

The Cognitive Lexicon

George Luke Dunbar

Ph D

Department of Psychology and Centre for Cognitive Science

University of Edinburgh

1988



I hereby declare that I have composed this thesis myself, and that the work is my own.

George L. Dunbar

30th September 1988

Acknowledgements.

I acknowledge the support and enthusiasm of my supervisors Terry Myers and Jim Miller. Jim's influence on this is more profound than he may realise. I'd like to thank the Centre for Cognitive Science (formerly School of Epistemics) for being a good place to work. SERC provided the dosh for most of the period of study, and the ESRC forked out for a short period of study in Italy - ta guys.

On a more personal note, I want to thank Gill and Romy for their support, even at the end. Spider Systems, thanks for the job. Professor Zampolli, thanks for Pisa.

Abstract

The thesis is constructed on two levels. At one level, it is an attempt to show how the interdisciplinary approach and methods of cognitive science can be used to advance understanding of the mental lexicon. To that end, the thesis begins by clarifying a methodological relationship among the techniques available.

At a second level, the thesis seeks to advance our understanding of a range of phenomena which illustrates the flexibility of lexical meaning. The underlying reasons for flexibility are analysed. Existing theories of lexical concepts are evaluated in the light of these arguments and are found wanting. An outline theory is proposed which takes into account integration and precization. This theory is then applied to a particular phenomenon - the mass—count alternation in English, and a variety of linguistic criteria are used to test the questions raised. A Space Grammar model of the theory is developed, and a means for translating it into Unification Categorical Grammar is provided. The model is then generalised to related linguistic phenomena. Several psychological experiments are proposed which could test the model. A concluding chapter summarises the model and discusses some of its implications.

The Cognitive Lexicon

Introduction	1
Part 1: Groundwork	2
Chapter 1 Prolegomenon	3
Chapter 2 The lexicon in time and in context	9
2.1 Semantic timing	9
2.2 Semantic development	15
2.2.1 Elaboration	15
2.2.2 Undifferentiated and non-egalitarian reference	19
2.2.3 Referential perspectives	30
2.3 Conclusions	32
Part 2: Characterising Lexical Concepts	33
Chapter 3 The importance of notation and system	34
3.1 Notation	34
3.2 Word meanings and distinctive features	39
3.2.1 Learnability	41
3.2.2 Derivation of relations	42
3.2.3 Interaction with context is inherent	45
3.2.4 More and less contrast	46
Chapter 4 Lexical concepts as prototypes	49
4.1 Prototypes can't be combined compositionally	52
4.2 Lexical indeterminacy of types	57
4.3 Coherence of prototypes	58
4.4 Instability of prototypes	59
4.5 Summary	60
Chapter 5 Lexical Concepts	61
5.1 Lexical concepts, theories and lexical flexibility	66
5.1.1 Diachronic change	66
5.1.2 Metaphor	67
5.1.3 Classical concepts	70
5.1.4 Contextual combination	72
5.1.5 Internal lexical structure	74

Part 3: From Lexicon to Phrase	75
Chapter 6 Dependency and Lexical Vagueness	76
6.1 Dependency and type emergence	77
6.1.1 Semantic tailoring and dependency	78
6.1.2 Transitivity and dependency	80
6.1.3 Aspect and type emergence	84
6.1.4 Vagueness and type emergence	87
Chapter 7 The mass-count distinction in English	89
7.1 What is it?	89
7.2 Is it lexical?	92
Chapter 8 A phrasal approach	109
8.1 Delimitation	115
8.2 Number in English	118
8.3 Relating the primitives to the mass-count distinction	119
Part 4: Theoretical Development	123
Chapter 9 Modelling and clarification of the analysis	124
9.1 Space Grammar	124
9.2 Unification Categorical Grammar	130
9.3 Classifiers and the partitive	133
9.4 Aspect and the mass-count distinction	149
9.5 Lexical concepts and theories	155
9.5.1 Lexical concepts	155
9.5.2 Theories	156
9.5.3 Theories, lexical concepts and the mass-count distinction	157
Chapter 10 The role of experimental psychology	160
10.1 Method	163
10.1.1 Procedure	163
10.1.2 Subjects	164
10.1.3 Materials	164
10.1.4 Design	167
10.2 Results	167
10.3 Discussion	172
10.4 Proposals for experiments	173
Chapter 11 Conclusions	177
References	185
Appendix A Materials randomisation program listing	208

Introduction.

This thesis is concerned to give a cognitively tractable account of lexical meaning. The starting point is a set of observations of the flexibility of lexical meaning. My aim is to understand these phenomena better, to understand how they relate to one another and to develop an interdisciplinary approach to the problem. I begin by looking at method. On the one hand, how can the tools of cognitive science be used to address the issues at hand. And on the other, to what extent does the way the problem is defined affect the theory that can be developed. I then consider alternative theories of lexical concepts. The notion of systemic contrast and the idea that lexical concepts are concepts are considered before I propose my own theory of lexical concepts. Issues raised by this discussion are then examined in relation to specific phenomena, especially the mass—count distinction in English, and a more general theory of lexical concepts is stated. Experiments which could contribute to the development of this theory are proposed before I make some concluding remarks.

Part 1.

Groundwork.

The chapters in this part of the thesis lay down a broad outline of the framework within which I hope to assemble a view of lexical meaning. The first chapter interrelates phenomena and research techniques in a general way. The second begins a process of defining the problem in such a way as to leave it solvable. There are certain classes of phenomenon which an adequate theory ought to take into account. Indeed, I take a stronger view. The task of building a satisfactory theory will be harder if these key phenomena are not taken into account.

Chapter 1.

Prolegomenon.

The purpose of this chapter is to make explicit a framework within the terms of which one might do interdisciplinary research on lexical meaning. Such research offers the potential to bring to bear the accumulated wisdom and methodological accumen of workers with diverse perspectives on the problem of explicating the nature of lexical meaning.

As Gibson has shown in relation to vision, movement (from one perspective to another) adds information rather than just complexity. And as J.S. Mill enjoins us in relation to ideas, a deeper understanding is facilitated by the consideration of a variety of points of view. That is, it is not simply a matter of avoiding the re-invention of the wheel: interdisciplinary activity may - in itself - yield new knowledge and, perhaps, new research techniques.

Having said these things, I will confine myself here to some general principles.

First, the research is defined in this case by the topic to which it is addressed. This can be made more specific. I am concerned with the role that lexical items play in establishing the meaning of a piece of text. That is, I have in mind the characterization of meaning advanced by Stephen Ullmann (Ullmann 1962). This conception will act as a focal point for the study which follows: the role of any information will be determined by its relevance to the goal of developing an understanding of the place of lexical items in the comprehension of language.

In the paragraphs that follow, I will briefly connect with some departments of study, sketching the relation of each to the whole. I begin by listing some key phenomena which at the same time appear interesting and offer the prospect of shedding light on the central topic. I then look at the principal methods available. It will be a matter of stipulation that what I seek to characterize is the nature of lexical concepts: a lexical concept is that part of

the information associated with a form that contributes to the interpretation of a text.¹

This definition of lexical concepts turns out to exclude the notion of a prototype, where that is defined as a (possibly structured) set of default values along certain dimensions to which exemplars can be compared and rated for typicality in terms of some similarity metric (examples departing least from the defaults being the most prototypical). This is because such a definition assumes a similarity metric which is independent of knowledge, fails to provide a stable or combinable representation and provides a representation that is too specific. Prototypes are more like the result of an interpretative act than its seed.

The following paragraphs outline some characteristics of words, and attempt to place them in an overall framework.

It is well known that "the word" has many aspects, and as a term can lead to confusion. For instance, it has form (acoustic, orthographic or even kinetic²), morphological structure and distributional characteristics as well as a denotation. It is worth remarking that all these contribute to meaning as it has been defined here. Ullmann (Weinreich 1963b) has noted that signs can vary in terms of their "transparency" at each of these levels — that is, in the directness of their relation to their denotation. In addition to simple cases of onomatopoea, sound symbolism appears to have an effect (Taylor and Taylor, 1962) at what we can guess is the level of synaesthetic metaphor (Gardener 1974). Furthermore some words exhibit their etymology and can be said to be, to that extent, more transparent semantically. Compare *grandfather* with *uncle* for example. On distribution, there is a substantive controversy as to whether, for instance, the part of speech of a word influences the character of its denotation. Note that this is precisely the question of linguistic relativity as it applies to syntax: does it make a difference cognitively whether a language picks out a given entity using one part of speech or another?

¹ The term "text" should be construed to include both written and spoken language, of course.

² That is, in the case of sign language. I set aside for the time being the observation that, for example, spoken language has kinetic, articulatory and visual aspects in addition to its acoustic persona.

