-property from the first stage, and must be able to play the role of the NP entity in the discourse scenario. For example, the conversation between Rudy and his friends might have been preceded by the following:

Rudy says, "Most of these ballet-dancers are rather clumsy; they seem to lack grace."

Here, then, the instantiation of the person who is "attractive as a ballet-dancer" would be as the more precise person who is attractive as a ballet-dancer in terms of style. So we have

the property of "style" being highlighted as the thematic dimension. When we combine this with the pre-property "attractive", it produces a more specific pre-property, "attractive in style". Taking the intersection of this and the pre-property extant at the end of Stage I produces a more precise range of properties: "attractive as a ballet-dancer as regards style" or "attractive in style as a ballet-dancer".

At the end of this stage, then, the NP sense can support the Type I Perspectival categorisation of the attractive ballet-dancer as a ballet-dancer. It can also support the relativistic description of the ballet-dancer:

The ballet-dancer is attractive as a ballet-dancer, with respect to style

This does not imply that the person is attractive in any other way, either as a ballet-dancer or as a person. And this is precisely the intuition we wish to capture.

V.4.2. An Account of the Second Reading:

Stage I: Combination

This Stage for the second reading differs slightly from that for the first reading. We firstly have the ascription of the head N's central-essence and central-diagnostic properties to the NP sense. The pre-property "attractive" is also ascribed. However, it is not applied to the central-diagnostic properties of the head N; it is, rather, left as the rather unspecific pre-property.

Again, at the end of this stage, we are able to effect an Affirmative categorisation of the NP entity as a member of the head N, under a Type I Perspective.

Stage II: Extension through Implicit Attachment

Accessing the instantiation is, again, constrained by the scenario information. Had Rudy's conversation proceeded thus:

Rudy says, "I had thought that this ballet troupe had a reputation for having rather more beautiful ballet-dancers; it seems I was wrong."

then we might have instantiated the "ballet-dancer who is attractive in some unspecified way" as a "ballet-dancer who is attractive physically". So applying the pre-property of "attractive" to the concept of a person considered as a physical body, we get the pre-property, "attractive as regards physical appearance". This is then ascribed to the NP sense.

A more precise thematic dimension might be provided by knowledge of Reg's profession - that of a hairdresser. So we might apply "attractive" to "coiffure", yielding the pre-property

"attractive as regards coiffure". The intersection of this and the previous pre-property produces a set of properties we might express as "attractive appearance as regards coiffure" or "attractive coiffure as regards appearance". These ascriptions then allow us to make the relativistic property description:

The ballet-dancer is attractive as a physical person, with respect to coiffure

V.4.3. An Account of the Third Reading:

Stage I: Combination

The combination of *attractive* and the "activity" reading of *ballet-dancer* has two immediate effects. Firstly, the central-essence properties of the head N are Undercut: if the modification is of only the activity per se, then we have no grounds for ascribing the denial or the aptness of the central-essence properties to the NP sense. Secondly, the application of the pre-property "attractive" to the central-diagnostic properties of the N yields a more specific pre-property, "attractive at ballet-dancing", which is ascribed to the NP.

At this point, the N type of "ballet-dancer" will be an indiscernable approximation to the NP peg with respect to central-diagnostic properties. Thus we can categorise the attractive ballet-dancer as a ballet-dancer under a Type II Perspective, satisfying schema PC":

The attractive ballet-dancer is a ballet-dancer with respect to performance of the activity

Here, we might view the possible other N roles of the peg as including, for example, "actor (who occasionally dances)", "dancer (as a hobby)", and so on. These will also be indiscernable approximations to the NP entity under the operant criterion, thus supporting categorisations:

The attractive ballet-dancer is an actor (who occasionally dances) with respect to performance of the activity

The attractive ballet-dancer is a dancer (as a hobby) with respect to performance of the activity

The property description of the person will be indeterminate:

The ballet-dancer is attractive at ballet-dancing in some unspecified way

Stage II: Extension by Implicit Attachment

Determining the instantiation is guided by the result of the combination stage, and by the scenario. Both "actor" and "hobbyist" would be suitable instantiations. In an appropriate context, then, we might select "actor (who occasionally dances)" as the instantiation; and "actor" as the concomitant NI.

Also necessary is the determination of the appropriate thematic dimension. This must also be effected by the instantiation: recall the account of the first reading of attributives, in which the "style" dimension was utilised as a result of the instantiation as an attractive ballet-dancer in terms of style. This again occurs here, so the instantiation is as an actor whose ballet-dancing has an attractive style.

So the base is the concept "actor", and the target is the NP sense. Recall that, for Equivocators, our information-lack demands that the matching can only be effected at central-diagnostic level. Both have to perform, have appropriate skills and so forth: so there is a good match. The result is that the NP person is attractive in style of ballet-dancing as an actor. This occurs through first constructing the pre-property of "attractive style" from the thematic dimension and the A pre-property; this is then intersected with the "attractive in ballet-dancing" pre-property, to form the pre-property, "attractive in ballet-dancing as regards style"; and this pre-property is then applied to the NI central-diagnostic properties to form the more specific properties, "attractive in ballet-dancing as an actor as regards style". This has two implications: firstly, the person may not necessarily have an attractive ballet-dancing style as this would be manifest in a professional ballet-dancer. Secondly, this does not imply that the actor who has an attractive style as a ballet-dancer, has such an attractive style as an actor. The account does not rule out either of these possibilities, but it does not assume that they will hold by default. This is, I think, in accordance with our intuitions.

At the end of this stage, then, the NP sense comprises the detailed specification of the pre-property which it incorporated at the end of Stage I. As such, we can effect the same categorisations, under more precise Type II Perspectives:

The attractive ballet-dancer is a ballet-dancer with respect to style in performance of the activity, as an actor

The attractive ballet-dancer is an actor (who occasionally dances) with respect to style in performance of the activity

The property description of the person will be more specific:

The ballet-dancer is attractive at ballet-dancing with respect to style, as an actor

Stage III: Further Extension by Context

As with the Equivocators, there are two possible outcomes regarding the central-essence properties of the head N: Privative or Affirmative. The property description will not be altered in this extension, whilst the appropriate Type I Perspective will be defined. Since this follows exactly the same pattern as the Equivocators, it is unnecessary to go into this

Stage.

V.5. Summary: a Scale of Privativity

The accounts given of the problem classes allow us to range the categorisation phenomena attending upon them along a "scale" or "continuum" of privativeness. The fundamental dimensions along which all of the combinations considered vary are as follows. Firstly, the kind of effect that the combination has upon the head N's central-essence properties. Secondly, the kind of effect upon the head N's central-diagnostic properties. Thirdly, the type of perspective under which the NP entity can be categorised as a member of the head N type. The permutations of the various points along these dimensions result in the ability to place the combination types along a scale of privativity, from the affirmative at one end through to the most privative at the other. This is as follows:

**TypeA...TypeB...TypeC1...TypeC2...TypeD1...TypeD2...TypeD3...
TypeD4...Attributiveiii...Attributiveii...Attributivei...Predicating A's**

Taking the first issue, there is a transition from the definite ascription of the head N's central-essence properties to the NP (in Attributive readings i and ii), through the Undercutting of these properties (Attributive iii, D1, D2, D3, and D4), through their Type II Rebuttal (C1, C2, and B), to their Type I Rebuttal (A). Taking the ascription of the central-diagnostic properties, there is a similar gradation. From a direct ascription of properties (Attributive i and ii), through an ascription of pre-properties (Attributive iii, D1, D2, D3, D4, C1, and C2), through an Undercutting (B), to a Rebuttal (A).

The scale appears to indicate that, rather than being two discrete options, affirmativity and privativity should be viewed as gradable notions which shade into each other through the phenomena of attributivity. If the strictly affirmative interpretations incorporate only the Attributives i and ii (as well as the predicating A's), then we can see that privativity can extend from the Pure Privatives along to Attributive iii. This has an interesting implication. I have already argued that privativity can arise in a functional manner: it is not the exclusive province of a restricted type of modifiers, but can arise from the interaction of erstwhile affirmative modifiers and appropriate N's. But in the case of Attributives, we have the same combination giving rise to three discriminable interpretations, two of which are affirmative and one privative. So the privativity of Attributive iii is very much a matter of informational and contextual requirements. The possibility of having either an affirmative or a privative interpretation of a single combination, without any obvious intrusion of metaphor, seems to be a final blow to the assumptions of Sense Selection.

As noted in II.3, the typology of privatives, and the scale of privativeness, are independent of the proposed mechanism of concept combination. In order to evaluate the categorisation outcomes in this way, it would be necessary to consider the categorisation behaviours (permissible perspectives) predicted at the end of each stage of the combination mechanism. That is, the possibilities of adopting different categorisation perspectives under different informational states must be considered, with the question of whether these perspectives coincide with those provided in the typology.

CHAPTER VI

Some Implications, Connections, and Suggestions for Future Research

In this chapter, I consider some connections with other areas of ongoing research, suggesting some possibilities for future research and extensions of the current view. I also comment on some implications of the current project.

VI.1. Some Connections with Other Research

VI.1.1. Situation Theory, Constraints and Background Conditions

There are two ways in which the current view parallels some central conceptions of Situation Theory (Barwise 1989) Firstly, the notion of a constraint's being relativised to background conditions, thus supporting conditional inferences, parallels the idea of a C-relation. Secondly, judgements' being relative to such background conditions parallels the relativity of categorisation to a perspective.

Situation theory takes meaning as being essentially relational. This implies that much of meaning, and our knowledge of the world, must be local: it is relative to the obtaining of appropriate background conditions. In Situation-theoretic terms, such meaning may reside in conditional constraints. A constraint is a vrelation holding between two types of situation, whereby if we know that a token of the first type holds, we can infer that one of the second also holds. That is, situations of type 1 "involve" those of type 2. So a type 1 situation carries meaning in virtue of the relation between its type and the other type. This is expressed as:

$$S_1 \Rightarrow_2 S$$

Conditional constraints are expressed as:

$$S_1 \Rightarrow S_2 \mid B$$

where **B** = the background condition, such that, provided we have a background condition of this type, the constraint holds.

Constraints, then, are concerned with the extraction of information from a situation. If the background conditions hold, and if the constraint is one to which an agent is attuned and which forms part of her focus situation, then the agent will extract the information S_2 from

S_1 . Where a background condition is constant (all situations that arise are of type **B**), there is no reason why an agent should become aware of the dependence of a constraint on **B**. That is, within the confines of **B**, the constraint is viewed as absolute. So any defeat or failure of a constraint operates by Positive Undermining: we do not first check that all of the required background-conditions hold; rather, so long as the current situation is of appropriate **B** type, or can be assumed to be of that type, the constraint holds. **B** does not express all of the interdependent facts that might be necessary for a constraint to hold, but only the causally or conventionally pertinent ones. An example given by Barwise is that, if Claire (his daughter) rubs her eyes, then she is sleepy:

$$S_1 \Rightarrow S_2$$

where

$$S_1 = \text{Claire rubs her eyes}$$

$$S_2 = \text{Claire is sleepy}$$

but this constraint only holds given appropriate background conditions:

$$S_1 \Rightarrow S_2 \mid B$$

where

$$B = \text{a situation where there is no pollen X present}$$

where pollen X is present, then the constraint may fail.

There are clear parallels between this and the idea of a C-related inference. Firstly, a conditional constraint falls short of an entailment relation, but also appears to be stronger than an inductive relation when it holds. That is, it is defeasible, but confers certainty when the appropriate background conditions obtain. Secondly, the relation is an internal one: where the background-conditions hold, it is treated as an unmediated relation. It is treated as absolute, as are C-relations: they are phenomenologically, if not theoretically, law-like. Thirdly, being relative to background conditions, conditional constraints can support the two crucial kinds of defeat. Consider Undercutting. If we find, regarding the above constraint, that pollen X is in fact present, then we must deny the inference from Claire's rubbing her eyes to her being tired. This does not mean that we can infer that she is not actually tired, simply that we cannot do so in the current background conditions. And we cannot deny that she has rubbed her eyes. So it is that the connection between the two situations is defeated: we have denied that rubbing her eyes involves Claire's being tired under the current background conditions. Again, consider Type II Rebuttal. If **B** stated that pollen X is present and that it is midday, then we might have reason not only to deny that the constraint holds, but also that Claire would be tired anyway. So we have a Type II Rebuttal. And fourthly, the commitment to Positive Undermining again parallels C-related inferences.

It seems, then, that some of the main qualities of conditional constraints mirror those of C-relations. C-relations have been proposed as a relation between the conditions of application of a concept and its conceptual content (in the object individuation condition), and between a lexical concept and its concept extension (in the sense generation condition). By hypothesis, a Situation-theoretic account of concepts might be advanced, utilising conditional constraints in the place of these C-relations. Such an account forms an interesting avenue for future research, and is currently being developed by Braisby (1989).

A second possible parallel between the current view and Situation Theory concerns the perspectival-relativity of categorisation judgements. I have argued that, in order to evaluate a particular categorisation, we must take account of the operant perspective, relative to which categorisations may or may not be warranted. Hence, the operation of the C-relations is relative to the appropriate perspective. Which is to say that the background-conditions relative to which a categorisation is evaluated are just the description in the operant perspective. Only if this description is satisfied by the object to be categorised, can the categorisation go through. On this interpretation, then, perspectives form part of the background-conditions to a categorisation, but are nonetheless absolutely crucial to the information-flow. They are therefore understood (implicitly or explicitly) by the partners to a particular communication.

Two distinct views of perspectival relations have been advocated, by Barwise, and by Perry (Perry 1986) respectively. Intuitively, Barwise's view is as follows. As noted in Chapter II, we can think of certain relations as inherently perspectival - for example, *left-of*. Barwise maintains that the apparently two-placed relation *left-of* should really be seen as a three-placed relation, where the third role is filled by a parameter denoting the perspective of the agent. So *left-of* from one position will differ from *left-of* from another, just because of this extra parameter. However, Barwise suggests the idea of "projecting" a relation along an undiscriminated dimension in order to account for such perspectival useage. If we note the perspectival parameter as *r*, then the idea is that, where *r* takes the same value for every application of the relation in a situation, that *r* is not discriminated in that situation. Hence, we can project the relation into another relation of one less argument roles. That is, we have projected out the perspectival role: this role may as well not exist, since it is completely undetectable in that situation. And Barwise comments that the projected relation (which we might think of as the relation as interpreted by the agent who is unaware of the perspectival role) is entirely independent of *r*: the projected relation does not need *r* in order to hold.

Applied to the current view, the idea would be that the perspective under which the categorisation of a stone lion as a lion holds, is projected out, leaving the categorisation perspective-free. But this carries the important implication that the perspective is undiscriminated in the current situation. But, whilst some perspectives may well be implicit and undiscriminated, the ones I have considered must all be discriminated. For they are not just important to the information flow concerning the objects, they are also salient in that flow. For example, without knowing the kinds of perspectives required, Fred cannot reply to the zoo-keeper or the school-girl. It is only with respect to these perspectives that their questions, and Fred's responses, are coherent in the situation. So rather than the categorisation of a stone lion as a lion being independent of some projected-out perspective, the categorisation is in fact crucially dependent upon that perspective; for the categorisation to hold, and for communicative success, the perspectival description must be satisfied and must be discriminated. A further problem, to which I shall return below, is that the projection analysis appears to require a bivalent analysis of perspectival-relativity: either a relation is or is not relative to such a perspective; if it is not, then the perspective is projected out. That is, perspectives are all-or-none. This does not accord with the flexibility in my account; a "non-perspectival" categorisation, on my view, would be a Type I Perspective. But since this incorporates a description based upon the lay-theoretic accounts of a domain, this carries no presumption of either (absolute) correctness, nor of being non-perspectival. It is simply a description based on a more inclusive criterion.

The second approach to perspectival-relativity in Situation Theory is Perry's account of "unarticulated constituents". If we consider the categorisation judgement,

The stone lion is a lion with respect to shape, as a statue
which I have argued might underly a statement like

The stone lion is a lion

then the statement contains the articulated constituents, whilst the perspective description, "with respect to shape, as a statue", comprises the unarticulated constituent(s). That they are unarticulated is a result of their being non-explicit. Perry argues that we can have unarticulated constituents in thought as well as language: our mental representations can have constituents of which we are not, and perhaps could not be, aware. Perry assumes that the theorist of the particular domain is able to specify the contents and constituents of a particular thought, which includes aspects that no lay theory would include. So we somehow employ, as unarticulated constituents, facts of which we constitutively could have no knowledge; we know the unknowable. This reflects the strongly Realist version of Situation Theory adhered to by Perry, in contrast to the less stringent version of Barwise. The final point about unarticulated constituents is that they are to be represented as arguments in, for

example, a particular classification, in the same way as articulated constituents.

This version of perspectival relativity also has implications at odds with the current view. Firstly, the commitment to the Manifestation Constraint means that unarticulated constituents of thoughts are inadmissible: although we have knowledge of which we are not aware, such knowledge is not of things of which we could not be aware, nor knowledge detailed by a particular scientific theory. The current view of perspectives is tied to lay theories alone. Secondly, the representation of the perspectival constituents in the same way as articulated constituents, means that, although the former are given by the background they do not form a part of it. If we can represent unarticulated constituents in a categorisation judgement (in an "infor", in Situation-theoretic parlance) as part of that judgement or statement, then the implication is that we could obtain a conjunction of the perspective and the categorisation-information as premises for the inference to the categorisation-judgement. That is, the constraint schema

$$S_1 \Rightarrow S_2 \mid B$$

would become

$$S_1 \ \& \ B \Rightarrow S_2$$

which is precisely what we do not want, since it fails to express the relativity of the whole constraint to *B* (Barwise, 1986). It seems, then, that Perry's approach would also be an inapposite route.