Different contexts may vary in the focus they place on these levels, but typically for linguistic signs the symbolic function (ie. the denotative relation) is to the fore. Nevertheless, other aspects of the word may be drawn upon in a manner closely analogous to what Weinreich described as the hypersemanticization of language (Weinreich 1963a). That is, in both cases greater than usual weight is given to one or more aspects: the word is pregnant with meaning. The complement of hypersemanticization is the phenomenon of desemanticization undergone by morphs which, for example, form a compound.

Having mentioned synaesthesia, I cannot avoid observing that concepts (which are traditionally taken to be the denotation of words in psychological accounts of meaning) have strata too. Most obviously they exist, to use Piagetian terminology, on both the behavioural and reflective planes of thought. But, on analogy with the brain which has primitive structures underlying the neocortex, concepts must tap reservoirs of passion as well. This aspect of meaning is to some extent dealt with in the work of Osgood.

Anomaly is characterized in a later chapter as a lack of transparency in the combination of words, and this is attributed to the content of the background knowledge a language user brings to bear on his task. Background knowledge is characterized as a set of naive theories (some transient) about reality and the current situation which give conceptual flesh to the skeletal lexical concept. This approach takes categorization to be strictly secondary in relation to lexical concepts, being primarily a function of knowledge of the world. Lexical concepts are held to be incapable in themselves of sustaining reasoning, which is viewed as a conceptual phenomenon.

It is an important fact about language comprehension that it takes place in real time. The comprehender must dynamically allocate processing resources to the interpretation of the words that are successively uttered. Additionally, in this same window of time, the senses of words must be combined. Indeed, Chomsky's arguments concerning creativity show convincingly that first, some stable lexical representation and second some regular method of combination must exist. Clearly, the rapid pace of comprehension limits the process of inference (which must be required to access background knowledge), and we can safely predict that a word's linguistic neighbours (among other contextual constraints) serve to gate this process in line with the Conversational Principle of "relevance". Moreover, there exist syntactic devices (broadly, "foregrounding") which serve to direct the focus of the comprehender's attention to particular items. Clearly, an adequate representation must take

these things into account. In particular, how do lexical concepts interact with the knowledge of the world which must underpin reasoning on the one hand, and linguistic context on the other? And what consequences do these functional responsibilities have for the structure of the lexical concept?

I have said above that concepts are frequently proposed as the denotation of lexical items. It is desirable, however, to connect the lexicon to external reality as well, since that appears to be what language is "about". Note that on the account I develop, the naive theories of the language user are held to provide that link and that, nevertheless, one requires to develop a fairly complex notion of self (ie. internal life) in order to describe and distinguish reference to properties and experiences of the self. Furthermore, it is valuable to fashion a distinction between the indexical and the depictive functions of language, even though these are very often mingled in individual utterances of a given word.

Before ending this chapter, I wish to make a few brief remarks about method and, especially, how the research tools of different disciplines can be used in relation to each other. Science I take to be defined by a way of proceeding to conclusions (viz. rationally and with a willingness to address uncertainties) which is shared by any reasonable person, whether he describes himself as a scientist or not. The secondary element that distinguishes science from other rational enquiries is the relationship between clarification (ie. philosophy) and empirical evidence; in science, the latter makes the running. And in Cognitive Science, clarification still has a major role to play. As well as philosophy, the apparatus of formal modelling and computer modelling (a-intelligence) can aid the process of clarification, with their enforcement of high standards of rigour adding something to mere armchair reflection. The other advantages of modelling have been well described by, for example, Chomsky (Chomsky 1965) and Hill (Hill 1987), and I expect that it's unnecessary for me to repeat them here.

On the empirical side, there are several sources of possible evidence. There is "linguistic" evidence; that is, sentences uttered (or constructed artificially) and the intuitions of native speakers about certain properties of those sentences (such as grammaticality or paraphrase relations). The storehouse of information amassed by traditional linguists and modern typological work are worthy of special mention here. (It is not necessary, however, to assume that it will be helpful to make a sharp distinction between what is traditionally classed as semantic information on the one hand, and syntactic information on the other (see Dunbar

1984).) As well as the kind of behavioural data generated in experimental psychology, subjects can also provide more extended reflections on their subjective experience of words.

In relation to the cognitive lexicon, and the type of meaning I am concerned with, let us first make a distinction between two meanings that a word might be said to have. On the one hand, there is the contribution that a word makes to a perceived meaning or understanding of an utterance. That is, a word plays a role in the cognitive event of comprehension, and its meaning to the comprehender is what it adds to the message. On the other hand, there is the meaning which users of language ascribe to a word through a metalinguistic process of dividing up the overall meanings of utterances among their constituents. This is lexicography. The difficulty with a lexicographic approach is precisely that it tends to view the word in isolation and at leisure, hence omitting from consideration the rapid march of time and the swift intrusion of other words upon the scene. As the physicist Bohm observed, when the analysis of any system excludes a factor whose influence upon the system is systematic, that factor will appear to have a random effect. In other words, theoretical advance depends on the analysis including any influential factors. This is a central doctrine of ecological psychology, and so this thesis can be seen as influenced by that movement to the extent that I emphasise the neglected factors of time, contextual interaction, combination and dependency.

Cognitive linguistics, a relatively new way of doing grammar which owes a great deal to Uriel Weinreich, provides descriptive tools which attempt to represent the structure of language and meaning in a psychologically realistic manner. It can be thought of as founded on a re-working of Chomsky's competence-performance distinction; the domain of idealization is contracted by drawing into the scope of competence such factors as the nature of perception and memory.³ Such an approach is, obviously, in tune with the perspective adopted here: if modelling is to lead to clarity then the model should encompass the determinants of the phenomenon being modelled.

³ It is crucial to realise that cognitive grammars are not typically presented as theories of performance and also that, indeed, they vary as to which "psychological" factors they set out to incorporate.

Given this, the relation between linguistic and behavioural data is complex, and we cannot simply regard the latter as providing a test of the hypotheses generated by the former: they access the phenomenon at different levels. I follow Chomsky in assuming that regularities at the linguistic level must arise from regularities at the cognitive level, but that the relationship between the levels may be sufficiently abstract to preclude testing the models from one with data from the other. Rather, they provide complementary data, with linguistic evidence establishing the general framework which behavioural evidence can then refine. At a practical level, each will inspire the other periodically, in the manner of a dialectic.

This chapter has outlined the framework within which the thesis which follows has been constructed. The optimal marriage of the traditional disciplines which feed into Cognitive Science has not yet been made, I am sure. Nevertheless, I hope that the relationship I adopt will turn out to be both coherent and productive: it is for the reader to judge.

Chapter 2.

The lexicon in time and in context.

My goal is to clarify the nature of the semantic representations which are associated with words and which mediate between form and meaning in language understanding. Two criteria will be emphasised: an adequate lexical entry must (1) be able to fulfil a function and (2) do so in a manner consonant with available psychological data. The function is to supply semantic information allowing the hearer to comprehend the speaker. This chapter outlines some factors whose role in perception influences the way I characterise lexical meaning.

As a first step, let me briefly review data available from studies of speech perception and reading, which bears on the issue of "semantic timing". By semantic timing, I mean the pattern in time of the availability of meaning. To summarise in advance, it will emerge that meaning becomes available very quickly, but that meaning evolves over time with processing exhibiting a periodicity linked to the structure of the discourse. After examining this, I will discuss the important role of development in the comprehension process, suggesting that variation in the degree of analyticity of processing is a property of human perception, and of human perception of language in particular. This evidence has important consequences for our understanding of the relationship of word meaning to comprehension. First, then, semantic timing

2.1. Semantic timing

Several studies, using a variety of experimental methods, have highlighted the role of semantic information in the recognition of words in speech and text. It appears that the processing system uses semantic interpretation of the earlier parts of a discourse to generate

expectations about subsequent parts.⁴

Miller and Isard 1963, extending the work of Miller, Heise & Lichten, found that words were recognised more accurately if they appeared in strings that made semantic sense as well as just possessing syntactic order. This early finding implies that the meaning of words in a string is employed in the identification of their neighbours, which is a particular example of the role of top-down information (information about what the perceiver expects) in the perception of language. Similarly, Marslen-Wilson and Tyler 1980 present evidence that subjects could identify words that rhymed with a target word faster in a string of "normal" prose which was semantically constrained than in nonsensical "syntactic prose". Furthermore, this superiority is apparent from the beginning of a clause if it is preceded by a context sentence.

More recently, Ford (1983, reported by Kennedy, note 1) found that a task in which subjects had to continuously make lexical decisions (ie. to decide for each stimulus whether or not it was a real word), reaction time was reduced if successive items cohered semantically rather than just syntactically.

Marslen-Wilson has developed a speech-shadowing task for psycholinguistic research which demands that subjects repeat auditorily presented strings verbatim. Using this task, Marslen-Wilson 1973 and Marslen-Wilson and Welsh 1978 found a class of errors (made even by subjects whose shadowing lagged by only 250ms) which they call "constructional errors". These involve the subject's output deviating from the input but in a way which conforms to local syntactic and (especially relevant here) semantic restrictions. The more recent of these studies found that subjects corrected phonetic errors in the input string more often when the identity of the word had been more strongly constrained syntactically and semantically.

Morton and Long 1976 using a phoneme monitoring paradigm (see eg. Foss 1970) reported that the higher the transitional probability between a pair of words (the likelihood of the pairing having occurred in the past), the faster reaction time was to a phoneme on the second word, indicating that information about syntactic and semantic links between adjacent words (of which transitional probability is a manifestation) frees processing capacity to

⁴ Parts of this section appeared originally in my undergraduate project, Dunbar 1984.

facilitate monitoring or, in other words, that less processing is required to identify a word if it is strongly linked to the previous word. Once more, this implicates the employment of the meaning of one word in the recognition of its neighbours.

Another task which yields evidence supporting on-line access to word meaning is that of listening to a passage being read, knowing that somewhere in the reading there is a mispronunciation which must be detected. Cole and Jakimik 1980 report their work of 1978 in which subjects reacted faster when the mispronounced word was highly likely to have followed the previous one (likelihood defined as transitional probability in the same way as Morton and Long above) - this was in addition to the finding that RT was faster if the mispronounced word had been implied in a previous sentence or suggested in the story title. Cole and Perfetti placed the same words in predictable and unpredictable contexts and found that variously aged schoolchildren and college students all reacted 200ms faster to a word being mispronounced if it appeared in a predictable context, suggesting that these subjects made the predictions "on-line".

Indeed, Marslen-Wilson and his colleagues have demonstrated that, in context, words are often identified before sufficient acoustic-phonetic information has become available to identify the word (see Marslen-Wilson 1986). On average, identification of words takes around 200ms, but 200ms of acoustic-phonetic information is typically insufficient to isolate a single item and tends to leave a large cohort of possible word candidates. Gating tasks, in which words are gradually revealed to subjects in, for example, 50ms increments, also indicate that recognition requires less of a word in sentential context than out of it. This phenomenon is known as "early selection". For semantic expectations to be able to play a role, of course, semantic information about competing items must be available during assessment, and these researchers have demonstrated that associates of contextually appropriate items do not receive preferential facilitation until after the recognition point. (This result was obtained using a cross-modal priming task in which subjects heard sentences containing a word-form; visually presented associates of any item consistent with the current acoustic-phonetic information were easier to differentiate from non-words, but only up to the recognition point). We should note, however, that - as Foss 1982 points out - it is only the relatively enduring post-recognition facilitation that is likely to play a role in assisting perception of subsequent material.

Further evidence for meaning being processed word by word lies in MacKay and Bever 1967's finding that the further right in a string a lexical ambiguity is located, the longer it takes to spot the ambiguity.

We will not be concerned here with particular models of speech recognition and their relative merits, since it is of little relevance to present concerns whether, for example, top-down information can influence acoustic-phonetic analysis. The key point is that context provides "a structured interpretative framework against which the senses associated with different word-forms can be assessed" (Marslen-Wilson, 1986 p40) and that words contribute to that framework rapidly. Taken together, these lines of evidence demonstrate that meaning is extracted from words as they are identified one by one from left to right. The evidence does not imply, however, that the whole meaning of a string is evolved in this process, and I turn now to evidence that elaborative and integrative processing of meaning continues in time.

It ought to be pointed out that there is evidence of "late recognition" of words. For example Shillcock, Altmann and Bard (1987), demonstrate that a large proportion of lexical items are not recognised until a point in time subsequent to the offset of following words.⁵ Clearly, if words are not identified they cannot contribute to meaning. Similarly, Warren and Sherman 1974 have found that perception of the word [*cough*]-eel in the sentence:

- (1) It was found that the ?eel was on the axle/shoe/etc.

depended on which noun appeared in the prepositional phrase. This finding of a "backwards Warren effect" is important since it suggests that identification of an item can be suspended pending disambiguation by context without derailing the entire comprehension process. (See also Ganong 1980). Mackay and Bever's (1967) evidence that it takes longer to find a second meaning for a multiply ambiguous sentence than for a singly ambiguous one was interpreted by them to imply that the identification of ambiguous words was suspended until context biasing one meaning or the other was found, on the grounds that in a multiply ambiguous sentence it is difficult to converge on a single interpretation.

⁵ It would be of considerable interest to ascertain whether words likely to be recognised late nevertheless still produce the priming effects observed in the studies of lexical access discussed above.

Metaphorically, the notion is of trying to ascend a muddy slope rather than proceeding step by step up a staircase: in a multiply ambiguous sentence there are fewer footholds. In sum, we have three pieces of evidence that the comprehension of language need not (indeed, on occasion, cannot) proceed word by word.

There is also considerable evidence of an association between the phrase structure of a discourse and processing effort. That is, the human processor appears to deal with the incoming information in phrasal chunks. Such a model of perception makes it easier to see scope for context sensitive processes of integration and elaboration (see below). The data supporting this assertion is of essentially two types: on the one hand, events near phrasal (particularly clausal) boundaries tend to be perceived as being at the boundary (eg. Bever, Garrett and Hurtig 1973, _____), or **better** at the boundary (Wingfield Klein 1971, Caplan 1972, Foss 1969)⁶. On the other hand, eye movement studies find longer fixations at the ends of syntactic units (clauses in particular) (eg. Just and Carpenter 1980).

The evidence for clausal processing is reviewed in Bever, Garrett and Hurtig (1973), Fodor, Bever and Garrett 1974 and Marslen-Wilson, Tyler and Seidenberg 1978. The paradigmatic click-insertion experiments required subjects to listen to strings into which a non-linguistic acoustic signal (a click) had been placed. The subject had to identify the location of the click. Garrett, Weber and Fodor () found that the click was perceived in a location nearer the major syntactic boundary in their experimental sentences than it had actually been. This suggests that the clause is in some sense a perceptual unit. Further evidence is provided by Wingfield and Klein (1971) who report that changes of channel are located more accurately if they occur at clause boundaries.

As Marslen-Wilson et al (1978) point out, this evidence suggests that the surface structure of a sentence is perceptually real, but in the absence of direct evidence for what they term "clausal processing", it doesn't explain **why** it is perceptually real. They develop the thesis that the key factor is informational completeness: the completion of a clause typically rounds off a set of previously generated expectations. They suggest that:

⁶ The event is typically the occurrence of a click in the acoustic signal, but in Wingfield and Klein (1971) it was a change of channel the signal was presented in (it changed to the right ear from the left).

"It is never the case that interpretative activity is not being attempted, but it will often be the case, as it were, that this activity will raise more questions than can be immediately answered." (Marslen-Wilson et al, 1978 p225).

The prediction they make is that identification time differences between targets specified by either rhyme or category cues will be constant unless there is a change in the amount of contextual aid. That is, if semantic information is less available at the beginnings of clauses then the category task should suffer more, with the rhyme task being relatively unaffected, when a clause boundary is encountered. Also, differences in identification time either side of the clause boundary should be attenuated if the first clause is not informationally complete (eg. if it contains a cataphoric pronoun). And, indeed, it is found that the difference between category and rhyme monitoring is constant at all positions, and that there is slowing immediately after a clause boundary only if the clause is informationally complete. Thus processing does exhibit a periodicity, but it is as a function of informational completeness.

Townsend and Bever 1982 examine in some detail the notion that integration with semantic context depends on the formation of a "propositional unit". For example, they find that if an item suffixed with *-ing* forms a propositional unit (if it is a gerund, for example) then it closes perceptually and integrates with context, thus slowing continuation reading times for a subsequent inappropriate copula. On the other hand, if it does not form such a unit (if the stem is an adjective, for example) then such effects are not found.

Carrithers and Bever 1984 found, for a reading task, that fixation times are longer and less affected by word length at the ends of clauses (particularly those conforming to a canonical NVN schema; non-canonical ones get processed more fully on-line). Haberlandt and Graesser 1985 performed a multiple regression on reading times that were available on a word-by-word basis, thanks to their use of a serial, visually-extended, display format. They found interactions such that sentence imagery and the number of new concepts introduced by a sentence had a particular effect on the sentence-final word, with reading time for this word increasing linearly as the number of new concepts in the sentence increased, for example. They conclude, therefore, that the pause on the final word is occupied by integration of concepts.