A final point concerns the bivalent attitude to constraints. A constraint either fails or it does not: either the background-conditions obtain or they do not. This appears too inflexible to express the categorisation behaviour evidenced by *Attributives* and *Privatives*. If a perspective is the background-condition for a categorisation-judgement, and if perspectives can be of greater or lesser generality, then the background-conditions for a generalisable categorisation can obtain to greater or lesser degrees. That is, if I say, *That is a lion*, the background condition/perspective description may support the categorisation with varying degrees of generality. That is to say, depending upon the extent to which the perspective/background condition holds, we can make inferences of different degrees of generality. Which is another argument that the perspective is crucial to the information-flow on any occasion, and cannot be ignored.

It seems, then, that the parallels between a Situation-theoretic view and some of the main constructs developed here, are strong. The extant views on perspectival-relativity, however, are at odds with the requirements of my position. A more promising development, in that it

locates perspectives in background conditions, is that currently being pursued by Seligman (Seligman 1989).

VI.1.2. Formalised Nonmonotonic Inference

The position I have advanced is largely non-formal. However, the question arises as to whether the defeasibility mechanisms could not have been expressed in terms of formalised nonmonotonic inference. There are, indeed, parallels between the phenomena I have considered and some basic issues in formalised nonmonotonic inference.

C-related inferences are made on the assumption of "total evidence" (Baker, 1977): all relevant evidence is assumed to be taken into account before the inference goes through. This mirrors the "closed world assumption": as long as someone has no information to the contrary, she will assume she has all of the information relevant for a particular domain, as well as the negation of any item she does not have (Ginsberg 1987). For example, if I ask an agent for a particular ticket for a show, she might say that such a ticket was not available. This is not because she has an exhaustive list of all of the tickets that are unavailable; rather, she assumes she has an exhaustive list of all tickets that are available. As long as she has no information that the ticket is available, she will assume that it is not. Thus the onus of proof is on the challenger, as in C-related inferences. C-related inferences have been advanced as crucial to the relation of concept extension. They operate on the closed world assumption, to enable the ascription of properties, by default inheritance, to the sense. Such ascriptions are made so long as there is no information to the contrary.

A further important phenomenon in nonmonotonic reasoning is the "Yale Shooting Problem"; although this concerns reasoning about actions and change, it is relevant to the kinds of combinations I have considered. To see this, we need first to note the "frame problem": unless we have information to the contrary, we assume that things in general stay the same when we perform an action. For example, when I walk across town I assume that my hat stays the same colour, even though I cannot see it while it is on my head. The difficulty is to discern which aspects of the environment remain constant, and thus may be expressed by "frame axioms". Now, the Yale Shooting Problem. If I were to load a gun, aim at Arnold, and shoot it, we might have conflict between a reasonable inference and a frame axiom. For example, there might be an inference rule stating that, when we shoot a loaded gun at a person, the result is a state of affairs in which that person is no longer alive. But there would also be a frame axiom stating that Arnold's being alive is not affected by most actions, including my loading a gun and aiming it at him. We can view this as a situation's

giving rise to two incompatible "extensions" (new states of information): one in which Arnold dies, and one in which he remains alive. Similarly, we might say that Privative combinations might give rise to just such an incompatibility: one concept extension based upon the priority of the N content; and one based upon the priority of the A content.

There are two approaches to the formalisation of nonmonotonic inference: one that draws default conclusions in the absence of information to the contrary, and another that "minimises" the extent of a predicate or relation. The distinction is comparable to that in first-order logic between proof-theoretic and model-theoretic entailment. Drawing default inferences in the absence of negating information parallels trying to find a proof of the negation of a default inference; if no such proof exists, the inference is deemed consistent and assumed true. And a minimisation of p entails q if q holds in all minimal models of p .

Proof-based approaches include Default Logic (Reiter 1980), Nonmonotonic Logic (McDermott and Doyle 1980), and Autoepistemic Logic (Moore 1985). These views are based directly on the consistency of default inferences with the rest of the knowledge-base. They formalise a default like *lions eat meat* by formalising the more complex: *If Leo is a lion, and it is consistent that Leo can eat meat, then Leo can eat meat*. In addition, they have fixed-point semantics': in order to evaluate a particular inference, we need to construct a fixed point in the knowledge base. This has two aspects. Firstly, whatever form for default rules is used, the application of these rules to the knowledge base does not yield any new conclusions to be added to it; secondly, all first-order logical consequences have been drawn. Both aspects result in an augmentation, or extension, of the original knowledge base. Only once we have constructed a fixed point can a default inference be evaluated: if its support is consistent with the augmented knowledge base, it is assumed true. Within this over-all commonality, the three views have different ways of defining default rules to capture "consistency", and different resulting ways of explicating nonmonotonic inference. Reiter defines inference rules of the form:

$$\frac{a : b}{c}$$

where a is the precondition of the rule (e.g., *Leo is a lion*); b is the clause that is checked for consistency with the knowledge-base (*it is consistent that Leo eats meat*); and c is added to the knowledge-base if b is consistent (here, c will be equivalent to b). In contrast, McDermott & Doyle advocate a modal approach, in which the above default rule would be expressed as:

$$a \ \& \ Mb \rightarrow c$$

where M is a modal operator such that Mp means "maybe p ", or " p is consistent with

everything else that is known". As a result of difficulties in providing a clear semantics for the M operator, Moore replaced it with a necessity operator, L : Mp is then equivalent to $\neg L\neg p$. This expresses the intuition that, if Leo is a lion and we don't believe that Leo cannot eat meat, then Leo can eat meat. Now, Moore distinguishes between "typicality" and "beliefs about one's beliefs". The former involves concluding that Leo can eat meat because it is not known that Leo cannot eat meat, and "typically, lions can eat meat". This expresses the picture of Reiter and McDermott & Doyle. In contrast, the latter involves concluding that Leo can eat meat on the basis that it is not known that Leo cannot eat meat, and "I would know it if Leo could not eat meat". The difference is that Moore's form of nonmonotonic reasoning is meant to be sound: the conclusions drawn are logically valid. If an inference fails, it is not because of a deficiency in the inference rule; rather, it will likely be because the premises were wrong. In contrast, McDermott & Doyle's logic draws defeasible conclusions: plausible inferences drawn on the basis of inconclusive evidence, in the absence of information to the contrary. They are not logically valid; we can be wrong in a way that does not depend upon a denial of the premises. Similarly, on Reiter's default inferences, I may find that Leo is not a typical lion; here, the belief that typical lions eat meat is not changed; rather, my evaluation of Leo's typicality alters.

Might these positions satisfy the Manifestation Constraint, and so account satisfactorily for the nonmonotonicity noted in the problem classes? One initial favourable point is that the proof-based approaches have difficulty in supporting transitivity and contraposition, as does the C-relation.

Another point is that, although Autoepistemic Logic appears to be the most successful of the proof-based methods, it is the one which mirrors least adequately the kinds of inferences I have postulated. Rather, the inferences involved in the problem classes are closer to those expressed in Nonmonotonic Logic; the typicality default of Reiter would comprise a subpart of the inferences in concept extension.

However, these somewhat hopeful comments are undermined by the nature of a fixed-point semantics. Despite the views' seeking to express the intuition of default reasoning (essentially, Positive Undermining), they depend, in the construction of the fixed point, on Negative Undermining. Constructing the fixed point amounts to examining the knowledge base for defeaters prior to making the inference. And this misplaces the emphasis. And note that, even were this plausible for narrowly circumscribed, hard molecular structures, it would be more difficult for more diffuse and holistic knowledge structures. The potential for degrees of specificity and depth of content ascribed to our representations in processing

casts further doubt on the fixed-point assumption.

If the proof-based views are inappropriate, what then of minimisation (McCarthy 1980)? McCarthy's "circumscription" formalisation describes default rules in terms of "abnormality". For example, a lion can eat meat unless it is abnormal in some way, which covers its having no teeth, being very young, or being a stone lion. Circumscription accepts as abnormal only those objects that are known to be abnormal. Defaults are expressed in first order logic:

Allx: bird(x) & ¬ abn(x) → canfly(x)

The abnormality predicate **abn** is expressed in the axioms of the knowledge base. The axiomatisation, **A**, then includes **abn: A(abn)**. If we consider any other predicate in the knowledge base, say **p**, the idea is that the minimisation of **abn** is defined as the assumption that **abn** is the smallest predicate for which **A(p)** holds. That is,

A(abn) & Allp: ¬[A(p) & p < abn]

where **p < abn** means that the extension (set of entities) of **p** is properly contained in the extension of **abn**. On minimisation, this can never be the case. In circumscription, the above formula is added to the knowledge base. There is no requirement for the solution of fixed-point equations. This, then, is one critical advantage of minimisation over the proof-based approaches. A further point is that, in accepting as abnormal only those objects that are known to be abnormal, circumscription attempts to flesh out the closed world assumption. So it addresses the general concern of C-related inferences.

However, both transitivity and contraposition find ready expressions in circumscription. The latter would mean that all non-meat-eaters are by default not lions; any lion that does not eat meat must then be an abnormal lion. Two points should be noted. Firstly, this kind of defeat reflected is too strong for the kinds of cases I have examined. Secondly, it gains its apparent intuitiveness from the **abn** predicate. But this predicate appears to be a rather blunt instrument to bring to bear on the intricacies of concept combination. For example, how would this predicate fall out as regards the denial of the different kinds of properties I have distinguished? For example, normal lions might live in Africa (a non-central property): are all non-African lions then abnormal? Most lions have a heart: would one with a pacemaker be abnormal? If so, is its abnormality commensurate with that of the non-African lions? A position on the effects of denying different kinds of properties (for example, on the ways in which the entity can be categorised by the appropriate nouns), which would be a detailed exposition of the **abn** predicate, is required. But if the **abn** predicate is thus delineated so as to allow many different kinds of abnormality, it is not obvious that this would then result in a predicate whose extension was minimal with respect to all of the

other predicates in the language.

So it is, then, that one avenue for future research would be to attempt to express the current position in terms of formalised nonmonotonic inference. The particular approach might be to address it through minimisation, by expanding upon the *abn* predicate in the way noted. As a result of doing so, we might then find that *abn* as regards one set of properties is not equivalent to *abn* regarding another set. This might then require us to consider defeats which do not flow from contraposition, and their effect on circumscription.

VI.1.3. Degree and Measure Adjectives

A central topic of this study has been the relativity of A meaning; I have claimed that a similar kind of relativity occurs in both Attributives and Privatives. I have not considered those A's in which a further type of relativity is often said to operate: degree and measure A's. Although the paradigm cases of Attributivity have been, for example, *attractive* or *good*, I have emphasised their attributiveness rather than their measure qualities. This emphasis stemmed from interest in the categorising type of a NP with the A as constituent. In this section, I will note some parallels between the current approach to the problem classes, and a view of degree A's Klein (1980). It will emerge that although we may detect a similarity in spirit, the approaches are at odds.

Klein's position uses supervaluations within a Montagovian intensional logic framework. His starting-point is to take degree and measure A's as vague predicates. That is, they are extensional, and their meaning, though "relative" in some sense, is not relative to the meaning of the modified noun. Consider the difference between

a good car

and

a fast car

Siegel has argued that the former requires knowledge of the meaning of *car* in order to know what the meaning of *good* is; the latter does not require comparable knowledge, she claims, because all we need to know is how fast things called cars usually go. And this is not a part of the meaning of *car*. The relativity of measure adjectives is thus held to be to the selection of a comparison class, which may or may not coincide with the extension of the modified common noun. What, then, is a comparison class? Klein argues that it is a subset of the universe of discourse which is selected relative to a context of use. It is represented not as part of the logical structure of A meaning, but as part of the contextual coordinate. An example will clarify this. One aspect of the supervaluation is a partial

context-dependent interpretation. If we consider any degree A (such as *tall*), Klein argues, then the interpretation of this A at any context partitions the universe of discourse, U, into three sets. The sets comprise denotations that are **positive**, **negative** and **neutral** with regard to the property *tall*; so the application of the A *tall* to members of the first and second groups will result in truth-values of 1 and 0 respectively; the third will be unvalued. However, supposing we are interested in the relative heights of two people in the neutral extension-gap, how could we evaluate them? This is achieved through completions of the partial models, determined by the partially defined characteristic functions. This also indicates the way in which the comparative may be defined in terms of the positive form of the A: if x is in the positive set with respect to U, and y is in either of the other two sets, then x is taller than y will be true. However, if both people are in the positive or negative sets, the situation is a little more difficult. And this is where Klein's comparison classes have a crucial role. He argues that we must "focus" on the particular set of which they are both members, using this subset of U, rather than U itself, as the comparison class. The A function then partitions this new subset; if this still results in x and y being members of the same subset of the comparison class, then we can iterate this focussing mechanism, applying the A to each new comparison class, until we have achieved a sufficient degree of detail to differentiate them.

Several crucial implications of this approach are pertinent. Firstly, the meaning of the A remains constant (it is the same function); all that alters is the comparison class relative to which it is evaluated. Secondly, if we seek to categorise a group of objects by the A *tall*, we may find that a particular object can be categorised as a member of the group of tall objects relative to a narrow comparison class, but cannot be so categorised relative to a broader class. A categorisation claim would be true in the first context, and false in the second. Thirdly, the evaluation of the A can be more or less specific; for example, if we categorise a group of people through the A *tall*, then we can apply the A to the whole group, resulting in an extension-gap; this may suffice for our purposes: we may want to apply only a coarse criterion of height, or we may be interested only in a person who is categorised by the first application of the A. Our categorisation could then be made more precise by the re-application of the A to the comparison class of the extension-gap. And so on. A fourth implication concerns the connection between the initial comparison class for A application, and those classes to which the A might be re-applied. Klein's claim is that the latter must always be proper subsets of the former; change in comparison class is thus a focussing on a class of less generality.

What parallels can we discern between this view and the current one? We should first note

that the commitment to realism is explicit in Klein's approach. However, his goal is not to provide an account of the psychological mechanism for understanding degree A's, so the Manifestation Constraint would be inappropriate. But this does mean that parallels must be ones in spirit and not in detail. The first crucial point is that the evaluation of the A is not just relative to the particular N which it is modifying (e.g., *people* as modified by *tall*), but relative to the particular N considered in a particular way. That is, the A is evaluated relative to the comparison class relevant for the N; but this comparison class may vary from one context to another. *Tall* evaluated relative to *people* may differ between a context of species-comparison (where the comparison class is the species), and say, one of objects-in-this-room comparison (the comparison class being the people in this room). In a parallel manner, I have argued that the evaluation of modifiers is relative not only to the particular modified N, but to that N considered in a particular way, as a member of an implicitly or explicitly attached type. A second parallel is the sense in which an overt categorisation might be true relative to one comparison class, and false relative to another. On my view, we have seen that a categorisation may be permissible relative to one perspective, but not relative to another. And perspectives involve descriptions derived from the N entity considered as a member of the implicitly or explicitly attached type. And thirdly, such evaluations of A's can be made more precise by the re-application of the A function to successively narrower comparison classes. In a similar way, I have adapted Bartsch's idea of the iteration of pre-properties, being successively applied to different ways of viewing the NP object, before combination.

Despite these parallels, there are also some differences in spirit. The first is the assumption that A's are best treated as predicates. This requires placing the comparison class in the contextual coordinate, and hence implies that the A's meaning does not vary across contexts (even though the evaluation of the function varies between different comparison classes). The view that I have put forward takes A's to be noun modifiers. I think that we might want to say that noun-modifier theories, although they result in a view whose operation is less transparent than predicate theories, do nonetheless have evidence to commend them. Hoepelman argues that English sentences with A's in predicative position have an "underlying source" that is more complex than their surface form. Consider cases of "conjunction reduction" involving A's:

John is good and bad *

John is a good cobbler and a bad darts player

John is good or good *

John is a good painter or a good sculpter

Here, it just appears to be clear that the A is modifying a common N. Further, Hoepelman notes cases of multiple modification by A's, in which the A's are obligatorily followed by a N or N's:

Ulysses is a big fat juicy greek *

Ulysses is a big fat juicy greek tomato

So it might seem natural to view A's as modifiers of N's. And the evidence of attributivity appears to caution against treating prenominal A's as predicates: there is not necessarily an equivalence between the following pairs:

Rudy is an attractive ballet-dancer

Rudy the ballet-dancer is attractive

Fido is a good dog

Fido the dog is good

And Potter & Faulconer found evidence that, whereas prenominal modifiers behaved according to Sense Generation, post-nominal modifiers acted more like independent predicates. The difference may, in formal terms, be more or less one of notation, as Klein (1986) argues; but in terms of cognitive mechanism it is a crucial difference. We are left, then, with the possibility that we treat A's either solely as N modifiers, or allow for different treatments depending upon whether they are pre- or post-nominal.

A second difference is in the kind of relativity involved: whereas Klein's view is that A's are evaluated relative to an extensional comparison class, mine has been that they are evaluated relative to the meaning of some N type. Klein's view echoes Siegel's, and thus implies that for different types of A there will be different types of relativity. The first conception would be for the relative readings of attributives: the meaning of the adjective is relative to the intension of the modified noun. The second conception is in the measure adjectives: the meaning of the adjective is relative to a comparison class (which may or may not coincide with the extension of the modified noun). There is then, an intensional and an extensional conception of relativity. These might be seen as parallel to the context-sensitivity of sense and of reference, noted in Chapter II. But it was also noted there that a proper consideration of the context-sensitivity of sense required an incorporation of referent knowledge - that the variability of sense might be strongly influenced by the instantiation of the combination in a context. And that the selection of an instantiation, as an indication of the flexibility of reference, was constrained by the concepts' content. If this is the case, then the strong assumption that we can have two radically divergent conceptions of relativity, in which the intensional and the extensional are completely independent, may be difficult to hold in any cognitively plausible theory.

A third difference is in the degree of flexibility permitted: I have already noted that a shift in comparison class involves focussing on a subset of the original comparison class. This has two aspects. Firstly, it implies that a shift will not change the sortal type under which the N is to be considered; that is, there may be no provision for the flexibility of implicit attachment. Secondly, it implies that a shift is always concurrent with an increase in specificity; that is, in shifting comparison class, we are focussing on a narrower class than before, and hence judgments can be more precise and detailed. It will be clear that the current view carries no commitment to either assumption.

Bearing in mind such differences, then, is there any way in which future research might incorporate the insights of Klein's approach into the current view? We could view degree A's as expressing pre-properties that combine with N's in order to specify detailed properties. So *tall* would express a pre-property, which would be specified differently for different N's: it would be evaluated differently for *insect*, *person*, *building*, and so forth. Rather than shifts in comparison class being to narrower classes, we could allow that shifts might occur across types. And the evaluation would not be with reference to an extensional class, but an intensional (in the sense noted earlier) concept. We could, then, allow that different contexts demand different instantiations, which then give rise to different implicitly attached N concepts. Each of these concepts would contain their own norm for size, and other degree properties. For example, it is part of the meaning of *tree* that it takes a certain characteristic range of sizes, as for *car* or *piano*; the instantiating of *large piano* might be a concert grand or somesuch. It is not immediately clear how the comparative might operate under the current view; I suspect a great deal of additional machinery would be necessary.

VI.2. Some Implications of the Current View

In this section, I draw out some of the implications of the current view for some of the issues broached in the first two chapters. I begin, though, with a speculation concerning the import of privativity and attributivity, as manifestations of the ability to adopt different perspectives on a single entity.

VI.2.1. Privativity, Attributivity and Language Success

A common argument made by proponents of semantic realism is based upon Putnam's notion of "language success" (Putnam 1978). The argument is used to advance two claims (Dowty 1979). The first is that, only on a realist view could we account for the way in which language aids our interactions in and with the world. And the second is that it does

this in such a way that meanings (qua intensions) are not, and need not be, mentally representable. Let us examine this argument.

Putnam argues that we must distinguish between theories of language understanding and theories of language success. The former is the province of psychology, and the latter that of a theory of reference. Language success is concerned with the contribution that our linguistic behaviour makes to the success of our total behaviour. Recall that realism assumes that a part of the determinate constitution of the world is a determinate word-world reference relation; this premise requires that there can be no role for cognition in a theory of reference, and correspondingly none in explaining language success. Consider two agents, **A** and **B**. The basic idea is that the use of language allows **B** to take advantage of the interactions with the environment that **A** has experienced but that **B** has not. Dowty comments,

the reception of a (true) linguistic expression from another speaker-hearer is a kind of short-cut direct interaction with the environment (thanks to the underlying language-environment correspondence and the assumption that speakers ordinarily utter only true sentences).

(Dowty, 1979: 378-379).

The argument is then that such language success confers evolutionary advantage on those species that possess language, as compared to non-linguistic species.

Let me be clear about the claim being made. It is that language success and language understanding are entirely independent, though complementary, domains; that language success is entirely explained by a theory of reference, and that the latter is completely separate from cognition. So if we can argue that a theory of reference requires a cognitive input for it to get off the ground, this would mean that the explanation of language success would also be linked to cognition. I think we can make such a case on both a synchronic and a diachronic/evolutionary analysis.

Let me take the evolutionary aspect first. It is often claimed that the kinds of terms whose use requires the least cognitive input are those often construed as rigid designators: natural kind terms and proper names. The Kripke-Putnam thesis of rigid designation is well known: that these terms pick out the same individual or type at every possible index, and hence that the intensions of such terms are not fixed by any common representations on the part of language-users. Agents do not have access to the essentialist definition of, say, *gold* or *lion*. However, Partee (Partee and Bach 1981) has argued that the underlying idea of the approach - of a causal-historical chain leading from an initial dubbing of entities to current use - itself demands cognitive faculties. That is, it is not possible to name a whole species

simply by baptising an individual member with a name; rather, there must be some supposition that an individual is a member of the species, and an intention to apply the name to the whole species. The latter is crucial in providing a theoretical similarity basis for extending the application of the term to new exemplars. The argument, then, is that the historical chain is crucially related to the abilities of language users.

In the synchronic case, the arguments of Chapter II indicated that a simple sense-reference relation, unmediated by world knowledge, cannot account for either the vagaries of reference, nor the context-sensitivity of sense. This is most clearly evident in the categorisation of Privative NP-type entities as members of their head N's. That is, the very notion of a perspectival relation between a word and an entity is ruled out by the realist view noted above; on this view, the relation between word and world holds determinately, and therefore without doubt. On a perspectival view, the relation is mediated by the appropriate description and covering sortal, and it holds only relative to these components. Again, then, the theory of reference demands interaction with a theory of understanding.

Two points must be emphasised, both relating to the assumption made by Dowty that speakers ordinarily utter only true sentences. This has two ramifications. Firstly, the notion of truth is a classical realist one, such the categorisation of an entity as a certain type, or of two entities as being the same, can hold if and only if all of the identity properties of the types in question are true of the entities. That is, a classical view of identity and truth means that a perspectival categorisation or identity statement is not permissible, since they would both involve descriptions that fall short of unique identity conditions. That is, my claim is that it is surely of crucial evolutionary import that people have the ability to adopt different perspectives on a particular entity, and can make predications and categorisations of that entity from such perspectives. And it is this kind of ability that underpins our flexible responses to the environment and to others, and this, rather than a rigid acknowledgement of a determinate set of facts, must be crucial to language success in Putnam's terms. Such a perspective-adopting ability is crucial in the development of a sense of subject and object - in arriving at a conception of there being an external world apart from one's own conception. Further, and this is the second point, even if the conception of truth were altered, it seems that the claim would be inappropriate. The assumption is that we only tell the truth, and this accounts for language success. But it is clear that we do not only tell the truth, and that this fact can also be of crucial evolutionary advantage for people who form part of a particular gene pool. That is, genetically non-identical groups can obtain advantage by deception. Language is as much a tool for misinforming as informing.

How might we view this? For successful communication, information is always relative to background conditions; on the current view, categorisation is thereby relative to a perspective. This means that the same statement can either inform or misinform according to whether the appropriate background conditions are satisfied. Consider Privatives. If I were to say, *there is a lion in the bushes*, then depending upon the perspective, I will making different categorisations of the entity in the bushes; and depending upon another agent's being aware of that perspective, she will gain information or misinformation from my statement. Having the perspective implicit facilitates rapid communication, but also leaves open the possibility of miscommunication and deception. For example, by leading the other agent to believe that the operant perspective is of, for example, a real lion, then we can lead them to act in particular ways for our own ends. Where the entity in the bushes is a statue, then she would derive misinformation, and could be deceived; where it is a real lion, then she has information, and can act accordingly. The ability to alter or left unsaid the background conditions relative to which one is making a statement is crucial to the art of manipulation, for literary, humorous or less honourable ends. Again, then, we can see that language success is not simply a matter of reporting the truth about the world; it is tied up with (mis)information, (mis)communication, and cognition.

It seems, then, that we can make a plausible counter-argument to the Putnam/Dowty claims, and hence argue that language success could be accounted for in anti-realism, and that the former crucially requires language understanding.

VI.2.2. Some Other Implications

A crucial implication of the current view is that many important semantic distinctions must be made at a level which cannot be reduced either to referents in the world, or to discourse or mental representations/concepts. They involve objects under guises, or pegs and alecs: they involve perspectives on referents. There are several important facets to this. Firstly, the interpretation of a N or NP will not simply be a denotation, nor a mental concept; rather, it will be a relation between the sense ascribed to the expression, and the referent to which it refers, through that sense, on the current occasion. And relating such a sense to a referent is part of adopting a perspective on that referent. This leads on to the second implication, regarding the ontology of any formal semantic system set up to account for categorisation in a cognitively plausible manner. This ontology must include perspectives: we must be able to make distinctions between contents and beliefs through distinguishing between perspectives under which beliefs or statements regarding an object are made. Only in this way will we be able to avoid making either too few distinctions (if we take account only of

referents) or too many (if we take account of every nuance of sense-ascriptions). Further, the definition of a perspective, as a relation between a description and a referent, and the generation of that description through the satisfaction of relations between lexical concepts and concept extensions, suggests that such an ontology would also need to find a place for relations. As is argued in Situation-theoretic approaches, relations should not be defined relative to prior and more basic objects; rather, relations are themselves as basic as objects. In the same way that relations and objects are dialectically interdefined (forming a local holism), so objects and perspectives are interdefined. That is, we cannot assert that either has semantic or ontological priority: there is no way of seeing objects outside of perspectives, and no way of adopting perspectives that do not involve objects. It is in this sense that anti-realism gives rise to what Kant labelled "empirical realism", regarding the objects of everyday experience: we have no way of viewing the world that does not include those objects, and no way of viewing those objects that is outside of our perspectival scheme (Peacocke 1986).

If the interpretation of N or NP is not a "naked" referent, but one under a perspective, this suggests a generalisation to other expressions. A perspective is simply that body of information that is taken into account in making judgements, relative to which the judgement is appropriate or not. There are two avenues for generalisation. Firstly, we can consider the interpretation of other assertions, and argue that these must be evaluated relative to an appropriate perspective. In this way, we can consider an appropriate perspective as conferring warrant on an assertion. For example, if I make a claim that would hold relative to certain background-conditions, and I am aware of the kind of circumstances under which those conditions might hold (rather than being aware of the actual conditions of their holding: see I.3.1), then that claim is assertible. However, if I am also aware that those conditions do in fact hold, then the claim is warrantably assertible. So matching a perspective to an assertion is critical if the assertion is to be warranted. If the claim made in the assertion outstrips the information incorporated in the perspective, and we nonetheless know what conditions would confer warrant on the claim, then it is assertible but not warranted. What is the implication of this? Recall that warrant has been proposed as an anti-realist replacement for truth. This means that the central semantic conception is crucially informational; it is concerned not with the state of affairs in the world, but with the interaction between an agent and that state of affairs. The arguments regarding verification, and the fact that the properties that comprise a perspective's description are publicly specifiable, suffice, I think, to avert any charges of solipsism (since we are tying meaning to verification) or psychologism (since meaning is also tied to cognition). Warrant, then, is relative to a perspective. Hence, we might construe the categorisation schemas AC', PC' and PC'' as stating

constraints on warrant conditions for the appropriate types of categorisation.

The second avenue for generalisation of the notion of a perspective also adds to the first. Here, we can expand the idea of the information taken into account, and relative to which judgements are made. Heretofore, I have considered only particular agents' having different perspectives at different times. If we consider the behaviour of such an agent through time, however, we might note qualitative as well as quantitative shifts in perspective: in addition to moving from Type I to Type II Perspectives within a particular range of knowledge, she might actually alter her knowledge, thus adopting perspectives of the same type, but with different contents. For example, as she becomes more interested in different kinds of wine, she might grow to distinguish different grape varieties, and then to distinguish particular growing conditions' producing differences in a single grape; and so on. There are two corollaries to this. Firstly, we can see that, without any modification, this generalises into a description of Putnam's notion of the division of linguistic labour. That is, different sub-groups of a society, as a result of their roles in that society, have different expertises. A goldsmith views gold in terms of its purity, its value, and so on; a miner in different terms; and an overdressed capitalist views it as a means of display. Each has his or her own range of knowledge concerning gold, sufficient to their own ends. And they will make assertions regarding gold relative to such a knowledge base. They can also, of course, adopt the perspective of any of the other groups in the society, in order to make particular judgements; they have, however, their own base-line. So warrant may be relative to a perspective as determined through a particular knowledge-base. And there is clearly no reason why such divisions should only be determined through different expertise; different groups in societies are defined for a bewildering array of reasons. Different perspectives are not then defined relative to their approximation of scientific essentialism. We can accept the division of linguistic labour, and the historical theory of reference, without adhering to Putnam's assumption that natural kinds have real, scientifically discernable ontological essences. That is, the division of linguistic labour is not a division of ontological access. This is the essence of Salmon's (Salmon 1982) critique of the Putnam-Kripke stance: that we cannot derive a commitment to metaphysical essentialism (or indeed any metaphysical commitment) from considerations of the epistemology and psychology of language use - unless the former are built into the latter as hidden premises, as Salmon suggests occurs in the Putnam-Kripke argument. Hence, we should consider the metaphysical and epistemological sides separately, and can support the latter whilst denying the former. The epistemological divisions in meaning, then, are not to be derived from metaphysical presumptions. Rather, divisions may occur along the lines proposed by Farr & Moscovici (Farr & Moscovici 1985) or Sperber (Sperber 1985). That is, we should consider an "epidemiology of

representations", grouped according to the intersection of a host of social, economic and cultural indicators. A second corollary of using different knowledge-bases to define different perspectives is that the central/diagnostic distinctions noted earlier might vary in content. That is, for a scientific investigator in an area, the range of defeasible diagnostic indicators for individuation may expand to cover most of the central properties as defined by a less sophisticated observer. Someone may, therefore, both partition a common range of properties differently, and add her own range of properties on the basis of her particular expertise.

And this links with an issue regarding Criterial relations. We may be able to discern a way of clarifying an apparent difficulty with the Criterion/Symptom distinction. Wittgenstein (1953) comments that it is often difficult to state whether a particular property is Symptom or Criterion, and that what appears to be Symptom on one occasion can seem like Criterion on another. This is often interpreted as a debilitating difficulty for the view. However, on the approach adopted here, this very fluctuation is a manifestation of the virtue of flexibility. I have already commented on the variation of central and diagnostic properties depending upon the perspective or expertise of an agent. This is one way in which the Criteria/Symptom distinction may apply to the same set of properties in different ways. Another way is that an agent can adopt, I have argued, Type I or Type II Perspectives; put simply, what is Criterial for a Type II Perspective may be Symptom for a Type I. So relative to a particular Type I or II Perspective, expertise may result in an agent's taking a different range of properties as Criterial; and comparing Type I with Type II, the same properties may be Criterial or Symptom. This, then, is another indication of the local nature of meaning-relations. The same content and the same term may be in different relations depending upon local requirements.

Having mentioned a generalisation of the notion of a perspective, we are able to note a slightly different view of the conferment of warrant. Heretofore, I have alluded to the evaluation of warrant with respect to the "facts of the matter". Now we can see that these "facts" may well be another perspective, whose purchase on reality is thought to be greater than one's own. And this is where the crucial distinction between recognising the *circumstances* of verification, and recognising simply *when* a verification might be made (see I.3.1.2), is important. Consider an itinerant shopper aspre, looking for a piece of jewellery made of gold. When he enters a jeweller's shop, he is warranted in asserting that those items labelled "gold" are in fact made of gold. This is simply because of his place in the matrix of social relations: someone with more expertise, who might themselves be able to actually *effect* a verification of *this is made of gold*, has already labelled it thus. So our

shopper is simply recognising when his assertion might be warranted. And again, this is a defeasible warrant: the jeweller might be unscrupulous. The idea, then, is that the "facts of the matter" might simply be someone else's perspectival description, more inclusive or incisive than one's own. In order, then, to allow for the warrant of a categorisation, the perspectival description must be evaluated against the broader perspectival description. And this means that we must allow for a host of different ways of relating and transforming our initial description so as to test it against the "facts" description.

A further implication concerns the notion of an "intensional" approach to meaning. I have advocated an intensional stance, insofar as this involves considering the content internal to a word's meaning, rather than focussing exclusively on the referents of words. However, this notion is not coterminous with the formal semantic notion of "intension" as utilised in, say, Montague Grammar. To see this, consider the basic idea of categorising an entity under a perspective. On an intensional logic approach, an intensional context produces an opaque context, which is the only alternative to extensionality and transparency. That is, if we cannot categorise an entity as member of a particular category on the basis of all of its properties, then we cannot do so at all. It cannot allow for a limited "quantifying-in": a Privative combination would define such a context, and would therefore produce the characteristic behaviours of privatives on the standard view. What I have claimed is that a view that takes perspectives as basic need not give rise to such a dichotomous approach to categorisation - that we can have combinations that are clearly not completely extensional (a stone lion is not a lion from every viewpoint), but which are also not opaque.

A final range of implications concerns the connections between concepts, senses, word-meanings, and lay theories of the world. The view I have advocated suggests that we should consider the four notions to be independent.

Concepts have traditionally been taken as the stable contents of thoughts; they are representations of objects that allow us to ignore minor differences and act according to continuities and invariants across different contexts. As such, they define the categorisation-conditions for being a certain type of entity, both for type-type and type-token categorisations. As representations of entities, they are the stuff in terms of which theories of the world are represented. So they are the stuff of knowledge-representation and thought.