It is a common finding that reading times for sentences are speeded by appropriate prior context. Observing that prior research has often used reading times for entire sentences, Sharkey and Sharkey 1987 set out to examine whether the locus of facilitation is integration with prior context or whether it has some other basis, by presenting texts a phrase at a time. They found that knowledge-based facilitation (that is, faster reading times for entities implicated in a previously activated script) occurred only for entities in the sentence-final position. The failure to find facilitation for other positions was not confined to sentence-initial position. Overall, these results are taken by the authors as support for a buffer model in which concepts are chunked (at clause boundaries) before being integrated into some discourse representation.

One implication of a relatively strong version of the word-by-word hypothesis (that as soon as a word is recognized it makes its full contribution to the meaning of the utterance) would be that that contribution can be determined at the time of recognition. This suggests that words have a fixed meaning which does not vary as a function of subsequent context, which - as many students of word meaning have observed - is an untenable view. There are many ways in which the meaning of a word may change subtly, and I will review the relevant data below. The evidence from the studies reviewed above, however, is clearly consistent with the view that meaning develops over time. It will be an aim of this project to provide a characterization of lexical meaning that allows for this.

2.2. Semantic development

The motivation for my interest in the perception of language becomes clear now. There are some processes which it would be desirable to find time in perceptual activity for: processes of elaboration and integration. The brief review has indicated that such time does exist. The next two sections contain some more evidence and some more ideas on how that time should be filled; the third outlines the general relationship between lexical semantics and "reference" that will be assumed.

2.2.1. Elaboration

One relevant feature of human performance is that word meanings are **flexible**. That is, on different occasions of use the sense of a word is capable of varying. Half, Ortony and Anderson 1976 describe, for example, changes in the sense of RED according to the identity of the noun it combines with. Barclay et al. 1974 discuss variation in the quality of a given property as a cue for a noun according to the nature of the verb; for example, *heavy* serves as a better cue for *piano* when the verb was *lift*. Barsalou 1982 demonstrates that the verb phrase in a sentence can affect the availability of properties of the grammatical subject. Let's examine this evidence in more detail.

Barsalou (1982), in rejecting the extreme views that on the one hand the whole meaning of a word is context-independent (CI)⁷ and that it is, on the other, a function solely of context (ie context-dependent, CD), used a semantic facilitation task to test for CI and CD components of meaning. Several seconds after having read a sentence, in which one noun phrase was signalled as the target by virtue of being underlined, subjects were cued with a "semantic property" whose ownership by the target they had to verify. The dependent variable was the time taken to make the verification, and Barsalou found that a certain subset of properties received verification times on average 145 ms slower when prior context had not reinforced them than when it had. These properties he classified as CD and opposed to properties verified equally quickly whether or not context reinforced them - CI properties.

This work can be directly related to that of Barclay et al. 1974 who reviewed then current theories of lexical representation and decried their emphasis on the inflexibility of word meaning. They assert, for example, that the representation of BALL in the two sentences :

(2) the man hit the ball

(3) the man hid behind the ball

is different. Although obviously the same in as much as the meaning is distinct from "formal dance", the representations are nevertheless said to differ in "size". The notion is that

⁷ Barsalou attributes this view to Katz who argued in fact that while context had a role, that role was too unwieldy to formalize, and was attacked for that stance by Bolinger 1965 who accused him of a false dualism and suggested that such "knowledge of the world" meaning as whether or not an entity typically wore shoes could be just as well expressed by semantic markers (binary features).

distinctiveness is not absolute but is a function of context. This important theme is discussed in more detail in the next two chapters. In an experiment, the meaning of the verb in sentences was found to affect which semantic properties functioned as effective memory probes for the object noun phrase. For example, the probe "heavy" cued "piano" more effectively if it had been lifted than if it had been "played". They found, moreover, that a given probe-verb combination couldn't retrieve any old noun phrase ("heavy" could not cue an "infant" which had been "lifted" as effectively as it could a "piano"), hence demonstrating that the property is a flexible possession of the noun and not an invariant attribute of the verb. The authors propose that :

"the way in which an unambiguous noun was psychologically instantiated was governed by the relevance of each of its semantic properties to the event described as a whole" p479.

In considering the Barsalou and Barclay et al studies, it needs to be remembered that there is a considerable delay between initial perception and subsequent response. That is, they employ successive tasks, evidence from which points to the conclusion that post-"on-line" processing develops the meaning of a word relative to its linguistic (and, probably, wider) context. This conclusion can be related to Miller 1977's argument that knowledge of the world operates on lexical entries to relate them to context and one another. Miller has, like many others, also argued (Miller 1978) that the extensive lists of senses of particular words which lexicographers turn out (see eg Nida 1975) are an artifact of not abstracting the analysis from contextual considerations.

Barsalou suggests two ways in which CD properties may be activated: either by inference rules (similar to Miller, 1977) or via the lexical entries which can release these portions of meaning if triggered by associated context. For example, he suggests that the meaning of EAT can be manipulated by the meaning of whatever is eaten:

(4) he ate the steak/soup/sandwich

Barsalou states :

"Given the existence of CD properties, the meaning of a word is not a fixed set of properties that is activated as whole every time the respective word is activated" p87.

Kunzendorf 1976 asked subjects to verify that species such as IRIS were "flowers" or "plants". The exemplars were presented in a list which was made up of flowers and either trees or weapons, and he found that subjects were faster to confirm exemplars as plants in the weapons list than they were to confirm them as flowers in the trees list. Kunzendorf argued that static theories of categorization could not explain why in certain circumstances it was easier to identify an entity as a member of a more general category, and that it was necessary to conclude that in different contexts different sets of feature would be attended to. Specifically, being a plant would distinguish a flower from a weapon but not a tree and so its plant-hood would be more salient in the former case. With this study we have to bear in mind, however, the opportunity that the nature of the materials affords the subject to develop a task-specific strategy.

Half et al (1976) presented their subjects with 19 sentences containing the adjective RED. Subjects were required to judge the relative redness of each of the 19 "red" objects denoted in the sentences. They find that when the rankings of the upper and lower bounds of redness in each context were established, subjects had rarely judged either pairs with overlapping intervals to be different, or pairs without overlap to be "equally red". This suggests that the different instantiations of red can be represented as ordered intervals. They also found that in around 20% of cases, the interval associated with a context **included** the interval for another context, and in a handful of cases this other context also contained the interval for a further context. This finding makes it unlikely that the interval is the same width in all contexts, and rules out the possibility that in context RED could be represented as a point. This is convincing evidence the context influences the range of plausible or likely values a concept may take, delimiting an interval within the potential range of the concept in isolation.

Anderson and McGraw 1973 argue that general terms are instantiated as specific exemplars. In their first experiment, they explicitly asked subjects to form and report "a vivid concrete picture", but in the second no biasing instructions were given. In this second experiment, subjects' recall of previously presented sentences was cued by low frequency exemplars of the category whose perceptual similarity to the high frequency associate presumed to be instantiating the category was an independent variable. They found that the more similar low frequency associate was a significantly better cue. For example, WOLF enhanced recall of

(5) The animal [dog] ran under the bush.

more than SQUIRREL. Anderson et al. 1976 extend this work, arguing that instantiation allows selection from an indefinitely large family of potential meanings "a sense which permits a coherent overall interpretation of the message" (p667). They present evidence that particular terms serve as better cues than the original more general terms for sentences like:

(6) The woman [actress] was outstanding in the theatre.

Crucially, this advantage of particular terms did not obtain for control sentences containing similar lexical material arranged in different relations, thus ruling out a possible explanation for the effect in terms of associations between the cue and the lexical material. The authors are careful to note, however, that in free recall particular terms were substituted for general ones only about 11% of the time, implying some representation of the general term. From a formal point of view, too, we should be hesitant about representing a general term as a specific example, since for well-known Berkleyan reasons the latter may not provide an appropriate set of inferences.

This argument is analogous to the argument against some propositional accounts of text representation: that is, they fail to account, for example, for Garnham's finding that "fried" is a better memory probe for the phrase "she cooked the chips" than "cooked" itself. What it is suggesting is that when the "conglomerate structure" or "mental model" is put together it is not assembled like a lego model in which the constituents do not affect one another's structure. I am emphasising, then, what Gombrich 1960 terms "the beholder's share"; the linguistic signal under-specifies the intended interpretation, and the hearer must work it out.

2.2.2. Undifferentiated and non-egalitarian reference

The concept of reference is central to most if not all theories of meaning, and a view on reference will be elaborated in the following section. For now I will merely assume that the reference of an expression is that which it picks out. Since I won't be presenting an extensional account of reference, this account cannot be faulted on the non-existence of unicorns, or water on the moon, for example. One difference between this and an extensional account of reference is that here the speaker is unaware of what is being picked out and of exactly how it is being picked out.

The discussion up to now has emphasised that, at the level of the word, linguistic symbols (necessarily) underdetermine the referent. But, at the level of the sentence, something more precise is articulated. That is, complete utterances make something (a description, an intention, a desire, whatever) clear. It is, of course, the case, however, that the meaning of a sentence is not always conveyed with uniform precision, and in this section I will argue that the concept of undifferentiated reference can help us understand some well-known phenomena of this type. What I shall seek to argue is that because sentences as psychological objects exist only relative to a process of comprehension⁸, and because this process exists only relative to a motivated hearer, the ploy of idealizing away from these processes and motivations risks distorting our view of how best to represent lexical meaning. Specifically, these factors lead to degrees of differentiation of reference, and this phenomenon in turn has deep implications for the way we choose to represent meaning.

First of all, consider the sentence:

- (7) How many animals of each kind did Moses take on the ark?

Erickson and Mattson 1981 conduct three experiments in which they establish that many people fail to notice the error in this sentence even if

- i) time pressure was removed;
- ii) they were shown an example containing an analogous error;
- iii) they read the sentence out loud;
- iv) the inconsistent name had a different phonological structure;
- v) the inconsistent name was placed in a focal position.

The authors describe this as an instance where people don't accurately comprehend the input even "when the understander possesses all the knowledge necessary for correct understanding" (p541). They do find, however, that the inconsistent name must be semantically similar to the correct one. For example, Noah and Moses are both Old Testament grandees. What I would claim is happening is that preliminary analysis indicates reference to an event; the hearer's goal is to retrieve a property of that event (the number of animals of each kind) and so the reference is not further differentiated because further differentiation is not relevant. If this is correct, then we are forced to distinguish between consequential "reference to" and instrumental "reference with". That is, the processing of the meaning of

⁸ Clearly sentences exist relative to processes of production as well.

the constituents of the sentence is geared to the task in hand, and we can exclude as the "lexical entry" for *Moses* the sum of our knowledge of him in toto. The processes of comprehension, then, *use* words as clues to or filters of information that may be germane.

Let us briefly consider what one might call "the thesis of egalitarian and fully differentiated reference" to which the account I am describing stands in contrast. Ray Jackendoff's thought-provoking book (Jackendoff 1983) develops the idea that one can use language to refer to more than just "things" (see also Jackendoff 1978). In traditional set theoretic semantics, nounphrases are taken as logical constants (constants unless intensional logic is employed, in which case interpretation is relative to a model) which refer to entities in the world. But the "meaning" of a sentence is a truth value (or set of truth conditions) and the semantics of all other constituents is built up recursively from just these types.⁹ Jackendoff argues the necessity of enriching the range of types of logical constant so that, for example, a prepositional phrase can refer to a #path# a #direction# or a #place#. His reasoning is that for pragmatic anaphora¹⁰ to succeed, the listener must be able to "project" (see Jackendoff (1983) pp22-29 for an elucidation of this concept) an entity corresponding to the pronoun. For example, to comprehend the utterance "Johnny asked me to bring that when I go there tomorrow" the hearer must be able to project some #thing# which *that* refers to and some #place# which the pronoun *there* is used to refer to. He argues that syntactically differentiable mechanisms are available for pragmatically referring to, for quantifying over, and for expressing a relation of identity with #places#, #directions#, #actions#, #events#, #manners# and #amounts#. This he takes as prima facie evidence for the existence for corresponding ontological categories. This view is lent some support by Jane Hill's discovery (personal communication), that in her computer model of the acquisition of early language, (Hill 1983, Hill 1986), the appearance of a distinction between *where?* and *what?* questions in the adult input forced the formation of two distinct categories of entity.

⁹ It is, of course, the case that non-standard logics can employ different ontologies. Indeed, we will have cause to examine ones which use both different notions of entity and of truth in the course of this investigation.

¹⁰ Pragmatic anaphora involves using a pronoun to refer without furnishing an antecedent. The hearer must recover the entity from his knowledge of the context. For example, "will you pass me that?" has a pragmatic pronoun as its final constituent.

On Jackendoff's account, therefore:

"a simple sentence like... Floyd broke a glass violently... includes five referring expressions, picking out two #things#, an #event#, an #action#, and a #manner#." (Jackendoff 1983, p58).

Referential properties are associated with every major phrasal constituent (defined by an X-bar theory of phrase structure), unless any are explicitly marked to the contrary by the surface form of a sentence (eg by negation: "there is no unicorn in the concourse"). He argues that it is not legitimate to reduce these categories to #thing# because there is no more straightforward correlation between #things# and things than #events# and events: projecting a #thing# involves "complex mental computations" (p47). Thus the logicians' (and consequently many linguists') assumption that a semantics founded on the notion of individual constant ("#thing#") has a solid and unproblematic base is false. It's false because so much goes into "a psychological treatment of what standard logic treats as the simplest possible expression" and because "the idealization to a set of necessary and sufficient conditions is inadequate." (p47).

It is entirely reasonable to hold that there are different sorts of entity that can be referred to, and that proposition will not be disputed (see also Langacker 1982b, Miller and Johnson-Laird 1976). But although Jackendoff makes a strong case that it is not necessarily correct to assume that the "thing" is ontologically *a priori*, it may nevertheless be that it, or some other category, is in some sense ontologically primary. For example, Keil 1979 claims that the ontological categories develop out of one another by a process of differentiation during childhood, the development of events from objects being one instance he cites. That is, although #things/objects# are not trivially correlated with reality, their correlation may be simpler than that of events. Another alternative would be to argue instead that the concept of "scenario" is primary, and that early language involves relatively undifferentiated reference to entities participating in a scenario. It seems, then, that while Jackendoff is right to draw attention to and to question the reductionist assumption of standard logic, he is insufficiently hesitant about substituting an assumption of his own: we can identify positions between Jackendoff and logic which suggests that his position is potentially open to empirical question.

The other observation that we must make of Jackendoff's treatment is that - because it explicitly omits processing considerations - it does not confront the issue of understanding being relative to processing. Nor does it consider the effect of having a specific goal of

understanding upon the representation evolved. And, as we will see, both these factors lead us to the views first that the meaning of a sentence does not necessarily consist of referents for each major phrasal constituent and second that not all referents should be represented at the same level.

I have suggested that the fundamental entity in language might be the scenario: can this claim be justified with respect to language acquisition? It has been argued that children's single-word utterances denote complex informational structures: for example, *Shoe!* might mean "mummy is putting my shoe on" when uttered in an appropriate context. The analysis of holophrases is one aspect of the debate in the literature over the methodological strategy of "rich interpretation", and few investigators would claim that single-word utterances are as syntactically rich as Bloom 1970 for example, once suggested when proposing reduction transformations for two-word utterances. But many have held the view, nevertheless, that the child is trying to refer to some participants in a process and to convey a set of relationships among them (eg Greenfield and Smith 1976). In a holophrase, then, a single symbol refers to several entities.

More recent studies, however, have been taken by their authors' as suggesting a different interpretation of the earliest utterances of children. Bates et al. 1983 describe early meanings as being components of language games. For example, the child Carlotta (12 m) would say *bam* always at a certain moment in a tower demolishing game. These authors argue that at this stage the child doesn't **have** her words, she **does** them. They describe development as a decontextualization process in which a word begins to stand for the rest of the script.

Nelson 1983 argues that scripts are conceptually basic, and characterises language development as the "cannibalization" of these scripts to provide parts that can be employed in new situations. (She hypothesises that this can begin when a single object plays several parts within a script). The script is a first layer of representation; analysis produces further layers, but the script is preserved. As evidence that the child's representation is in terms of events, she cites Church (1966) who describes how a 12mth old responded to the word *bath* by taking off her clothes, turning on the taps etc. Nelson cites also the Hallidayan "pure performatives"; under this analysis one talks of a "naming script" which is a part of a "reading script", for example. Vygotsky 1962 argues, rather earlier, for a strikingly similar position:

"Semantically, the child starts from the whole, from a meaningful complex, and only later begins to master the separate semantic units, the meanings of words, and

to divide his formerly undifferentiated thought into those units." (p126).

We might paraphrase this as: language development is a process of differentiation of reference. Maximally undifferentiated reference, by contrast, arises when a syntactically complex utterance is taken as referring to only one entity.