The organisation of world-knowledge into hierarchical structures allow that concepts might also fulfill certain of the functions of lexical semantics. For example, they might support the basic lexical meaning relations of hyponymy and antonymy, which then determine

inferential connections, and thus define analyticity and necessity for the words so related.

As representations of the knowledge regarding types of referents, concepts have also been invoked as the representation of the meanings for the terms that ordinarily denote such referents. Concepts are thereby accessed or constructed whenever the associated word is used. And a concept thus constitutes the categorisation- or individuation-conditions of entities labelled by such a word, where such conditions are very often different from the conditions on being a particular type of entity, as represented in lay theories. They also, then, allow for flexibility and context-sensitive specificity.

We can isolate, then, three disparate functions often ascribed to concepts. The first is to represent the properties of (types of) referents; the second is to act as the meaning representation of a lexical item; and the third is to relate such items to referents, in categorisation. Most cognitive approaches to concepts have advocated or assumed a unitary view in that a single construct is to perform all three functions. This results in a tension between the three requirements: how can a single construct both represent our stable world-knowledge, and fulfill the flexibility of the mental representations of word meanings? Or again, how can we allow for the flexibility of categorisation-behaviour whilst assuming that this is based upon concepts that are stable? On the view advocated here, the functions are performed by three different constructs.

On the current view, a sharp distinction is made between the lexical concept and the sense or word meaning. Only in restricted, limit cases will the two be equivalent. The lexical concept, as I have argued, is the stable semantic contribution of the lexical entry of a word to the interpretation of that word (or a phrase of which it forms a part); it comprises central properties of the type of referent to which the N usually refers. As such, it is this which forms the stable core of thought about such entities, and comprises categorisation-conditions for a target entity to be categorised as of a type. Such properties, as well as forming the lexical concept for a word, then, also constitute part of the world-knowledge or lay theory regarding referents of the particular type. Use of the lexical concept automatically provides access to other, non-central properties regarding the type, as a result of the position and relations of the type in the lay theory. The first function are thus fulfilled by the lexical concept.

The second function is fulfilled by the word-meaning. That is, the relations of antonymy, hyponymy, and thus of necessity and analyticity, are defined over the relation between the perspectival description and the facts of the matter. Only those inferences that are

warranted by the facts will be supportable. So the inferences that follow from the warranted application of a term will define the analytic properties of the word's meaning. That is, those inferences that are made certain by the warranted application of the concept will be analytic of that concept. For the conditions that I have considered, the facts of the matter happen to be defined by the sense description. So, for these conditions, the analytic is given by the statements from perspective that are supported by the sense. So analyticity is here relative to the sense. This has far-reaching implications. senses and not lexical concepts. It means that the kinds of semantic inferences that are permissible, required and denied on any particular occasion, are determined by the generated sense, and thus may or may not correspond to those that might be supported by the lexical concept. This has a crucial implication that notions of analyticity and (semantic) necessity must be relativised to the particular local occasion of use, and its contextual parameters - since these are the very constraints on the generation of a sense. This sense may, on occasion, correspond closely to the lexical concept; in this case, the semantic relations will be those defined over the kinds of referents to which the term usually applies. This means that the advantage of the concept-as-word-meaning picture - that stable referent-relations underpin stable meaning-relations - is provided for on the current view, but only as one of many possibilities for the relation between the lexical concept and the generated sense. On other occasions, the sense diverges radically from the lexical concept, so that the inferences that we can make on the basis of the use or application of a term do not correspond with those regarding the term's referent. Clear examples of this are the Privative combinations. We might view this position as a move towards separating the analytic from the (ontologically) necessary: the latter may approximate the former only in the limit cases where the sense does not extend the lexical concept. And even here, "necessary" does not connote essentialism. Analyticity, then, flows from the notion of necessity noted in I.3.2: so any evidence making a superordinate certain is uncontroversible evidence for the subordinate. This is a clear implication of holding the semantic and the ontological separate. The sense, then, performs the second function, of the mental representation of word-meaning.

The third function, of relating the mental representation to the referent, or of providing the warranted categorisation conditions for a term, is fulfilled by perspectives. And the evaluation of a perspectival description yields the word-meaning, as the relation between the intended categorisation and the facts of the matter.

Two further implications of this division of lexical labour should be noted. Firstly, I have argued (II.2.7) roughly in favour of the idea that conceptual coherence is theory-driven (which applies to both lexical concepts and senses). A criticism of this view, advanced by

McCauley (McCauley 1987), is that it results in a vicious circularity of theory and concept. Theories are that which structure and relate concepts, yet those very theories are held to be made up of concepts. This criticism goes through only if we assume that words are always to be represented only by one sense; if this is not the case, then the circularity disappears; if the sense can be a radical alteration of the concept, then there just is no circularity. Secondly, I argued in favour of a non-compositional and non-monotonic account of concept combination (II.2.4). The arguments advanced still hold if we focus only on the lexical concepts involved. However, if we examine the picture with respect to senses, I think that a more straightforwardly compositional picture could well obtain. That is, once we have generated the senses either for words or phrases (where each subpart of an expression constrains generation of the other subparts), their manner of combination will be compositional. That is, the interpretation is compositional as regards the senses, but not the lexical concepts. A surface non-compositionality gives way to a deeper compositionality.

The implications considered in this final section illustrate the effect of taking the anti-realist Manifestation Constraint seriously, as a cognitive constraint on semantics. The rest of the thesis, similarly, aimed to integrate philosophical and cognitive-scientific requirements into a naturalised approach to semantics. It is hoped that the study as a whole contributes to such a development in the direction of semantic modesty.

References

- Aarts, J. and Calbert, P. (1979) *Metaphor and Non-Metaphor: The Semantics of Adjective-Noun Combinations*, 1. Tübingen: Max Niemeyer Verlag.
- Anderson, R. C., Pichert, J. W., Goetz, E. T., Schallert, D. L., Stevens, K. V. and Trollip, S. R. (1976) Instantiation of General Terms. *Journal of Verbal Learning and Verbal Behaviour*, **15**, 667-679.
- Appiah, A. (1984) Verification and the Manifestation of Meaning. *Proceedings of the Aristotelian Society Supplementary Volume*, **LIX**, 17-31.
- Armstrong, S. L., Gleitman, L. R. and Gleitman, H. (1983) What Some Concepts Might Not Be. *Cognition*, **13**, 263-308.
- Baker, G. P. (1974) Criteria: a New Foundation for Semantics. *Ratio*, **16**, 156-189.
- Baker, G. P. (1977) Defeasibility and Meaning. Chapter 2 in Hacker, P. M. S. and Raz, J. (eds.) *Law, Morality and Society: Essays in Honour of H. L. Hart*, 1. Oxford: Oxford University Press.
- Baker, G. P. and Hacker, P. M. S. (1984) *Scepticism, Rules and Language*, 1. Oxford: Basil Blackwell.
- Barclay, J. R., Bransford, J. D., Franks, J. J., McDonald, N. S. and Nitsch, K. (1974) Comprehension and Semantic Flexibility. *Journal of Verbal Learning and Verbal Behaviour*, **13**, 471-481.
- Barsalou, L. W. (1982) Context-independent and Context-dependent Information in Concepts. *Memory and Cognition*, **10**, 82-93.
- Barsalou, L. W. (1983) Ad Hoc Categories. *Memory & Cognition*, **11**, 211-227.

- Barsalou, L. W. and Sewell, D. R. (1984) Constructing Representations of Categories from Different Points of View. Emory Cognition Project •Report No. 2, Emory Cognition Project, Emory University, Atlanta, GA., 1984.
- Bartsch, R. (1986) Context-dependent Interpretation of Lexical Items. Chapter 1 in Groenendijk, J., de Jongh, D. and Stokhof, M. (eds.) *Foundations of Pragmatics and Lexical Semantics*, 1, pp1-27. Dordrecht: Foris.
- Bartsch, R. (1987) The Construction of Properties Under Perspectives. Report No. 87-08, Department of Mathematics and Computer Science, University of Amsterdam, Amsterdam, 1987.
- Barwise, J. (1989) *The Situation in Logic*, 1. Stanford, Ca.: Centre for the Study of Language and Information.
- Berlin, B. E. and Kay, P. L. (1969) In . Berkeley, California: University of California Press.
- Bolinger, D. (1967) Adjectives in English: Attribution and Predication. *Lingua*, 18, 1-34.
- Boyd, R. (1983) On the Current Status of the Issue of Scientific Realism. *Erkenntnis*, 19, 45-50.
- Braisby, N. (1989) Situating Word Meaning. Technical Report No. EUCCS/Working Papers Vol.3, Centre for Cognitive Science, University of Edinburgh, Edinburgh, 1989.
- Bruner, J. S., Goodnow, J. J. and Austin, G. A. (1956) *A Study of Thinking*, 1. New York: New York Science Editions, Inc..
- Burstein, M. H. (1988) Analogical Explanations Combining Inconsistent Mental Models. In *Conference of the American Society for Cognitive Science*, Montreal, August, 1988.

- Burstein, M. H. and Adelson, B. (1987) Analogical learning: mapping and integrating partial mental models. In *Conference of the American Society for Cognitive Science*, Washington, August, 1987.
- Carey, S. (1983) Semantic Development: the State of the Art. Chapter 12 in Wanner, E. and Gleitman, L. (eds.) *Language Acquisition: the State of the Art*, 1, pp347-389. Cambridge, Mass.: Cambridge University Press.
- Clark, H. (1983) Making Sense of Nonce Sense. Chapter 6 in d'Arcais, G. B. F. and Jarvella, R. J. (eds.) *The Process of Language Understanding*, 1. New York: Wiley.
- Clark, A. J. (1985) Anti-Realism and Recognition Capacities. *Philosophical Quarterly*, 35, 171-178.
- Clement, C. and Gentner, D. (1988) Systematicity as a constraint in analogical mapping. In *Conference of the American Society for Cognitive Science*, McGill University, Montreal, 1988.
- Cohen, B. and Murphy, G. L. (1984) Models of Concepts. *Cognitive Science*, 8, 27-58.
- Collins, A. M. and Quillian, M. R. (1969) Retrieval time from Semantic Memory. *Journal of Verbal Learning and Verbal Behaviour*, 8, 240-247.
- Craig, E. J. (1982) Meaning, Use and Privacy. *Mind*, 91, 541-564.
- Davidson, D. (1984) *Inquiries into Truth and Interpretation*, 1. Oxford: Oxford University Press.
- Donnellan, K. (1966) Reference and Definite Descriptions. *Philosophical Review*, LXXV, 281-304.
- Downing, P. A. (1977) On the Creation and Use of English Compound Nouns. *Language*, 53, 810-842.

- Dowty, D. (1979) *Word Meaning and Montague Grammar*, 1. Dordrecht: Reidel.
- Dowty, D., Wall, R. E. and Peters, S. (1981) *Introduction to Montague Semantics*, 1. Dordrecht: Reidel.
- Dummett, M. A. E. (1976) What is a Theory of Meaning? Part II. Chapter 3 in Evans, G. and McDowell, J. (eds.) *Truth and Meaning*, 1. Oxford: Clarendon.
- Dummett, M. A. E. (1978) *Truth and Other Enigmas*, 1. London: Duckworth.
- Erickson, T. B. and Mattson, M. E. (1981) From Words to Meanings: a Semantic Illusion. *Journal of Verbal Learning and Verbal Behaviour*, 20, 540-551.
- Farr, R. M. and Noscovici, S. M. (1985) *Social Representations*, 1. Cambridge: Cambridge University Press.
- Fodor, J. A. (1975) *The Language of Thought*, 1. New York: Thomas Crowell.
- Fodor, J. A. (1981) The Current Status of the Innateness Controversy. In *RePresentations*. Cambridge, Mass.: MIT Press.
- Fodor, J. A. (1987) *Psychosemantics*, 1. Cambridge, Mass.: MIT Press.
- Franks, B., Myers, T. F. and McGlashan, S. (1988) Defeasibility in Concept Combination: a Criterial Approach. In *10th Annual Conference of the Cognitive Science Society*, McGill University, Montreal, August, 1988.
- Garnham, A. (1979) Instantiation of Verbs. *Quarterly Journal of Experimental Psychology*, 31, 207-214.
- Geach, P. T. (1967) *Reference and Generality*, 2. Ithaca: Cornell University Press.

- Geach, P. (1967) Identity. *Review of Metaphysics*, 21, 3-12.
- Gentner, D. (1983) Structure-Mapping: a Theoretical Framework for Analogy. *Cognitive Science*, 7, 155-170.
- Gerard, A. B. and Mandler, J. M. (1983) Ontological Knowledge and Semantic Anomaly. *Journal of Verbal Learning and Verbal Behaviour*, 22, 105-120.
- Ginsberg, M. L. (ed.) (1987) *Readings in Nonmonotonic Reasoning*. Los Altos, Ca.: Morgan Kaufmann Publishers, Inc..
- Grandy, R. (1987) In Defense of Semantic Fields. Chapter 8 in LePore, E. (ed.) *New Directions in Semantics*, 1, pp259-281. London: Academic Press.
- Grayling, A. C. (1982) *Introduction to Philosophical Logic*, 1. Brighton, Sussex: Harvester.
- Grayling, A. C. (1985) *The Refutation of Scepticism*, 1. London: Duckworth.
- Grice, P. (1975) Logic and Conversation. Chapter 7 in Cole, P. (ed.) *Studies in Syntax*, Volume 3: *Speech Acts*, 1. New York: Academic Press.
- Half, H. M., Ortony, A. and Anderson, R. C. (1976) A Context-sensitive Representation of Word meanings. *Memory & Cognition*, 4, 378-383.
- Hampton, J. A. (1987) Inheritance of Attributes in Natural Concept Conjunctions. *Memory & Cognition*, 15, 55-71.
- Harman, G. (1986) *Change in View*, 1. Cambridge, Mass.: Bradford Books.
- Heim, I. (1982) The Semantics of Definite and Indefinite Noun Phrases. PhD Thesis, Linguistics, University of Mass..
- Hintikka, J. (1983) *The Game of Language*, 1. Dordrecht: Reidel.

- Hintikka, J. (1987) Truth-Conditional Meaning Theories. Chapter 7 in LePore, E. (ed.) *New Directions in Semantics*, 1, pp235-258. London: Academic Press.
- Hoepelman, J. (1983) Adjectives and Nouns: a New Calculus. In Bauerle, R., Schwarze, C. and Stechow, A. (eds.) *Meaning, Use and Interpretation of Language*, 1. Berlin: de Gruyter.
- Holyoak, K. (1985) The Pragmatics of Analogical Transfer. In *The Psychology of Learning and Motivation*, Volume 1, 1, pp59-87.
- Jackendoff, R. J. (1983) *Semantics and Cognition*. Cambridge, Mass.: MIT Press.
- Johnson-Laird, P. N. (1983) *Mental Models*. Cambridge: Cambridge University Press.
- Johnson-Laird, P. N. (1987) The Mental Representation of the Meaning of Words. *Cognition*, 25, 189-211.
- Jones, G. V. (1982) Stacks not Fuzzy Sets. *Cognition*, 12, 281-290.
- Kamp, J. A. W. (1975) Two Theories about Adjectives. Chapter 9 in Keenan, E. L. (ed.) *Formal Semantics of Natural Language*, 1, pp123-156. Cambridge: Cambridge University Press.
- Kamp, H. (1981) A Theory of Truth and Semantic Representation. In Groenendijk, J. A. G., Jansen, T. M. V. and Stokhoff, M. B. J. (eds.) *Formal Methods in the Study of Language*, 1, pp277-322. Amsterdam: Mathematical Centre Tracts.
- Kay, P. and McDaniell, P. (1978) On the Linguistic Significance of the Meanings of Basic Colour Terms. *Language*, 54, 610-646.
- Keil, F. (1979) *Semantic and Conceptual Development*, 1. Cambridge, Mass.: Harvard University Press.
- Keil, F. (1987) Conceptual Development and Category Structure. Chapter 7 in Neisser, U. (ed.) *Concepts and Conceptual Development*, 1, pp175-200. Cambridge: Cambridge University Press.

- Klein, E. (1980) A semantics for positive and comparative adjectives. *Linguistics and Philosophy*, 4, 1-45.
- Lakoff, G. (1987) *Women, Fire and Dangerous Things*, 1. Chicago: University of Chicago Press.
- Landman, F. (1986) *Towards a Theory of Information: On the Status of Partial Objects in Semantics*, 1. Dordrecht: Foris.
- Levi, J. (1978) *The Syntax and Semantics of Complex Nominals*, 1. New York: Academic Press.
- Levinson, S. C. (1983) *Pragmatics*, 1. Cambridge: Cambridge University Press.
- Luntley, M. (1988) *Language, Logic and Experience*, 1. London: Duckworth.
- Lyon, K. (1989) Context Effects in Categorisation. PhD Thesis, Psychology, University of Edinburgh.
- Marr, D. (1977) Artificial Intelligence: a Personal View. *Artificial Intelligence*, 9, 37-48.
- McCarthy, J. (1980) Circumscription: a form of non-monotonic reasoning. *Artificial Intelligence*, 13, 27-39.
- McCauley, R. N. (1987) The Role of Theories in a Theory of Concepts. Chapter 11 in Neisser, U. (ed.) *Concepts and Conceptual Development: Ecological and Intellectual Factors in Categorization*, pp288 - 310. Cambridge: Cambridge University Press.
- McDermott, D. and Doyle, J. (1980) Nonmonotonic logic I. *Artificial Intelligence*, 13, 41-72.
- McDowell, J. (1976) On 'The Reality of the Past'. Chapter 2 in Hookway, C. and Pettit, C. (eds.) *Action and Interpretation*, 1.