I wish to claim, however, that the possibility of an interpretation of a sentence not differentiating referents remains throughout childhood and into adult life, as in the case of the semantic illusion described above. In this I take a line parallel to that of Donaldson 1978:25 on "egocentrism" (cf Piaget 1926) and disembedded thought in general (Donaldson, Grieve and Pratt 1983).

"what is being claimed is that we are all egocentric through the whole of our lives in some situations and very well able to decentre in others. Piaget would not disagree with the claim that egocentrism is never wholly overcome.... [But] I want to argue that the difference between child and adult in this respect is less than he supposes...."

The first phenomenon illustrating this is the use by young children of "event probability" as a "preference rule" (cf Jackendoff 1983) in reaching an interpretation of active and passive sentences. Strohner and Nelson 1974, for example, find that between about 2 and 4 years of age, children make errors in the interpretation of active sentences which denote unlikely events.

(8) The mouse chased the cat.

is acted out as a moggy pursuing an mouse. Such errors must, argue Strohner and Nelson, be due to the employment of a "most likely scenario" strategy because neither a "competent" (cf Chomsky 1965:8ff) analysis nor an analysis based on an "N-V-N = agent-process-patient" word order strategy would yield errors. My claim is that the "probable event" strategy involves undifferentiated reference to a scenario (in this case, a "predation" scenario) rather than projection of the two entities in a given relationship, with projection based on the major syntactic phrases in the input string. Instead, the main verb summons a scenario (or script, schema, frame, whatever; cf Minsky 1975 and, Verschueren 1981) which fills slots (argument places) as appropriately as the remaining material (in the two NP's) allows. That is, the three conceptual constituents have no independent cognitive existence; rather they are emeshed in a set of relationships and the child takes the sentence as referring to a single, minimally differentiated, event.

Support for this view comes from the additional finding in Strohner and Nelson 1974 that 3 year old children will employ the word order strategy on what they term "reversible" active and passive sentences.

(9) The Spurs player kicked the Arsenal player.

This is a reversible sentence: the event is equally plausible whoever is assigned the role of agent. Thus, the word order strategy seems to be used at first as a fall-back position iff the probable event strategy fails to yield a clear decision. This could be modelled in structural terms by assigning greater weight to the former heuristic (cf Jackendoff, 1983 ch8) but let us, from a processing perspective, assume that the former is a "quick and dirty" analysis of the string, which passes "control" to a word order strategy only if it fails. If this assumption is correct, then we can see why the child's interpretation is relatively undifferentiated: it is based on only a preliminary analysis of the input.¹¹ What was Carlotta's only mechanism for dealing with words is developing into a processing heuristic for Strohner and Nelson's slightly older subjects.

I have argued, then, that aspects of children's performance on tasks involving the interpretation of sentences can be understood in terms of a phenomenon of undifferentiated reference arising out of a processing strategy. If this is correct, we can justifiably expect to encounter a similar phenomenon among adults, and I contend that this is so. Indeed, it pervades ordinary language use to an extent rarely revealed by formal analyses of sentences.

For example, Broadbent 1981 has the following to say about improbable sentences:

"... sentences are systematically understood in ways that sober application of the official structure of the language would stigmatize as incorrect ... in fact, hasty adult readers may have to re-read the original sentence to see exactly what was wrong." (p376)

He refers to a study by Herriot (Herriot 1969) showing increased reading times for improbable sentences for adults, which implies that the strategy has not totally vanished.

¹¹ It must be noted, of course, that the account here doesn't explain why children develop beyond this stage of accepting decisive preliminary analyses of strings (eg. they are eventually able to interpret improbable sentences correctly). Although Strohner and Nelson 1974 provide some intriguing evidence with a bearing on this issue, it is beyond the scope of the present work to pursue this.

This view of language understanding is a common-place among students of text comprehension (eg. Bransford and McCarrell 1977, Garnham 1981), who would generally argue that the representation of a text is a representation of an event (the event described in the text) rather than a text. That is, not a list of representations of successive sentences. The sort of evidence that is adduced in support of such claims is that, for example, memory contains more than the text provides at face value, leading to confusion among paraphrases of material that does appear with information that is only implicit.

There is also evidence (Pichert and Anderson 1977) that people remember a passage differently according to the perspective they are asked to take, which supports the notion that the motivation for processing a sentence can affect the way it is processed. Similarly, Barsalou 1981 has demonstrated that people form concepts according to their goals. An example of such a concept would be "food suitable for dieting". (Lakoff 1984) has demonstrated how languages institutionalise culturally important categories which can seem heterogenous to the cultural outsider.

It is standardly argued that referring expressions do not necessarily generate new referents in the model but will frequently be linked up with individuals already in the model (the central exemplar of this type of referring expression being, of course, the pronoun). This is modelled by Johnson-Laird 1983 in terms of a superficial propositional representation for each sentence of the text being operated on by a set of rules for creating a model (which, as has been noted in the literature (Garnham 1981, Reichgelt), is very like the creation of discourse representations in DRS theory). Such rules would stipulate, for example, when it was appropriate to introduce a discourse referent to the model. Indeed, several studies have been able to demonstrate the consequences for processing time of increasing the difficulty of bridging inferences required to make these connections (Clark and Haviland 1977, Sanford and Garrod 1981).

Here it is important, then, to take into account the purposes of the hearer in listening to a piece of language, and to recognise a distinction between "referring to" and "referring with". The former is the prototypical "reference" of linguistic and philosophical work. The italicized NP in:

(10) *The man in the corner* called out to us.

"refers to" a certain individual. But it is not always useful to project an entity

corresponding to every NP in a text: it may be, for example, that the reader's purpose is to learn the attitude of the writer towards some topic. Consider the following short passage:

General Sibson came into the room and poured a large brandy. He said: "those goats will get what they deserve." A wry smile crossed Peter's lips: "and you, will you get what you deserve when this war is over?"

In such a case the text is read as "referring to" the attitude in question "with" intermediate reference to concepts such as "goat" and "war", but neither the topic (the behaviour of officers and men in a war) nor the attitude is explicitly referred to. This is not to say that the referring expressions in such cases go unprocessed. To be sure, in order for the text to be interpreted, they must be processed. What is being claimed is that they are processed less fully (cf Bradshaw 1974, Lackner and Garrett 1973).

Conrad and Rips 1986 report a study of sentence-picture verification in which they systematically varied the pieces of information that were "given" by systematically manipulating a prior context-setting interrogative. On the assumption that those aspects of content that are extracted from the target sentence first will be compared with the picture first, they hypothesise that subjects will be faster to reject pictures that are discrepant with respect to the earliest information to be extracted. They found that disconfirmations were fastest when the false components included the given term. In their discussion, they suggest that given information receives prior but more superficial attention. Similarly, Engelkamp and Zimmer 1983 found that subjects shown pictures depicting a scene described by a preceding sentence were likely to prefer looking at the picture of the entity that had been placed in focus by clefting. They hypothesize that focussing devices result in "deeper" processing, in the sense of processing that brings about greater elaboration.

There is a close parallel with the cognitive asymmetry of the subject and object of a sentence. Evidence for this is reviewed in Engelkamp and Zimmer 1983. Anderson (1976) incorporates this observation in his ACT framework (by notating the 'subject of a proposition' differently from an 'argument of a relation'), arguing that the object of a sentence is not simply referring and rejecting what he calls flat propositional theories in which all nounphrases are treated as arguments to the verb. Strawson 1967 regards this asymmetry as "obvious" (p106), and it is also incorporated by RW Langacker who uses a reversal of foreground and background to account for the difference between active and passive sentences (Langacker 1982b). The psychological underpinning of passivization is discussed in Allerton 1982 (pp48ff).

Engelkamp and Zimmer 1983 also examine what happens when passivization and clefting suggest different foci ie. when they conflict as in sentences like:

(11) It is the motorcyclist by whom the girl is welcomed.

They found that gaze switching between the entities (as depicted in a drawing) was greater for sentences harbouring such conflict.

Langacker (1983) gives a more highly elaborated view of the theory developed in Langacker (1982), and in that more recent paper Langacker develops an argument that constructions related in transformational theories by the "dative shift" rule (and which, under the Katz-Postal hypothesis that transformational rules don't change meaning, would have to have the same meaning) differ in terms of the conventional imagery (or metaphor or construal) that they provide as a vehicle for the content words. Thus,

(12) John gave the box to Dave.

emphasises the "path" (traversed by the box), whereas

(13) John gave Dave the box.

emphasises the resultant state, in which Dave is in control of his new possession. Hence, the choice of construction is a matter of the "relative salience" (Langacker 1983 p49) of the entities referred to. He points out that this difference has consequences for the acceptability of instantiations of the construction: because it's difficult to construe a fence as a controlling possessor

(14) The shortstop threw the fence a ball.

is slightly odd. And because it is hard to imagine leaving a path that is not salient:

(15) I threw her the ball, but it never reached her.

and

(16) I went down the office, but never got there

are also anomalous.