- McGinn, C. (1980) Truth and Use. Chapter 2 in Platts, M. (ed.) *Reference, Truth and Reality*, 1, pp19-40. London: Routledge & Kegan Paul. •
- Medin, D. L. and Shoben, E. J. (1988) Context and Structure in Conceptual Combination. *Cognitive Psychology*, 20, 158-190.
- Medin, D. L. and Wattenmaker, W. D. (1987) Category Cohesiveness, Theories, and Cognitive Archaeology. Chapter 3 in Neisser, U. (ed.) *Concepts and Conceptual Development*, 1, pp25-62. Cambridge : Cambridge University Press.
- Miller, G. A. and Johnson-Laird, P. N. (1976) *Language and Perception*, 1. Cambridge: Cambridge University Press.
- Minsky, M. (1975) A Framework for Representing Knowledge. Chapter 6 in Winston, P. H. (ed.) *The Psychology of Computer Vision*. New York: McGraw-Hill.
- Montague, R. (1974) *Formal Philosophy*, 1. New Haven: Yale University Press.
- Moore, R. C. (1985) A Formal Theory of Knowledge and Action. Report No. CSLI-85-31, CSL & I, Stanford University, Stanford, Ca., September 1985, 1985.
- Murphy, G. L. and Wright, J. C. (1984) Change in Conceptual Structure with Expertise. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 10, 144-155.
- Murphy, G. L. and Medin, D. L. (1985) The Role of Theories in Conceptual Coherence. *Psychological Review*, 92, 289-316.
- Murphy, G. L. (1988) Comprehending Complex Concepts. *Cognitive Science*, 12, 529-562.
- Myers, T. F., Franks, B. and Braisby, N. (1989) Partiality and Coherence in Concept Combination. EUCCS No. 33, Centre for Cognitive Science, University of Edinburgh, Edinburgh, March, 1989.

- Neisser, U. (1987) From Direct Perception to Conceptual Structure. Chapter 2 in Neisser, U. (ed.) *Concepts and Conceptual Development*, 1, pp11-24. Cambridge : Cambridge University Press.
- Nunberg, G. D. (1978) *The Pragmatics of Reference*. Bloomington, Indiana: Indiana University Linguistics Club.
- Oden, G. C. (1977) Interpretation of Fuzzy Logical Information. *Journal of Experimental Psychology*, 3, 565-575.
- Osherson, D. N. and Smith, E. E. (1981) On the Adequacy of Prototype Theory as a Theory of Concepts. *Cognition*, 9, 35-58.
- Osherson, D. N. and Smith, E. E. (1982) Gradedness and Conceptual Combination. *Cognition*, 12, 291-318.
- Partee, B. (1981) Montague Grammar, Mental Representations and Reality. Chapter 8 in French, P., Uehling, T. and Wettstein, H. (eds.) *Contemporary Perspectives in the Philosophy of Language*, 1. Cambridge : Cambridge University Press.
- Peacocke, C. (1979) *Holistic Interpretation: Action, Space, Interpretation*, 1. Oxford: Clarendon.
- Peacocke, C. (1986) *Thoughts: an Essay on Content*, 1. Oxford: Basil Blackwell.
- Perry, J. (1986) Thought Without Representation. *Proceedings of the Aristotelian Society Supplementary Volume*, 60, 137-151.
- Phillips, C. (1978) Constructivism and Epistemology. *Journal of Philosophy*, 53.
- Platts, M. (1979) *Ways of Meaning*, 1. London: Routledge & Kegan Paul.
- Pollock, J. L. (1986) *Contemporary Theories of Knowledge*, 1. London: Hutchinson.

- Pollock, J. L. (1987) Defeasible Reasoning. *Cognitive Science*, **11**, 481-518.
- Potter, M. C. and Faulconer, B. A. (1979) Understanding Noun Phrases. *Journal of Verbal Learning and Verbal Behaviour*, **18**, 509-521.
- Putnam, H. (1975) The Meaning of 'Meaning'. Chapter 4 in Gunderson, K. (ed.) *Language, Mind and Knowledge*, 1. Minneapolis: University of Minnesota Press.
- Putnam, H. (1978) *Meaning and the Moral Sciences*, 1. Cambridge: Cambridge University Press. realism and antirealism.
- Quine, W. V. O. (1951) *Mathematical Logic*. Cambridge, Mass.: Harvard University Press.
- Quine, W. V. O. (1960) *Word and Object*. Cambridge, Mass.: MIT Press.
- Quinton, A. M. (1973) *The Nature of Things*, 1.
- Quirk, R., Greenbaum, S., Leech, G. and Svartik, J. (1972) *A Grammar of Contemporary English*, 1. London: Longman.
- Rasmussen, S. A. and Ravnkilde, J. (1982) Realism and Logic. *Synthese*, **52**, 379-437.
- Reiter, R. (1980) A Logic for Default Reasoning. *Artificial Intelligence*, **13**, 81-132.
- Rips, L. J. and Turnbull, W. (1980) How Big is Big?. *Cognition*, **8**, 145-174.
- Rosch, E. (1973) Natural Categories. *Cognitive Psychology*, **4**, 328-350.
- Rosch, E. (1975) Cognitive Representations of Semantic Categories. *Journal of Experimental Psychology: General*, **104**, 192-233.
- Rosch, E. (1978) Principles of Categorisation. Chapter 4 in Rosch, E. and Lloyd, B. (eds.) *Cognition and Categorisation*, 1. Hillsdale, New Jersey: Erlbaum.

- Roth, E. M. and Shoben, E. J. (1983) The Effect of Context on the Structure of Categories. *Cognitive Psychology*, 15, 346-378.
- Salmon, N. U. (1982) *Reference and Essence*. Oxford: Basil Blackwell.
- Seligman, J. (1989) Perspectives in Situation Theory. Technical Report No. EUCCS/Working Papers Vol.3, Centre for Cognitive Science, University of Edinburgh, Edinburgh, 1989.
- Siegel, M. E. A. (1979) Measure Adjectives in Montague Grammar. Chapter 6 in Davis, S. and Mithun, M. (eds.) *Linguistics, Philosophy and Montague Grammar*, 1, pp223-262. London: University of Texas Press.
- Skorstadt, J., Falkenheimer, B. and Gentner, D. (1987) Analogical processing: a simulation and empirical corroboration. In *Meeting of American Association for AI*, Seattle, August, 1987.
- Smith, E. E. and Medin, D. L. (1981) *Categories and Concepts*, 1. Cambridge, Mass.: Cambridge University Press.
- Smith, E. E. and Osherson, D. N. (1984) Concept Combination with Prototype Concepts. *Cognitive Science*, 8, 337-361.
- Sperber, D. (1985) *On Anthropological Knowledge*, 1. Cambridge: Cambridge University Press.
- Strawson, P. F. (1985) *Scepticism and Naturalism: Some Varieties*, 1. London: Methuen.
- Swinney, D. A. (1979) Lexical Access During Sentence Comprehension. *Journal of Verbal Learning and Verbal Behaviour*, 6, 645-659.
- Tennant, N. (1987) *Anti-Realism and Logic: Truth as Eternal*, 1. Oxford: Oxford University Press.

- Thagard, P. and Holyoak, K. (1988) Analogical problem-solving: a constraint-satisfaction approach. In *Conference of the American Cognitive Science Society*, McGill University, Montreal, August, 1988.
- Warren, B. (1984) *Classifying Adjectives*, 1. Gothenburg University Press.
- Weir, A. (1985) Against Holism. *Philosophical Quarterly*, 35, 225-244.
- Wittgenstein, L. (1953) *Philosophical Investigations*. Oxford: Basil Blackwell.
- Woods, W. A. (1975) What's in a Link: Foundations for Semantic Networks. In Bobrow, D. and Collins, A. M. (eds.) *Representation and Understanding*, 1. New York: Academic Press.
- Wright, C. (1980) Realism, Truth-Value Links, Other Minds and the Past. *Ratio*, 22, 112-132.
- Wright, C. (1987) *Realism, Meaning and Truth*, 1. Oxford: Basil Blackwell.
- Zadeh, L. (1965) Fuzzy Sets. *Information and Control*, 8, 338-353.
- Zadeh, L. (1982) A Note on Prototype Theory and Fuzzy Sets. *Cognition*, 12, 291-197.

Appendix: Publications

1. Franks, Myers & McGlashan (1988)
2. Myers, Franks & Braisby (1989)

**DEFEASIBILITY IN CONCEPT COMBINATION:
A CRITERIAL APPROACH**

**BRADLEY FRANKS TERRY MYERS SCOTT McGLASHAN
CENTRE FOR COGNITIVE SCIENCE
UNIVERSITY OF EDINBURGH**

CONCEPT COMBINATION AND DEFEASIBILITY

What are the semantic relations within and between lexical concepts that constrain their combination? The position advanced is that they are criterial relations. Such a view has advantages over extensional and standard intensional approaches, which fail to capture adequately the characteristics of defeasibility at play during such combinations. We will examine these questions with respect to the adjectival modification of nouns.

Approaches To Concept Combination

Treatments of concept combination divide into "extensional" and "intensional" theories. The major failings of both approaches can be traced to some shared assumptions. Extensional approaches include set-theoretic treatments of concept combination: concepts are treated as unanalyzed units, and the detailed interactions of the contents of the concepts are bypassed. "Intensional" views construe concepts according to their internal properties, and concept combination is reduced to the interactive mechanism for the inheritance of properties from the inputs. Most writers postulate some partitioning of the features into, for example, "necessary" and "characteristic" types: in combinations, the "necessary"/deductive attributes of both concepts are to be inherited by the conjunction, whilst the "characteristic" attributes might be negotiated.

The shared assumptions of the two views are as follows. Firstly, the crucial attributes are those which are deductively related to the concept: as Jackendoff (1983) has pointed out, a defining-features intensional account is formally equivalent to a meaning-postulate extensional account. Secondly, and linked to this: the semantic object is a "total object" (Landman, 1986). Total objects are crucial to classical logic and realist semantics. An object is equated with the set containing *all* of its properties: these properties are the deductively-related identity conditions. Total objects are thus predicated upon the assumption of total evidence or information, even though this may be beyond the grasp of a particular language-user. Thirdly, neither view can distinguish between the conditions of application of

a concept, and the content of that concept. Since they are (total) object-centred, the assumption is that the concept can be applied only if all of its identity conditions are fulfilled; and these are usually taken as conditions on being a particular kind of object.

Standard intensional approaches have a further quality which is founded upon the above three assumptions. The characteristic properties are construed as default properties. And an object-centred approach demands that where the characteristic features of a concept obtain, so must the necessary features.

The major difference between the extensional and intensional approaches is that the former takes the necessary attributes and operates over them as a set, in terms of formal operations only. The intensional approach operates upon the attributes themselves, allowing the further use of non-necessary attributes.

The Criterial Relation

The criterial relation (henceforth, "C-relation") is a semantic relation which may hold either between two different concepts, or between a single concept and its evidential conditions. The notion derives fairly directly from the later Wittgenstein (1953). Traditionally, criteria are said to have the following properties. Firstly, the relationship between the evidence (qua criteria) and the particular claim (qua application of a concept) is somewhere between deduction (since it confers certainty), and induction (it is defeasible). Secondly, as the criteria are necessarily good evidence for a claim, they fix the semantic content of that claim. Thirdly, to have satisfaction of the criteria for a claim is consistent with obtaining further information which overturns that claim. The relation is thus defeasible. Fourthly, criteria are generally held to be multiple. In addition to criteria a concept has symptoms, or S-relations. These are inductively related features, similar to default properties in that their defeat will not alter the identity of the particular claim. So both semantic relations are inherently defeasible.

This reflects the received criterial view. The approach adopted here extends the concepts according to an anti-realist semantics. Primarily, C-relations govern the conditions of application or use of a term, rather than conditions on identity of an object (cf. Tennant, 1987). The criteria for a concept are C-related to the mental representation of the object, which is itself S-related and C-related to its attributes, and to other concepts. C-relations are the conventionally underpinned semantic relations holding between the evidence for, and the application and evaluation of, a concept. Hence, C-relations embody partial objects based upon partial information: they concern how we may reliably "go beyond the given" evidence to infer a particular kind of conceptual entity. C-relations license defeasible

extensions of partial objects to less partial ones.

This allows us to specify the major difference between S-relations and default properties; since the properties are not tied to the identity conditions of objects, where the S-relations of a concept obtain, there is no necessity that the C-relations do so. Hence, a concept may be applied on the basis of S-relations alone, but with less warrant than a C-related application.

The final difference between the standard views and the one adopted here is that the central semantic relation is content-driven and non-transitive. This issues in a type of defeasibility which is not open to the other views.

Types of Defeasibility

Three kinds of defeasibility are pertinent: rebuttal, undercutting and default. The first two involve inter-conceptual conflict and defeat, the last one intra-conceptual. But firstly, we define a

Defeater: if P is evidence for Q, R is a defeater for this evidence iff:

- R is consistent with P; and
- (P & R) is not evidence for Q.

Crucially, anti-realism's denial of excluded middle denies that this is equivalent to evidence for (-Q).

Rebuttal: Denial of the Claim:

Rebutting Defeater: If P is prima facie good evidence for Q, R is a rebutting defeater for this iff:

- R is a defeater; and
- R is good evidence for (-Q).

Now, there are two types of rebutting defeaters:

Type I: R is a type I rebutting defeater iff:

- R is a rebutting defeater; and
- R is good evidence to support the claim that P would not be warrantably assertible unless Q were so. Hence, (-P).

(The second clause is, of course, a generalization of the definition of modus tollens in classical logic: -P unless Q).

Type II: R is a type II rebutting defeater iff:

- R is a rebutting defeater; and
- R is good evidence to deny the claim that P would not be warrantably assertible unless Q were so. Hence, P may be assertible.

(Here, the second clause generalizes the denial of modus tollens).

Any defeasible deduction will necessarily be a type I rebutting defeater, since modus ponens supports modus tollens. Hence, set-theoretic and standard intensional views can utilize only this type. How would this operate? A defeasible concept is applicable iff we can show that *none* of the defeating conditions obtains. So certainty is equated with necessity, and necessity is read as a deductive relation. Any weaker relation is inherently doubtful. The logical possibility of doubt in any situation is equipollent with an actual grounded doubt in the current situation.

Undercutting: Denial of the C-relation:

Defeat of C-relations occurs by an:

Undercutting Defeater: If P is prima facie good evidence for Q, R is an undercutting defeater for this iff:

R is a defeater (not a rebutting defeater); and

R is a good evidence to deny the claim that P would not be warrantably assertible unless Q were so.

(Note, in this case, the second clause cannot be a generalization of modus tollens, since we do not have an outright denial of Q).

This type of defeater, then, attacks the *connection* between P and Q rather than Q itself. The defeater is a reason for denying that we would not have the evidence unless the conclusion were true. This does not imply either $\neg Q$ or $\neg P$: it denies that a claim or generalization (Q) on the basis of P would be warrantably assertible. But the evidence itself is still assertible, and indeed the claim could still be so.

The epistemological assumptions upon which C-relations are predicated are a great deal more naturalized than their classical realist counterparts (Baker & Hacker, 1984). The obtaining of a C-relation provides certainty in the sense that it is conclusive evidence for the claim. And it is conclusive in the sense that the evidence cannot be improved upon, even if it is multiplied: and it can nonetheless be overturned. The burden of proof is also altered: on the classical view, there is a *requirement* to check and deny the (possibly open-ended) list of potential defeaters. On an anti-realist view, there is no such requirement: if the prima facie evidence supports the claim, then, if there are no available prima facie defeaters, the claim will go through - even though there is a possibility of defeat. So the logical possibility of doubt is not equivalent to the existence of an actual grounded doubt.

Default: Denial of Inductive Properties:

This is the approach exemplified in the S-relation and utilized in frame and prototype theory. A default property is either a parameter or a value of a parameter which is linked to a concept by an inductive

relation: it is a "typical" property. Where the property is overridden, the application of the concept itself is nonetheless still justified.

ADJECTIVE-NOUN COMBINATION

Noun Phrase Constructions

One type of construction which illuminates the different kinds of defeasible semantic relations are NPs where an A modifies the N, as in sentence-type 1, below; the crucial questions concern how this modification operates, especially in cases of conflict of A and N properties, and whether we can derive sentence-types 2 and 3 from the type-1. The sentence-types can be rendered:

1: This is a (A)(N) 2: This is A 3: This is a N

It will be evident that the derivation of type-3 sentences is equivalent to the derivation of the superordinate category of the N from that of the NP, with the preservation of the sortal type. For *privative* adjective-types, the inferences to type-2 and type-3 sentences are problematic: the A functions neither predicatively nor attributively. Consider the following case:

1: This is a fake Hogbin 2: This is fake 3: This is a Hogbin

Here, we cannot infer 3 from 1, although 2 is sensical. In terms of classical logic's set-theoretic treatment, privatives are modelled by a condition which requires that, if X is a member of the set of entities denoted by N, it is a member of the relative complement of the NP set, in the particular domain. (Hoepelman, 1983).

Now, the kinds of examples with which we are concerned are ones which we term "functional privatives": ones where the inference to type-3 sentences is made problematic by the *interaction* of the semantic properties of the N and A. Thus functioning privatively is *not* an intrinsic property of the type of adjective being used - rather, it stems from the conflict of semantic features and the principles of defeasibility by which they are resolved. Examples include:

This is a stone lion

This is a plastic flower

These adjectives can function affirmatively in other contexts:

This is a stone bridge

This is a plastic chair

They each allow inferences to their appropriate sentence types 2 and 3. In the functionally privative cases, the inference to type 3 is problematic, although not completely nonsensical.

Problems With The Deductive Approach

If we map a deductive relation over total objects, the only form of defeasibility open is type I rebuttal. So the denial of the evidence for the necessary conditions of a concept entails a denial of all of its conditions - i.e., of the concept itself. This is precisely the effect achieved by the set-theoretic relative complement method. Three problems emerge: firstly, it denies too much: with functional privatives, we would like to preserve at least some of the non-essential properties of the "lion" concept; since they support *modus tollens*, and require default properties to be dependent upon the obtaining of necessary conditions, standard views cannot allow this. Secondly, there is no account of the asymmetrical nature of the combination. We can argue that this will largely be at the behest of dependency relations, in which the C-relations of the A have primacy (cf. Anderson, 1986). Thirdly, the method can only tell us what the object *is not* rather than what it is: a "fake Hogbin" just is not a Hogbin, and a "stone lion" is not a lion: yet it seems that a "fake Hogbin" is somewhat more like a real Hogbin in appearance than is a Rothko. Again, a stone lion is more similar in shape to a lion than to a frog. The argument of the next section will be that the C-relation view can make good most of these flaws.

THE CRITERIAL APPROACH

We claim there are two different defeats at work here: one type II rebuttal, the other an undercutting. Ordinarily, a C-information based concept of "lion" will function as evidence for the sortal identity of an object: the conceptual evidence (P) is C-related to the identity claim (Q). But the combination of the concept "stone" (R) with P will undercut that C-relation. The combination of P with R involves type II rebuttal of the C-information lodged in P by the C-information in R. So, concept Q's use is C-related to certain kinds of information which is generated in support of the use: for example, ideas of "internal essence" of the identity of lions, derived from conventionally structured lay theories of the domain (see Murphy & Medin, 1985; Keil, 1987). If these properties were necessary/deductive conditions on a total semantic object, the overriding of them would function according to rebuttal type I; hence the term could not be applied. That this is not the case is clear from the ease with which we can understand "stone lion". The concept's use is also S-related to certain features, concerning "appearances". So we can justifiably utilize the term "lion" if we can generate appropriate C- and S-relations (i.e., P) on the basis of current context and information. But we have an undercutting of the connection between P and Q, by R: that is, the applicability of the term "lion" on the basis of P is denied. This denial of positive support for Q is not equivalent to an assertion of Q's falsity. And this also leaves P open to negotiation.

For undercutting, the combination (P & R) must be consistent; this is where the type II rebuttal of the noun's C-information, which motivates the undercutting, operates. Crucially, this type of defeat could

only be facilitated by a context of an undercutting, and not by a rebuttal of claim Q. The question is then, which has priority, P (the noun), or R (the adjective)? On the basis of linguistic dependency, the C-relations of "stone" (such as "inanimate", etc.) have the effect of overriding the C-relations of "lion" by a type II rebuttal. The full concept of "stone" is thereby rendered consistent with the undefeated S-relations of "lion". Two points are noteworthy here. First, the survival of some of the S-information is acceptable since S-relations do not require the obtaining of C-relations. Second, in this combination not all of the S-information survives: only that which is *required* by the C-relations of the "stone" concept. The other S-relations would either be directly negated (by default defeat), or rendered inconsistent (by general defeat). But the C-relations of "stone" positively require that there is some physical structure or shape: and it is these qualities of "lion" which are retained. We claim that the defeat of the C-information of "lion" by that of "stone" is type II rebuttal because it is a cancellation of the criterial *properties* of the noun; and the other defeat is by undercutting because it defeats a C-related *claim* made on the basis of such properties.

We have advocated a well-motivated intensional account of the defeasible semantic relations constraining concept combination. Such an account seems to be necessitated by the case of functional privatives. Its extrapolation to more straightforward concept combinations should be perspicuous.

REFERENCES

- Anderson, J. (1986), Suprasegmental dependencies, in J. Durand (ed.), *Dependency and Non-Linear Phonology*. London: Croom Helm.
- Baker, G.P., & Hacker, P.M.S. (1984), *Scepticism, Rules and Language*. Oxford: Blackwell.
- Hoepelman, J. (1983), Adjectives and nouns: a new calculus, in R. Bauerle, C. Schwarze & A. von Stechow (eds.), *Meaning, Use and Interpretation of Language*, Berlin: de Gruyter.
- Jackendoff, R. (1983), *Semantics and Cognition*. Cambridge, Mass.: MIT Press.
- Keil, F.C. (1987), Conceptual development and category structure, in U. Neisser (ed.), *Concepts and Conceptual Development*. Cambridge: CUP.
- Landman, F. (1986), *Towards a Theory of Information*. Groningen-Amsterdam Studies in Semantics, 6, Dordrecht: Foris Publications.
- Murphy, G.L., & Medin, D.L. (1985), The Role of theories in conceptual coherence, *Psychological Review*, 92 (3), 289-316.
- Tennant, N. (1987), *Anti-Realism and Logic*. Oxford: Clarendon Press.
- Wittgenstein, L. (1953), *Philosophical Investigations*. Oxford: Blackwell.

**Partiality and Coherence
in Concept Combination**

**Terry Myers
Bradley Franks
Nick Braisby**

Research Paper EUCCS/RP-33

March 1989

**Centre for Cognitive Science
University of Edinburgh
2, Buccleuch Place
Edinburgh EH8 9LW, Scotland**

**Copyright © 1989 by
Centre for Cognitive Science
University of Edinburgh**

Partiality and Coherence in Concept Combination

Terry Myers

Bradley Franks

Nick Braisby

Centre for Cognitive Science
University of Edinburgh

August 31, 1989

1 Introduction

This paper addresses the problems of lexical representation and coherence. Our discussion of these problems leads to what we call a sense generation approach. We will present two accounts from within this approach and illustrate them with respect to concept combination. In section 2 of this paper we describe the tension that exists between the two phenomena of semantic flexibility and specificity. There are accounts of lexical representation in which this tension raises major problems and there are others in which it provides a clue to resolving these problems. The former are characterised by what Clark (1983) calls the selection of senses and the latter by what he calls sense creation. In section 3 we discuss the implications of sense selection accounts for coherence and concept combination, indicating problems that arise. We rely heavily on Murphy & Medin's (1985) arguments and suggest some extensions. In section 4 we present the sense generation view (which is related to Clark's notion of sense creation) which avoids such difficulties. Central to this account is a consideration of partiality, reduction of partiality being brought about by constraints provided by the discourse and situational factors. In section 5 we will sketch two different approaches to concept combination within this framework. In section 6 we draw conclusions and raise questions regarding the nature of coherence and lexical representation.

2 Flexibility and Specificity

Consider the following example:

Rudy is at the ballet watching a single ballerina dancing on stage. Excitedly, he whispers to his friends, "Isn't she delightful?". His friend Ron agrees, replying, "Yes, she's very beautiful". His other friend, Reg, agrees too, saying, "Yes, what an exquisite pirouette". But Rudy replies, "I simply meant she's a wonderful person".

The order of authorship is arbitrary. The first two authors were supported by European Basic Research Action, ESPRIT project 3175. The second two authors were supported by ESRC studentships Nos. COO428622024 and COO428722001 respectively. We would like to express our deep gratitude to Robin Cooper, George Dunbar and other members of the Mental Lexicon Workshop for many useful discussions and insights. We would also like to thank Richard Cooper for his help in preparing the printed version of this paper.

Delightful in the above exchange seems to mean different things for the various individuals involved. From Ron's perspective, *delightful* means 'physically delightful'; from Reg's, *delightful* means 'dancing delightfully', while for Rudy *delightful* means 'delightful as a person'. From the perspective of an usher walking past who only hears Rudy's first comment, the meaning of *delightful* is indeterminate with respect to these specific meanings.

Suppose that we want to characterise the semantic component of the lexical entry for *delightful*. We might propose that the lexical entry is sufficiently general to encompass all the observed meanings. In which case, we require a lexical entry which is general enough to capture all the different meanings that can be conveyed by *delightful*, including those illustrated in the exchange. Even if we could specify such a general lexical entry for *delightful*, however, this would not then allow the possibility for *delightful* to convey, say, the meaning 'delightful as a person'. Yet, in use, *delightful* can convey just this. What is more, any of the meanings indicated above could be made even more specific if more contextually provided information were available. For example, Ron's meaning for *delightful*, 'physically delightful' could have been more specific: physically delightful with respect to coiffure' if we know that Ron is, say, a hairdresser. In principle, there is no bound to this degree of specificity and so a general lexical entry seems even more implausible. Tied to this idea of unboundedness, we also want to allow that meanings can be indeterminate: for example, the meaning for the usher, is non-specific with respect to the meanings for Rudy and his friends. In conclusion then, our treatment of the lexical entry as general, being motivated by the observation of flexibility, renders the lexical entry unable to provide for the specificity of meaning that we observe. Seemingly, we can't treat the phenomena of flexibility and specificity independently of one another: to do so produces the tension we have illustrated¹.

3 Sense selection

3.1 Weak and strong sense selection

For the purposes of a distinction to be made later, we will refer to the semantic contribution of the lexical entry as the "lexical concept", and the intended meaning as the "generated sense". One approach to characterising lexical entries is to treat the different meanings of a word as different lexical concepts. For example, under this view we might suppose that the lexical entry for *mother* includes several lexical concepts for *mother*, which we might label "foster mother", "step mother", and "biological mother" - in the same way as, say, *bank* is assumed to have a number of independent lexical concepts underlying its use. A strong version of this view would hold that for each and every different sense a different lexical concept is accessed from the lexical entry. A weaker version maintains that there may be independent, unrelated lexical concepts from which, on different occasions of use, different senses may be elaborated. For example, under the weak view, we could, in the context of a fairy tale, arrive at a sense for *mother* corresponding to "step-mother who is wicked" by elaborating the lexical concept for *mother* labelled "step-mother". Presumably, on the strong sense selection view we would select the appropriate sense, "step-mother who is wicked", relevant to the context and that could not be further elaborated.

¹A similar point has been made by, among others, George Dunbar. In Dunbar (1988) he suggests that, "while lexical meaning is flexible, lexical use can be very precise ... this is paradoxical" (p. 69).

Both versions of sense selection miss the crucial intuition that these senses are closely related; they are not at all like the different senses of *bank*. That is, *river bank* and *money bank* have clearly distinct meanings, and thus *bank* should have (at least) two different lexical concepts. In the case of *mother* the different lexical concepts we could hypothesise are clearly related. Accordingly, in the strong version, we have a granular picture of flexibility and specificity. Flexibility is granular in that the possible senses a word may have cannot alter and we have an impoverished view of flexibility. Specificity is also fixed by the set of senses in that senses can be no less or no more specific than those represented. The weak version allows senses to vary in terms of their specificity but we only have apparent flexibility: the range of possible senses is restricted by the initial choice of lexical concept.

3.2 Coherence in sense selection

In exploring the implications for coherence of the sense selection view we will treat Lakoff's (1987) analysis of *mother* within this account, although we accept that he may not actually be committed to such a view. That is, we might see each of his sub-models as a different lexical concept within the lexical entry for *mother*. A number of consequences then follow. In particular, we will argue that we cannot maintain both weak sense selection and a theory-based view of coherence. To see this we must turn to Murphy & Medin.

In Murphy & Medin's (1985) account we may distinguish between internal and external components of conceptual coherence. Both aspects derive from theories of the world. Internal coherence results from the relations between features for which the theory provides an explanation. For example, the various properties of chairs represented in the concept (its size, having a seat, its rigidity and so on) cohere in virtue of theories concerning the function of a chair. External relations between different concepts are also motivated by theories that support causal and explanatory connections between those concepts. Consider the concepts "cat" and "dog". According to Murphy & Medin these concepts are coherent to the extent that they are embedded in our theories of the world: the greater the number of connections between such concepts and the rest of our knowledge, the more coherent and stable they are. The connection of "cat" to many different aspects of our knowledge (such as our theories of domestic life, ownership, pleasure and so on) provides it with coherence over and above that which is provided by our theories of "cat" alone, which is the source of the concept's internal coherence. The same thing is true of complex concepts, for example "cat and dog". The concept is coherent as a result of there being many common theories of cats and dogs. Seemingly, on Murphy & Medin's view, we could not allow that complex concepts such as "elephants and lemons" or "stone lions and trout" have the same level of coherence as "cat and dog". This stems from the fact that such complex concepts has fewer common theories to relate the component concepts.

This view of the theories underlying coherence seems to be intimately connected to metaphysical issues. Accordingly, for Murphy and Medin, concepts result from the way that theories cut up the world. Consequently, categories that cut across ontological boundaries would not be very coherent. For example, "elephants and hopes" would not be considered coherent, unless this category were motivated by a theory. If this theory were consistent with one's other theories, the coherence of the concept would be enhanced. The claim of Murphy & Medin is that the flexibility of conceptual coherence derives from the flexibility

of particular theories. On our interpretation of their claim, this means that there are two mechanisms for allowing coherence to be flexible. Either, in certain contexts, coherence is enhanced by constructing an explanation or theory or, alternatively, the theories that underpin coherence are inherently flexible. It is not clear how these mechanisms for flexibility in coherence would work for examples like the following.

Imagine you are sitting on a park bench in London. In front of you is a stone statue of a lion. Consider how you would reply, if someone were to ask you, "Is that a lion?" Seemingly, there is one sense in which you might say "yes", and a related one in which you might say "no": and, given these senses, either reply could be appropriate, according to circumstances.

It seems, then, that the word *lion* can be used quite appropriately to talk of a statue of a lion. However, there is a question as to just how many of our theories of lions involve statues. Seemingly, our theories of lions are capable of cutting the world up so that we can categorically assert that a statue of a lion is not a lion. One approach to this phenomenon is to view it as a peculiarity of word use. However, we wish to preserve the intuition that there is an essential perspectival aspect to this case. That is, we can adopt a different perspective on a statue of a lion: we may simply view it as a lion. Additionally, in this case each of these perspectives seems equally coherent. As we see it, there are two possibilities for Murphy & Medin to account for this apparent flexibility in coherence. One is to invoke their suggestion regarding the flexible nature of theories. Another is to adopt the standpoint of weak sense selection, and thus hold that there are two lexical concepts for *lion*. In what follows we will explore the former possibility and conclude that the latter is their only plausible option. This in turn leads to an inconsistency.

Under the first option, that of invoking flexibility in theories, we envisage three possible mechanisms open to Murphy & Medin for achieving the flexibility of coherence. One mechanism that Murphy & Medin offer is that of the construction of an explanation in cases such as "elephants and hopes" where, seemingly, ontological boundaries are crossed. In the above case, apparently there is a similar crossing of ontological boundaries. Our theories of lions tell us that stone lions are not really lions, though presumably there are some theories of lions that are also theories of stone lions. Perhaps then it is that the coherence of the concept of lion in this case relies on a constructed explanation. However, since our stable, underlying theories concerning lions cannot all apply in the case of the stone lion, it must be that the use of lion in talking of a stone lion is less coherent than the use of lion in talking of a real lion. This follows because coherence in Murphy & Medin's view falls out of the number of explanatory links which structure the concept. Hence, on this view such constructed explanations necessarily lead to less coherent concepts than those stable explanations provided by the knowledge base. This is at odds with our intuitions. Seemingly both uses are, in this case, equally coherent. What we want is to allow that both ad hoc and stable explanations can provide for equal coherence. This is what we will argue for in section 4.2.

Murphy & Medin provide another mechanism to account for the flexibility in coherence that we observe. This is the possibility that we may have flexible theories. In the case of the stone lion, we might suppose that our theories concerning lions are flexible, accounting for the observed flexibility of coherence in use. The use of *lion* to talk of a stone lion can be coherent as we would want in virtue of the fact that the same theories concerning lions can

also be theories concerning stone lions. However, this does not accord with the intimate connection that theories have to metaphysical concerns. If our theories concerning lions are also theories concerning stone lions then we have no principled manner of cutting the world such as to differentiate between non-lions (like our stone lion) and real lions. We take it that this undermines the metaphysical position that Murphy & Medin adopt.

A third possibility that we believe Murphy & Medin may allow for is the option of theory change. This option entails that in order to preserve the coherence of the lion concept when employed to talk of a stone lion, the theories underlying the concept would have to change. This, however, seems problematic. Since they adopt the Quinian position of viewing theories as components of a large and intricate web, a change in one will have ramifications for the others. In this case we would have to change our theories concerning lions so as to allow the possibility of inanimate lions. Changes as radical as this, as they acknowledge, would require a global reorganisation of one's knowledge base. It would also require, under most interpretations, a divorcing of theories from metaphysics. This is precisely what we argue for. In section 5 we will present two views that are not tied to the assumptions that lead to these counter-intuitive consequences. As we have seen there are good reasons to suppose the kind of flexibility observed in coherence is not to be captured by the notion of flexibility in theories. We now turn to the second possibility we pointed to earlier.