Allerton (1982) similarly observes that valency relations and surface relations are independent, for purposes related to the given-new distinction and textual coherence. Of particular note are the relations between lexical items such as *elude* and *miss*, which will participate in truth-conditionally equivalent propositions, but arrange their arguments differently on the surface. A relevant finding in the psychological literature is Hampton's recent discovery (Hampton 1987) of the non-commutativity of typicality ratings for exemplars of complex concepts. Specifically, he found that subjects do not give the same rating for a given exemplar for putatively conjunctive categories of the following form:

(17) An X that is a Y.

(18) A Y that is an X.

Rather, greater weight was systematically given to the term appearing in the relative clause. Hampton suggests that this may be due to the primacy of new information.

Langacker develops a formal framework (Space Grammar) to permit representation of such distinctions, and some aspects of this system will be described in a later chapter. The relevant point here is that the difference between constructions (such as active v. passive, or unshifted v. shifted ditransitives) can be understood partly as a difference in the relative prominence given to the entities participating in the process being described - the same event can be portrayed differently.

In conclusion, I claim that making a distinction between differentiated and undifferentiated reference as the consequence of an interpretative act (and its processing corollary, the "to"- "with" distinction) can give us a theoretical handle on some superficially unconnected phenomena. This distinction seems likely to be related to a processing distinction and to arise out of the fact that language is **used**.

Many other investigators have proposed bi-partite models of processing along broadly these lines. Gibson 1950:211 says:

"the perception of everyday life is very often schematic ... but perception can become literal whenever the observer needs to discriminate."

Neisser 1967 makes an analogous distinction between two stages of processing, and Wason 1977 contends that the difficulties some subjects experience in coming to terms with the

solution of the 4-card problem arise because of the nature of the pre-analytic framework in which they first cast it.

It is likely to be something that intellectual development affects, with increasing maturity being reflected in increasing ability to differentiate reference. It is possibly something that is affected by cognitive style; one might predict, for example, that field independent subjects would be more prone to differentiating reference than field dependents. It is also, probably, affected by linguistic variables. For example, one might expect undifferentiated referents to receive less elaboration of meaning, and this could be used to test whether backgrounded information is relatively undifferentiated.

2.2.3. Referential perspectives

It has already been noted that different constructions can be used to describe a single situation. In this section I wish to build on this notion, and to discuss something of the relation between linguistic meaning and the world.

Strawson (1967) argues against the view that the meaning of an expression is a referent, or that the meaning of a sentence is a truth value. These are, instead, characteristics of some use of an expression. Meaning, he suggests, consists of general directions for the use of an expression; in other words, conventions governing use. Thus Strawson puts forward a view of meaning broadly compatible with the view being developed here; on a particular occasion of use, an expression will have a referent, and this referent is a more specific sort of thing than the meaning of the expression itself. Johnson-Laird 1987 puts forward a perspective on meaning that appears to be equivalent.

We should note that on occasion the world can pre-empt the process of inference and construction by making a referent perceptually available. This provides specific properties. Thus, for example, if I ask you to glance at "the first word on the page", many of the properties of that word are fixed (such as its identity, its form and so on). On the other hand, imagining the last word of Noam Chomsky's next book leaves the instantiation of the word WORD more open. We can define two extremes; at the one end the entity is entirely familiar; at the other it is entirely unknown. Obviously, definite articles and possessives are likely to anticipate instances of pre-emption. Strawson points out that some types of expression are more dependent on context than others; a pronoun does not pick out anything

without some context, for example. Nevertheless, given the degree of support from context that its type requires, any nounphrase can be used under conditions of extreme familiarity with the referent.

What, then, can an expression communicate if the referent is already known? Most obviously, it may form part of a sentence that predicates a new property of that entity, and in the extreme case the referential expression may function merely as a tag or pointer to the entity in question, having - in the sense in which J. S. Mill suggested that proper names were meaningless - no meaning. More typically, the referring expression will have some descriptive content and will contribute a fresh perspective on the entity in question. Thus we have the elegant variation of journalism, and the inelegant variation of ordinary conversation. As Strawson noted, the predication of a simple subject-predicate construction may be conveyed by a definite nounphrase. For example:

(19) Napoleon was *the greatest French soldier*.

Here the nounphrase is functioning at the opposite end of the spectrum from a tag. Attributive reference (Donnellan 1966, Johnson-laird and Garnham 1979) has essentially this character. The definite description in

(20) *The man who left his wig* (whoever he is) is guilty of robbery.

is not referential; rather, it attributes to a reference target (as yet unsighted) certain properties.

It can be argued that proper nouns are not like common nouns because one can ask about the reference or the meaning of common nouns, but only about the reference of proper nouns (Kempson 1977). Our notion of referential perspective and "degree of pre-emption" avoids this problem rather neatly. Because pre-emption is a matter of degree, there is no discontinuity. To make this argument persuasive, it should be pointed out that proper names can manipulate perspective in at least two senses. First, there is a difference between the expressions *Barry* and *Professor Richards* (which refer to the same individual and are proper nouns), which might, as a first approximation, be termed "honorific". Second, characteristic properties of Professor Richards can be attributed to Dr Myers by utterances like:

Terry is being a Barry.
having

doing
etc.

I have emphasised that referring expressions and sentences portray entities and situations, and that the speaker can choose to portray a given entity or situation in different ways by his choice of words or construction. In later chapters, I will look in detail at a particular example of the manipulation of perspective by referring expressions; the mass-count distinction. In general, entities can be referred to using either mass or count nounphrases in English. The difference is essentially one of perspective. For example, the phrase

(21) a vast expanse of table

highlights a particular aspect of the table in question.

2.3. Conclusions

The term "conclusions" is perhaps too conclusive to describe the speculations I put forward here. Clearly much remains to be done in working out a theory of psychological meaning which accounts for all the relations among symbol, world and mind. Nevertheless, I hope, in these first two chapters, to have sketched a useful survey of what I take to be relevant data, and to have drawn some appropriate conclusions regarding the process of interpretation. Interpretation is viewed here as a constructive process in which attention plays a critical role. Syntactic devices and lexical items provide constraints on the process of construction. We can describe two sorts of definition that might be given to a word. On the one hand, there is the meaning which users of language ascribe to a word through a metalinguistic process of dividing up the overall meanings of utterances among their constituents. This is lexicography. You start from the interpretation and apportion sections of it to the component phrases of the utterance. On the other hand, its meaning can be viewed as the contribution that a word makes to a perceived meaning or understanding of an utterance. That is, a word plays a role in the cognitive event of comprehension, and its meaning to the comprehender is what it adds to the message. On this view, which I take to be more compatible with the empirical findings I have outlined, the word does not so much provide material for the construction of a representation, but a tool. In particular, lexical concepts integrate with each other, and this process of integration extends over time and influences the elaboration of the meaning of the lexical concept.

Part 2.

Characterising Lexical Concepts.

Having outlined a framework in Part 1, I focus now on the key issues. In part, this is done by considering some important ideas which have been put forward as theories of word meaning. My approach to these theories is critical, but I aim to retain the valuable insights which one invariably finds associated with any major theory.

In chapter 4, I will discuss prototype theory and its limitations, but first I will tackle two issues which may otherwise distract us. Firstly, what relationships hold between meaning postulates, semantic components and, indeed, other formal devices such as semantic networks and predicates? And secondly, can a purely systemic account of lexical meaning suffice as a psychological theory?¹² These questions are closely entwined in the literature, but I will discuss them separately.

In the final chapter, I present an informal description of my own "Lexical Concepts" view. The rest of the thesis will explore this approach.

¹² The second part of this chapter will appear as Dunbar 1989.

Chapter 3.

The importance of notation and system.

In this chapter I briefly examine two issues which together seem to have occupied twentieth century lexical semanticists heavily. The first arises immediately in any practical attempt to do lexical semantics, and is concerned with the problem of representing meaning orthographically. The second concerns a theoretical concept that has been highly influential, the notion of systemic contrast. I make no secret of my view that neither of these questions provides a research path for profitable cognitive science.