The alternative position that we believe Murphy & Medin might adopt is the one that we call weak sense selection. That is, the above arguments will not go through if we assume that there are two lexical entries and hence two lexical concepts for *lion*. Aside from the arguments already presented against this option, it seems there are further considerations that would prohibit Murphy & Medin from taking this view. Following Quine (1960), we might like to distinguish between two lexical entries for *light*. This move seems legitimate. After all, it seems that the theories we have for light-weight objects do not have to be the same as those for light-coloured objects. For the case of *lion*, although lions and stone lions are differentiated by the theories provided by the knowledge base, lions and stone lions are nonetheless very closely related. It is presumably the case then that since we can relate lions and stone lions we have theories to do just this.

Returning to our interpretation of Lakoff, we can see that similar arguments apply. That is, if we tie Murphy & Medin's arguments to this assumed position then we may allow that the different sub-models for *mother* are related by what they call theories. Consider the following example:

Kim lives with her biological mother, Mary. Mary adopts a child, Jane, to whom Mary behaves in all respects as she does to Kim. Jane calls Mary "mother" and behaves towards her in the same way Kim does. Initially, Kim cannot accept that Mary is Jane's mother but after some time she does.

How is it that Kim's notion of mother appears to change? One possibility is that initially she only has the "biological mother" concept of mother and that she somehow acquires another concept of mother, the "adoptive mother" concept. Even though Kim starts off with the notion of mothers as biological mothers, one can envisage that Kim's theories of the world allow her to construct an explanation of how Jane can call Mary "mother". Indeed, she may come up with a new notion of mother, a mother who is like a biological

mother in all respects save the biological relations. That is, Kim, with her notion of a biological mother, equipped only with her theories of the world, can construct the notion of an adoptive mother. The fact that this seems so clearly to be the case undermines the starting point of weak sense selection. Namely, that there may be independent lexical concepts for *mother*. The fact that they may be related by our theories of the world indicates that they are not independent.

We have seen that to posit independent lexical entries is to undermine those very arguments that Murphy & Medin advance in favour of theories. Yet, we take those arguments as very good grounds to accept the theory-based account of concepts. Consequently, we also take it that what is needed to avoid the pitfalls of the sense selection accounts is an extension of this theory-based view. So, although we crucially need a way of describing the difference between lions and stone lions and mothers and adoptive mothers, we need a way of avoiding making an unprincipled distinction at the level of lexical entries.

There are several possible responses to this difficulty. One is to place constraints on the way theories may relate these independent lexical concepts. That is, we may want to place constraints on the possible extensions of any two of the lexical concepts for *mother*, such that one cannot be extended to form the other. However, this flies in the face of the observation that we can do just this. Kim *can* construct a notion of adoptive mother from her notion of biological mothers. So for this to be a viable option, more justification is needed. Another option is simply to reject the arguments of Murphy & Medin that the coherence of senses is determined by theories. The fact that biological and adoptive mothers may form a coherent complex concept has nothing to do with the theories that relate them. As stated earlier we agree with Murphy & Medin's general position. We believe that the only other alternative is to reject the assumptions on which weak sense selection is based. As a consequence, we can see that the interpretation we placed on Lakoff in order to discuss coherence must be rejected. Accordingly, we can only view Lakoff's sub-models as describing something other than lexical concepts.

4 Sense generation

As we have shown, we can characterise lexical entries according to the assumptions of sense selection. We believe that strong sense selection does not provide a viable characterisation. We have illustrated weak sense selection by choosing to interpret Lakoff in this manner. This position is challenged by the arguments in favour of theories as determinants of coherence. In this section we will present an alternative account of lexical representation that is consistent with such arguments.

4.1 Partiality

From our consideration of the arguments of Murphy & Medin concerning coherence, we require independent lexical concepts only in those cases where it is clear that they are unrelated by the kind of theories they discuss. For example, in the case of *bank*, we would require two lexical concepts, one for *river bank* and one for *money bank*. These, after all, seem unrelated unless we construct some ad hoc explanation. In the case of *mother*, though, we would not want to hypothesise independent lexical concepts for "foster mother", "step-mother", etc. Instead we would want to represent *mother* with a single lexical concept.

In order for this concept to apply in the cases of all types of mother, that is to capture the phenomenon of semantic flexibility, we would seemingly require it to be general with respect to these cases. As we indicated before, this appears to disallow the possibility of capturing the phenomenon of specificity. However, this conclusion is tied to the assumption that lexical concepts and senses are total objects. Under this assumption, the sense for *mother* cannot be extended or elaborated. That is, we can only have one sense for *mother*. However, that there *are* many different senses is undeniable.

An alternative is to view the various senses as extensions of a single lexical concept, a view that is related to that of Dunbar (1988). That is, senses are related to lexical concepts via a relation of concept extension². Returning to our example of Rudy at the ballet, the different meanings we may associate with *delightful* can be ordered in the following way. The usher's meaning is only as specific as the lexical concept, and this meaning is compatible with any of the possible extensions of the lexical concept. Ron's meaning, however, is more specific. It doesn't just indicate that there is something that is delightful, it also indicates what it is that is delightful, namely the dancer. Thus, for Ron, the information conveyed by *delightful* would appear to contain the information gleaned by the usher. Similarly, for Reg and Rudy with respect to the usher. Thus, the ordering is one of informational containment. However, if we compare the information conveyed by *delightful* from Ron's perspective with that conveyed from Reg's perspective, we cannot say that either is more informative. Thus the relation of informational containment is undefined for these two senses. It follows that the relation of informational containment is a partial one.

Our position then is that semantic flexibility can only be properly captured by considering the lexical concept as a partial object. This can be differently extended according to the interaction of various constraints on each occasion of use. We see such constraints as being provided by the discourse, the situation in which it resides and cognitive models which pertain to both the discourse and the situation. It will emerge that these are the very constraints that underpin the generation of a coherent sense. That being the case, we will see that coherence and flexibility are but two sides of the same coin.

4.2 Coherence in sense generation

We have so far considered the coherence of concepts from the point of view of Murphy & Medin's arguments. On their view coherence is intimately bound to metaphysical issues. As we have seen there are a number of problems associated with this approach. In what follows we will be concerned with the coherence not of concepts but of senses. In this way we avoid being committed to a necessary connection between coherence and metaphysics.

Taking lexical concepts as partial objects seems to provide us with too many degrees of freedom. So, for instance, we seem to require a constraint to prevent certain co-occurrences of features. Consider our lexical concept for *balloon*. We might want to ascribe to it the property of not being able to float in the air. Under the unconstrained account of the extension of partial objects, this is perfectly permissible. However, following on from the arguments of Murphy & Medin, we would want to claim that such an ascription could only

²From here on we use *concept extension* to mean that relation between a lexical concept and the same lexical concept some of whose features have been modified. Modification here include the cases of simple adding of features as well as the denial of features. In no sense do we mean *extension* to refer to the set of objects described by the lexical concept.

be valid if licensed by a theory. In this case, the balloon could be filled with a heavy gas. Hence, in order to constrain what extensions we allow, the process of extending a lexical concept to form a sense must be the very process by which that sense becomes coherent. That is, the process of extension must realise the kinds of theory envisaged by Murphy & Medin.

In order to illustrate the kinds of issues which emerge with respect to coherence in sense generation, we will consider the example of the stone lion, of section 3.2. According to our view, a given sense of a word derives from the corresponding lexical concept and may or may not be an extension of it. Consider describing a stone lion as a lion. We can account for this sense of lion in either of two ways. One possibility is that the lexical concept is general and no extension is required. From our preceding arguments this option would disallow the possibility of simultaneously satisfying the conditions of semantic flexibility and specificity. The only option then is to consider that the lexical concept is specific in the following sense: it contains, or may inherit, among others, feature descriptions such as "animate", "four-legged" and "golden-colour". This will capture the sort of specificity we require for some occasions of use, for example, where we are describing a lion at the zoo. There may, however, be circumstances in which a lexical concept would be extended further, perhaps in order to distinguish between lions and lionesses. More importantly though, as it stands, the lexical concept cannot meet the flexibility requirement. In order to resolve these problems, we will invoke a more general mechanism for concept extension. So, in the case of the stone lion being described as a lion, seemingly the feature "animate" present in the lexical concept is denied in the generated sense. Accordingly, what we require of the constraints which underpin concept extension is that they allow this sort of defeasibility. In section 5 the two views we present will adopt different methods of achieving this.

In principle, we allow that any property ascription provided by a lexical concept can be defeated. As a consequence, we cannot regard the various entities described by a given word as forming a category which has some metaphysical significance. That is, such categories, in contrast to those envisaged by Murphy & Medin, do not "carve the world at its joints". Returning to the example, since a stone lion is inanimate, we require that the extension of "lion" involves at least the defeat of the property ascription concerning animacy. In this way we can meet the requirement of semantic flexibility. Given the lexical concept we have for *lion*, there are circumstances in which it seemingly can't be used. However, via our defeasible mechanism for concept extension, we can override any feature which would prohibit its use. The important point here is that this overriding can only be licensed by theories, *ad hoc* or otherwise.

The mechanism as we have described it so far is a little simplistic. Consider again the case of the stone lion being talked of as a lion. In one sense, we may agree that the statue is a lion but in another we may protest that it is not. That is, we require some choice in the mechanism as spelled out. If we want to affirm that the statue is a lion then we will adopt the defeasible method of extension already outlined. If, however, we want to deny that the statue is a lion we will not extend the lexical concept in this defeasible way. Of course, we may allow the statue to be described as just that, a statue. Again, in such a case we wouldn't require any defeasible extension of the lexical concept for statue.

These different options for concept extension lead to different property ascriptions, that is,

different descriptions concerning the same entity. We take these descriptions of an entity to define different perspectives. For a particular cognitive agent, a perspective associates descriptions with given entities. So, if Fred approaches our stone lion adopting the "real lion" perspective, one that is defined by the description given by the lexical concept for lion, he won't categorise it as a lion. However, were he to adopt the "statue of a lion" perspective, one given by a description resulting from defeasible extension of the lexical concept for lion, he would categorise it as a lion. He could of course adopt a third perspective, the "stone statue" perspective, resulting from a non-defeasible extension of the lexical concept for statue. From this perspective, he would similarly categorise the stone lion as a statue.

We can now see that an adequate characterisation of coherence is not going to be provided solely by the theories which Murphy & Medin discuss. Instead, we will pursue the idea that theories may be local: that is, provided by ad hoc explanations, situational factors and informational requirements. If Fred were to exclaim, "This is not a lion", then we can only ascribe to him the perspective we labelled the "real lion" perspective. It is only this perspective that makes this use of lion coherent. If we attribute the "statue of a lion" perspective to Fred, then this particular use of lion would be rendered incoherent. It follows then that the coherence of a sense can only be defined relative to some perspective. So, clearly, perspectives may involve local or ad hoc theories: we don't carry around our theoretical luggage everywhere we go.

There are many constraints on concept extension that give rise to coherence. Some are to do with the extra-linguistic aspects of word use, such as the agents involved, the situations in which they find themselves and the particular informational requirements they have. Others are to do with the linguistic context, concept combination, for example. In what follows we will focus on the notion of informational requirement and examples of concept combination.

Consider the following puzzle taken from Braisby (1989):

Fred is sitting on a park bench in London. He knows that at the other end of the park there is a statue of a lion. A schoolgirl approaches him and, explaining that she has been given an assignment to sketch a lion, she asks Fred if he has seen one. Fred replies that he has and points her towards the statue. A little later an exhausted zoo-keeper appears and, explaining to Fred that a lion has escaped from the zoo, he also asks Fred if he has seen one. Fred replies that he hasn't.

Related to our discussion of perspectives, we would like to say that Fred is adopting different perspectives for different occasions. The reasons for choosing the particular perspective he does are to do with the nature of the information required of him. Fred can infer that the schoolgirl requires information germane to lions that one can draw and so adopts the relevant perspective in order to reply in the way that he does. The zoo-keeper, however, places a different informational requirement on Fred. It is this that similarly helps determine the perspective Fred adopts and so allows him to reply as he does to the zoo-keeper.

We would like to distinguish two aspects to informational requirement. One is appropriate to the degree of specificity a given sense may have. So, in the above example, in replying to the zoo-keeper, Fred means "real lion" by *lion*, a meaning given by the lexical concept

for lion. In replying to the schoolgirl, however, Fred has a more specific meaning in mind, namely *lion* as meaning "statue of a lion", a meaning that can be derived via defeasible extension from the lexical concept for lion. That is, the informational requirements of the agents involved help determine the degree of specificity of the senses of Fred's use of *lion*. In addition, these requirements lead to a difference in the flexibility of these meanings. That is, the meaning of Fred's use of *lion* in replying to the schoolgirl, is only flexible enough to describe statues of lions. The meaning of his use of "lion" in replying to the zoo-keeper, however, is different in terms of flexibility: it is only flexible to the degree that it can describe real lions. So the informational requirements of the agents involved can also lead to differences in the flexibility of meaning.

The other constraint on concept extension that we will consider is that provided by concept combination. In the next section we will expound on this within the general framework developed so far. We will do so via the application of two alternative views.

5 Partiality, coherence and concept combination

Our preceding arguments lead to a set of conclusions that we take as requirements on any theory of lexical representation. First and foremost is the commitment to sense generation, in which non-ambiguous words are represented by a single lexical entry, which contributes one lexical concept, and particular senses are generated from this by concept extension. Further, we take the lexical concept to contain a description of the central properties of entities which can be described by the corresponding word. Accordingly, the lexical concept is not general in the Quinian sense; that is, in itself it can only describe a subset of the range of entities which the word itself can describe. Quine's notion of generality would maintain that the lexical concept can describe all individuals the word is used to describe. Entities outside of this subset, but which nonetheless can be described by the word, require a different mechanism. The mechanism we propose is one of concept extension. We see this as being determined by theories, though not ones that carry metaphysical assumptions, *a la* Murphy & Medin. As a consequence, the extensions of concepts are senses, whose coherence is provided by the nature of the extension. The kinds of theories which underpin the generation of senses are those that allow for defeasibility. Similar to the "lion" case we discussed earlier, we will treat such extensions as defining perspectives, and accordingly coherence obtains relative to a perspective. Perspectives themselves are constrained by factors of the situation, local context and the informational requirements of the agents involved. The following two views are based upon these conclusions.

5.1 View 1

This section outlines a possibility for lexical representation developed by Braisby (1989). Braisby refers both to an implementation in C-prolog and also a situation-theoretic interpretation of this implementation. For purposes of brevity, we will not go into these matters here. What we will present is necessarily a simplification. However, the main points will still emerge. In that account a word meaning (or WORM) can be described in a similar way to that offered by the standard classical view. For example, the WORM for *lion* could be described as follows.

$$\left[\begin{array}{l} \text{animacy:animate} \\ \text{colour:golden} \\ \text{legs:4} \end{array} \right] \quad (1)$$

When Fred responds to the zoo-keeper, what he is claiming is that he hasn't seen an individual which this description satisfies. But this WORM does not help us in the case of the schoolgirl. For Fred has not seen *any* individual which can be described by this meaning. How then can he reply to the schoolgirl that he has seen a lion? In Braisby's (1989) account, this case implicates what are called Combinations of sc worms (or COWORMs). COWORMs can also be involved in word meaning. Indeed, a word's meaning may be given by the WORM which corresponds to it or by any one of a number of COWORMs. A COWORM is arrived at by combining a number of WORMs. Taking the "lion" case as an example, there is a combination of the WORMs for *lion* and *statue* as follows.

$$\begin{array}{l} \text{WORM(LION, P}_1) \\ \text{WORM(STATUE, P}_1, \text{P}_2) \end{array} \quad (2)$$

What the COWORM does is to relate various data structures. It does this by virtue of the fact that "statue" relates two types of data structures. That is, we may view it as an operator. Given a data structure of one particular type, we obtain a data structure of a different type. For example, this particular COWORM, relates the data structure above to one like the following.

$$\left[\begin{array}{l} \text{animacy:inanimate} \\ \text{colour:_} \\ \text{legs:4} \end{array} \right] \quad (3)$$

Here the "statue" WORM is being used to form a different kind of data structure. In particular, it tells us that a statue of a lion is inanimate and does not inherit the colour it has from the thing it represents. The data structure given by this COWORM describes perfectly what Fred *has* seen. Thus, Fred's reply to the schoolgirl is based on this (or a related) COWORM. The crucial aspect of this account is that it is COWORMs which involve a given WORM that may underly the meaning of that WORM's corresponding word. It is given that this is a condition on word use. The particular COWORM that is adopted by a speaker or hearer depends on a large number of factors, ones that have been discussed in the earlier parts of this paper. It is clear in this case though that it is the nature of the information sought from Fred that determines the particular WORMs and COWORMs he employs. This is not unlike the case of Rudy.

Let us suppose that each of our excited ballet lovers has the same WORM for the ballerina as follows.

$$\left[\begin{array}{l} \text{body: [shape: fine]} \\ \text{hair: [colour: brown]} \\ \text{dance: [move: pirouette]} \\ \text{personality: [temperament: quiet]} \\ \vdots \end{array} \right] \quad (4)$$

Here a subtlety in the process of combining WORMs is required. When "is delightful" is predicated of the ballerina by Rudy, his friends may do a number of things with the WORM that describes her. One option, that chosen by Reg, is to combine his WORM for the ballerina with that for *delightful* so the feature of delightfulness is added to that aspect of the WORM for ballerina concerning her dance as follows.

$$\left[\begin{array}{l} \text{body: [shape: fine]} \\ \text{hair: [colour: brown]} \\ \text{dance: [move: pirouette]} \\ \quad \quad \quad \text{is: delightful} \\ \text{personality: [temperament: quiet]} \\ \vdots \end{array} \right] \quad (5)$$

In a similar fashion, Ron combines his WORMs for the ballerina and *delightful*, resulting in a data structure in which "delightful" is a feature of the body attribute of the WORM for the ballerina. Similarly for Rudy. How, though, are we to capture the usher's interpretation?

The usher has a meaning for the ballerina that is the same as our ballet lovers, let us say. However, she has many possibilities open to her. She may combine her WORM of *delightful* with that for the ballerina in any number of ways. The number of options open to her are determined in two ways. Firstly, by the number of complex attributes of the WORM for the ballerina. Secondly, by the number of additional WORMs that she could employ. Similar to the "lion" case where Fred used the additional WORM of "statue" according to the appropriateness of the circumstances, so the usher has similar options too. If, for example, the usher is aware that certain members of the audience play all manner of language-games, then the meaning of *delightful* is simply indeterminate. That is, there are many possible COWORMs that we may take to define these games and the usher has no way of deciding which COWORM to choose. It is this indeterminacy that renders her unable to say what Rudy means.

Our other example involves Kim, Mary and Jane and the problems Kim has when Jane comes to stay. In particular, we want to say that Kim has a notion of what it is to be a biological mother but not an adoptive mother. However, she must come to acquire such

a notion. The crucial aspect of the general discussion has been that we do not want to conclude that Kim ends up with more than one lexical entry for *mother*. Again, we will use the mechanism of COWORMs.

We can describe Kim's WORM for *mother* as follows.

$$\left[\begin{array}{l} \text{gender:female} \\ \text{genetics:gen} \\ \text{child: } \left[\text{genetics:gen} \right] \end{array} \right] \quad (6)$$

How is it that given this WORM for *mother* Kim can later accept Jane's utterances concerning Mary? Again, we will claim that Kim learns of "adopting", which expresses a relation that Kim can utilise in forming the following COWORM.

$$\begin{array}{l} \text{WORM(MOTHER, P}_1\text{)} \\ \text{WORM(ADOPTIVE, P}_1\text{, P}_2\text{)} \end{array} \quad (7)$$

This COWORM will give a new kind of data structure, of a type related by the WORM for *adoptive*. This particular WORM acts to deny the genetic connection between parents and children. Accordingly, we get the following data structure which allows Kim to make sense of Jane's use of *mother*.

$$\left[\begin{array}{l} \text{gender:female} \\ \text{genetics:gen1} \\ \text{child: } \left[\text{genetics:gen2} \right] \end{array} \right] \quad (8)$$

Within Braisby's framework WORMs and COWORMs can be taken as defining different perspectives. Thus, we might say that Fred has two perspectives: one where he uses the WORM "lion" in isolation, one where he uses it in combination with the WORM for *statue*. Similarly, what Kim has learned is not a new lexical concept or WORM for "mother" but a new way of Combining WORMs (or a new perspective) and similarly with our long forgotten ballet lovers. In a similar vein, we may envisage "concept combinations" being accommodated in this fashion. Let us take as an example, "stone lion". It seems that we would like to say that, at the very least, "stone" implies inanimate while "lion" implies animate. Such an intuitively appealing view precludes the standard set intersective account. Alternatively, the feature structures which describe the component WORMs will fail to unify. Here, we would expect one or other of the component WORMs to be extended so that unification will succeed between the extended senses. Just as before, a Combination is required. Combining the WORMs for "lion" and "statue" renders a sense for *lion* that implies inanimacy. The feature structure described by this COWORM can then unify with that for "stone". However, just as we may view this as taking a particular perspective on what "stone lion" means, we can imagine that this combination allows for a number of different perspectives to be taken. So, although we may take the "statue of a lion" perspective defined by the COWORM of "lion" and "statue", we could also take the "petrified lion" perspective on "stone lion" which is defined by the combined WORMs of 'lion' and 'petrified'. Given suitable intonation and stress, we might also adopt the perspective which gives the meaning "lion that sells stones" which is what we would want to say in the case of "coal merchant" which means "one who sells coal". And so on.

We have seen that by considering the meaning of a word as being described by different data structures on different occasions of use, we open up the possibility of capturing flexibility and specificity. In addition, coherence is captured by the use of COWORMs which we may think of as embodying atomic, local theories. This is consistent with the objectives of the sense generation position.

5.2 View 2

This section presents some of the central aspects of the view of lexical representation and combination sketched in Franks et. al., (1988), and developed in Franks (1989). The outline given here is a simplification of that work in that it will focus primarily upon concept combination. We will consider the way in which one concept's extension constrains that of the other in a head-modifier relationship. In this way, the combined extensions result in a single sense for the noun phrase. The properties descriptions of the lexical concepts correspond to "Central" properties, derive from conventionalised lay theories. These include Central-essence properties, which relate to the ontological or functional essence of an entity.

Let us again consider our ballet buffs. Suppose that the ballerina is known to them as "Netty". Suppose further that Rudy's first comment had been:

(1) Netty is a delightful ballet-dancer.

Each of the meanings of Rudy's comment, for the ballet-buffs, can be seen to interpret (1) in different ways. On a simple account, we might say that this combination involves the concept for *ballet-dancer* being extended to incorporate the property of delightful. Hence, we might then say that the interpretation of (1) is (2).

(2) Netty dances ballet delightfully.

This is Reg's interpretation. It is based upon an assumption of Explicit Semantic Attachment, in which the major source of the properties of the sense for a NP is the explicitly given head noun, and the properties contributed by the modifier will be evaluated with respect to this noun (Platts, 1979). In (1), the explicitly attached noun is "ballet-dancer". Franks argues that, in order to capture semantic flexibility we must allow for Implicit Semantic Attachment: that is, agents may evaluate the property of being delightful with respect to a different noun concept. How is this effected? The initial combination of the modifier with the head noun unifies their properties and results in a sense which carries the meaning that Netty, who is a ballet-dancer, is delightful in some unspecified way. This is the sense which is generated by the usher. Her sense is compatible with many different extensions, in which the general property of being delightful is determined as a more specific property by evaluating it with respect to an implicitly attached noun concept. For example, Ron's meaning involves a choice of a noun concept such as "people seen in terms of their bodies", which results in an appropriate specification of being delightful. We then have this sense for (1):

(3) Netty is a ballet-dancer who is delightful to look at.

And Rudy's sense for (1) will similarly involve the selection of another implicitly attached N concept. In these cases the different agents are adopting different perspectives on the ballerina. We noted in 4.2. that a perspective defines a description of an entity. Franks argues that perspectives have two components: the categorising sortal noun or nominal phrase (either implicitly or explicitly attached to the construction which describes the entity), and the range of properties which are used to make the categorisation. In this case, we have perspectives whose major contrast is given by their different covering sortals. In contrast, the perspective of the usher has a less specific range of properties, and no particular sortal noun. Consequently, for the usher, the meaning of Rudy's phrase is indeterminate. The flexibility in perspectives will be further evident in the consideration of *stone lion* and similar combinations.

If combining concepts constrain the extension of the head noun, then they may act as defeaters for some of the properties in the noun's lexical concept. This is exactly what happens in what Franks labels Negating Functional Privative combinations, such as *stone lion* and *wooden banana*. These behave in a similar way to Negating Privatives Proper, which involve adjectives such as *fake* and *false*. The latter have usually been defined as denying such inferences as that from

(4) This is a fake gun.

to

(5) This is a gun.

Franks argues that this conclusion omits the crucially perspectival nature of such categorisations, which we will now examine.

Consider the combination "fake gun". The lexical concept for *gun* might include such Central-essence properties as "internal mechanism for propelling bullets" and "barrell for directing bullets", and other properties such as "trigger mechanism", "barrell mounted over trigger", "colour and weight of metal", and so forth. The lexical concept for *fake* specifies that it will defeat the Central-essence properties of "gun", and allow the ascription of some of the "appearance" properties of "gun" to the sense for *fake gun*. The modifier, then, acts as a Rebutting defeater for the noun's central-essence properties, that is we can ascribe the negation of the properties to the sense. After this combination, where the modifier has priority, we find that the sense for *fake gun* parallels the usher's for *delightful ballet-dancer*. That is, it includes only the general property that the entity has some of the appearance properties of the head noun. But we cannot say in precisely which way it looks like a real gun. It is consistent with many different extensions of this 'appearance' property, through the selection of a particular implicitly attached noun. This is consistent with our intuition that there are many possible ways for a gun to be fake. An entity of the type described by this sense would be indiscernable from a range of entities of the type of possible implicitly attached noun concepts, which might be labelled "toy gun", "replica gun" or "model gun". This means that we might adopt a categorising perspective on such an entity based on the narrow range of "appearance" properties, with respect to any of these nouns, as well as the head noun. That is, we might categorise it thus:

(6) The fake gun is a gun (or toy/replica/model gun) with respect to appearance.

Since this perspective does not incorporate any Central-essence properties, it could not be generalized to a broader range of properties, and hence the categorization is restricted in its applicability. We will refer to such restricted perspectives as Type II Perspectives.

If there is an informational requirement to make the sense more specific, then our agent will access an implicitly attached noun in order to extend the sense. This choice is crucially related to situational factors. If we know that the entity described by *fake gun* has been used to overcome some intruder in a house, we might then select "replica" as the appropriate noun type. This will be used to specify the detail of the "appearance" property, and the sense will also inherit the Central-essence properties of this noun. So the sense for *fake gun* includes properties to the effect that such an entity is a replica of a gun, which "looks like" a real gun in the way that a replica does. Now, we might adopt two perspectives on the entity described by the noun phrase. The first is the Type II Perspective:

(7) The fake gun is a gun with respect to appearance (as a replica).

The second is a Type I Perspective:

(8) The fake gun is a replica with respect to central-essence properties.

A Type I Perspective is one in which the categorised entity has the central-essence properties of the categorising sortal noun.

It seems that exactly the same kind of process occurs for a combination like *stone lion*. Here, the sense can define two different perspectives. Fred's situational context would facilitate a choice of "statue" rather than "rock-formation" or "ornament" as the implicitly attached noun. The Type II Perspective is:

(9) The stone lion is a lion with respect to shape (as a statue).

This is the perspective of the schoolgirl, which Fred adopts in order to respond as he does. The Type I Perspective is:

(10) The stone lion is a statue with respect to central-essence properties.

The perspective of the zoo-keeper would be a Type I: to satisfy his informational requirements, the entity described by *lion* would need to be a lion with respect to central-essence properties. This clearly cannot be supported by the sense generated for *stone lion*. In order for Fred to reply "yes" to any agent asking about the presence of lions, the perspective of the questioner must incorporate a range of categorising properties which subsumes the range of properties in the sense for *stone lion*. The schoolgirl's perspective satisfies this, but the zoo-keeper's does not. Franks argues that this is a requirement on the coherent use of a categorisation statement.

Our third example concerns Kim's quandary. This might be approached through a consideration of another type of privative combination, the Equivocating Privatives. These include combinations such as *blue orange* and *straight banana*, as well as ones involving adjectives like *apparent* and *alleged*. Equivocators are combinations in which we are unable to state whether the entity described by the noun phrase is in fact a member of the head noun or not. We simply have insufficient information. The modifier Undercuts the head

N's Central-essence properties. That is, on current information, we are unable to ascribe these properties to the sense: it does not mean, unlike in the case of Negators, that we can ascribe the negation of these properties to the sense. If our information changed, we might reverse the Undercutting.

Consider the combination "apparent friend". The Central-essence properties of the concept for *friend* might include "loyalty", "affection", "absence of self-seeking", whilst the other properties include "solicitous behaviour", "smiles", "various vague kindnesses". And the Modifier's lexical concept specifies the Undercutting of the head noun's Central-essence properties, and an ascription to the noun phrase sense of some of the non-central-essence properties. Having Undercut the Central-essence properties of the head noun, we are again left with a sense which is indeterminate as regards exactly how the "appearance" properties of the head noun are manifest. After accessing an appropriate implicitly attached noun, such as *deceiver*, the general property is made more specific. This sense then defines a Type II Perspective on an entity described by the noun phrase:

(11) This apparent friend is a friend (or deceiver) with respect to appearance.

In order to decide whether this apparent friend is a real friend, we need to extend our information through some process of inquiry. The outcome of this will determine the appropriate combinations of perspectives. In an Affirmative outcome, the apparent friend is a real friend, thus reversing the Undercutting of the Central-essence properties of the head noun, and allowing for a Type I Perspective:

(12) This apparent friend is a friend with respect to central-essence properties.

In a Privative outcome, the sense defines the same combination of perspectives as in the Negating Privatives case.

Let us return to Kim's quandary. This follows a pattern similar to Equivocators. In trying to understand Jane's use of the word *mother*, through trying to adopt Jane's perspective, Kim will Undercut the central-essence properties of her concept of *mother*. Hence she could use this sense to define Jane's perspective, which is of Type II:

(13) Mary is Jane's mother with respect to behaviour.

This allows her to have some understanding of how Jane can call Mary "mother". However, Kim may learn of "adopting", which enables her to determine that, from Jane's perspective, the sense for *mother* involves a Rebuttal of the central-essence properties of the lexical concept for *mother*. But as in the Negators case, this does not prevent Jane from calling Mary "mother": it simply means that she must adopt a particular type of perspective to do so. As in View 1, this implies that Kim is not learning a new lexical concept for *mother*, as in our interpretation of Lakoff, but that she is able to adopt a new perspective - a Type II Perspective - on her mother as Jane's mother.

On this view, then, the flexibility and specificity of a wide range of combinations are addressed by the twin mechanisms of defeasible extension and implicit semantic attachment. Both are viewed within the context of the perspective-relative nature of categorisation. Accordingly, coherence is relative to the two aspects of a perspective, and both must be appropriate for a coherent categorisation.

6 Conclusions

In this paper, we have outlined a general account of lexical representation and presented two approaches within this account and the way they treat the problems of coherence and concept combination. Moreover, we have demonstrated that both approaches promise an adequate treatment of the crucial phenomena of semantic flexibility and specificity. These approaches have been motivated by arguments concerning coherence in sense selection.

Strong sense selection has unfortunate consequences especially regarding the combinatorial explosion of interpretation associated with complex expressions. We take it that this position is not only unfortunate but untenable. Weak sense selection, which is assumed in a number of linguistic approaches, avoids an explosion of this magnitude. However, we have argued that the position is undermined by a consideration of coherence. We concur with the general position of Murphy & Medin that coherence is determined by theories. As a consequence, we expect the relations between senses to be expressed in any theory of lexical representation. This conclusion is in direct opposition to the starting position of weak sense selection. Our conclusion is that such a starting position is unprincipled. In order to have a principled account of the phenomena we have discussed, it seems a sense generation account is required.

An important implication of the sense generation approach is that lexical concepts should be regarded as partial. The specificity of meaning we need is brought about by the possibilities of concept extension. The mechanism of concept extension, in allowing for the defeat of properties, also provides for the desired flexibility of meaning. Importantly, the constraints on extension are provided by theories which, as we have seen, underpin the coherence of the generated sense. We have strongly argued that theories are local: they may be provided by the immediate situation. In addition, agents can choose to utilise certain theories in preference to others on the basis of their informational requirements. This at once extricates us from Murphy & Medin's connection between metaphysics and theories. It also, crucially, leads us to the conclusion that an agent adopting different theories in his cognitive activities is entertaining different perspectives. Consequently, the examples which we have employed throughout this paper are essentially perspectival in nature as, indeed, is word meaning.

References

- Braisby, Nick. 1989. Situating Word Meaning. Paper presented at the Conference on Situation Theory and its Applications, Asilomar, CA.
- Clark, Herbert H. 1983. Making Sense of Nonce Sense. In d'Arcais, G. B. Flores. and Jarvella, Ray J. (eds.). *The Process of Language Understanding*. Chichester: John Wiley and Sons. 297-331.
- Dunbar, George L. 1988. The Cognitive Lexicon. PhD. Thesis. Centre for Cognitive Science, University of Edinburgh.
- Murphy, Gregory. L. and Douglas. L. Medin. 1985. The role of theories in conceptual coherence. *Psychological Review*, 92, 289-316.

- Franks, Bradley. 1989. *Concept Combination: Towards an Account of Privatives*. Edinburgh Research Paper in Cognitive Science.
- Franks, Bradley, Terry Myers, and Scott McGlashan. 1988. *Defeasibility in Concept Combination*. In *Proceedings of the Tenth Annual Conference of the American Cognitive Science Society*, Montreal.
- Lakoff, George. 1987. *Cognitive Models and prototype theory*. In Neisser, U. (ed.). *Concepts and Conceptual Development: Ecological and Intellectual Factors in Categorization*, 63-100. Cambridge: Cambridge University Press.
- Quine, Willard. V. O. 1960. *Word and Object*. Cambridge, Mass.: MIT.
- Platts, Michael. 1979. *Ways of Meaning*. London: Routledge and Kegan Paul.