3.1. Notation.

The principal differences among notations concern the interpretation given to elements and the processing assumptions associated with the notation. Semantic components, for example, are generally interpreted as abstract concepts and are, in particular, held to be distinct from elements in the object language. Meaning postulates, on the other hand, are relations of entailment among words of the language under study. Thus:

(22) *man*: [+HUMAN] [+MALE] [+ADULT]

and

(23) *man* → *male*

are, respectively, componential and meaning postulate analyses of the English word *man*. It is sometimes argued that a meaning postulate analysis has the advantages of permitting partial definitions and of avoiding the need to assume that words are interpreted by decomposition into semantic primitives (Fodor et al. 1980, Johnson-Laird 1983, Lyons 1981). (For further evidence that "componentially" more complex items carry no heavier processing burden, see Kintsch 1974). Others take the view that these assumptions are independent of the notation used (Bierwisch 1970, Leech 1981). Leech (1981) argues, rather, that describing semantic components as abstract concepts has the advantage that the same

components can be used in the analysis of different languages, whereas separate postulates would have to be provided for French *chien* English *dog* and German *Hund*. Lyons would argue strongly, however, that it is not for linguistics to thirl itself to a particular set of assumptions about the nature of mind (ie. that there are concepts or that these concepts may be universal) (Lyons 1968, Lyons 1977, Lyons 1981). (Leech 1981 specifically avows a weak universalist position). Kempson 1977:ch11 is sceptical about any claim that the difference between components and postulates is more than superficial.

Allan 1986, preferring the term predicate to "meaning postulate"¹³, concludes that predicates are equivalent to components and goes on to interrelate entailment and set membership (Allan, 1986 pp172-173). Indeed, a typical componential scheme will include some redundancy rules like [+BOVINE] → [+ANIMATE] fleshing out basic lexical entries; such rules draw the two formalisms still closer.

It has been asserted that a componential analysis is limited (relative, presumably, to both network and meaning postulate accounts) in not being able to account for relational attributes (Cohen 1983), but Leech 1981 shows how to encode both polar and relative oppositions in terms of components.

Comparisons have also been made in the psychological literature between semantic network theories (eg Anderson and Bower 1973, Anderson 1976, Collins and Quillian 1969, Rumelhart and Norman) on the one hand, and "set-theoretic" (componential) ones on the other (eg. Smith, Shoben and Rips 1974). Semantic networks represent words as embedded in hierarchies; the meaning of a word is assumed to be the sum of the properties it inherits by way of connection to other words. In more sophisticated versions, links other than the ubiquitous *is-a* are permitted between nodes. Hollan 1975 argues that the difference between nets and sets is merely notational and points out that a network can be defined as a set of linked nodes which is equivalent to a set of (concept, feature) pairs.¹⁴ Rips, Shoben & Smith's (1975) reply does not contest this structural isomorphism, arguing only that the distinction matters with respect to processing considerations (what operations consume resources and whether there need be complete decomposition).

¹³ His reason for preferring the term predicate is quite simply that he wishes the elements of his analysis to be part of a metalanguage rather than the object language.

These authors conclude that:

"Apart from these substantive matters, there seems to be little reason to prefer networks to sets, except as a matter of theoretical convenience" (p157).

Anderson (1976) argues that the principal advantage of nets is that they represent a convenient notation for analyzing (memory) search, and in chapter 5 he proposes a specific procedure for translating network representations into linear ones as required.

Johnson-laird (1983) suggests that:

"In the absence of any principled constraints on the processes that can be employed in setting up or interrogating semantic networks, the empirical content of the general theory of semantic networks is unclear" (p214).

He does, however, commit himself to the view that meaning postulates are distinct from components, claiming that only the latter can account for certain experimental data. Johnson-Laird's supposed refutation of meaning postulates rests on the mistaken assumption that the notational device carries theoretical weight (Johnson-Laird 1987). To carry it through, he is forced to prescribe that "[subjects] *should* search for postulates of [a specific] form" (p193, my italics).

A meaning postulate account can work, however, if it is not burdened with the additional assumptions prescribed, but rather is granted some others.¹⁵ Specifically, let us assume that the subject engages in parallel in serial searches of two stacks of meaning postulates:

Stack 1.

$\forall x (x \in \{\text{jug, knife, ...}\}) \rightarrow x \text{ is not a substance}$

$\forall x (x \in \{\text{sherry, ammonia, ...}\}) \rightarrow x \text{ is a substance}$

$\forall x (x \in \{\text{toffee, wood, ...}\}) \rightarrow x \text{ is a solid}$

¹⁴ In fact, Hollan points out that a function assigning a value to each edge of the digraph is also necessary for equivalence to the feature theory defined by Smith, Shoben and Rips 1974. This is because Smith et al's Feature Comparison Model takes variation in criteriality into account (see below).

¹⁵ This model was first presented to a Mental Lexicon workshop in the Centre for Cognitive Science, Edinburgh, on 16th July, 1987.

Stack 2.

$\forall x (x \in \{\text{vase, syringe, ...}\}) \rightarrow x$ is not connected with consumption

$\forall x (x \in \{\text{jug, toffee, ...}\}) \rightarrow x$ is connected with consumption

$\forall x (x \in \{\text{toffee, plate, ...}\}) \rightarrow x$ is connected with eating

$\forall x (x \in \{\text{toffee, cake, ...}\}) \rightarrow x$ is consummable

Searching of each stack terminates when the key postulate is arrived at. The fundamental assumption that makes this model work is that search proceeds from less specific to more specific meaning postulates. One further assumption: for each rule an item passes (apart from negative rules), add 0.1 to the probability of recall. For example, *toffee* passes two rules in Stack 1, and three in Stack 2; while *plate* passes no rules in Stack 1, but two in Stack 2. The predicted probabilities of recall are tabulated along with the percentage of words correctly recalled, as reported by Johnson-Laird, Gibbs and de Mowbray 1978, in the following table:

Table 1. Comparison of Predicted and Obtained Recall

Word	Predicted probability	Obtained percentage recall
vase	0	8.1
jug	0.1	10.6
plate	0.2	16.2
petrol	0.1	10.6
beer	0.2	21.5
coal	0.2	21.5
toffee	0.5	50

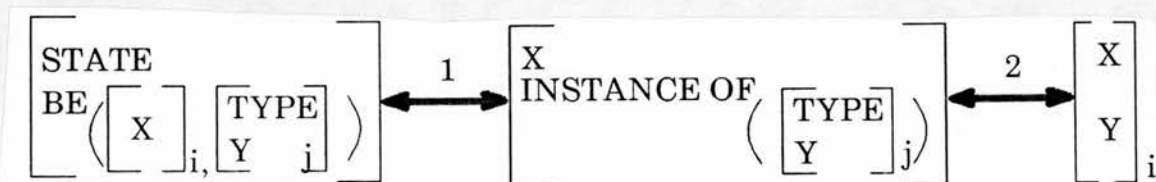
This model gives a fairly good fit to the recall data, accounting for a little over 95% of the variance of the means reported ($r = 0.9766$). The object is not, however, to propose a serious model of behaviour, but simply to demonstrate the general point that meaning postulates could handle such data.

These arguments serve to confirm our suspicion that the theoretical context is of greater significance than the notation. Collins and Loftus 1975 clinch it when they say that:

"Network models were developed as a method of representing features in a computer" (p410).

Jackendoff approaches representation from a different angle, arguing that in principle semantic relations are to be derived by processes of conceptual induction (Jackendoff 1983:ch6). He argues for his approach from two directions. Firstly, generic categorization judgements can behave like ordinary categorization, in that, for example, there may be fuzzy margins. And secondly, an initial *a priori* deductive relationship between some categories (required to get the ball rolling), cannot simply be assumed. (This second argument will also be used against systemic accounts of meaning in the discussion below). It is, therefore, necessary to allow what he calls "TYPE-inclusion" judgements to be made on the same basis as judgements that particular tokens are instances of certain types. Sense relations, such as entailment and hyponymy can then (following Katz), largely be derived from generic categorization judgements.

Nevertheless, he eventually reaches a view similar to ours with regard to notation. In addition to categorization judgements, he provides inference rules which he describes as handling **deductive** property inheritance (that is, the same thing as meaning postulates and the rest). The two rules interrelating three types of representation are shown below (where "X" is some type or token):



Syntactically, 2 is a sort of "property raising" transformation; any entity that is an instance of some category can have that category as a property.

Because he emphasises a Grammatical Constraint (ibid., pp13-16), Jackendoff rejects some notations because they do not map easily onto the form of English. Jackendoff's theory differs, then, in according importance to the form of the notation as well as in his invocation of the notion "judgement", of which more later. He argues, however, that network/meaning postulate theories are notational variants of markers (ibid., p123). In closely relating attribution, inference and sense relations, he adopts a position in line with the consensus of previous research outlined above.

In short, then, the content of any device for representing the meaning of words stems largely from associated assumptions, rather than the form of the device. Here I will talk in terms of parameters and their values, but let me emphasise that this does not amount to a hypothesis about lexical meaning. In order to avoid confusion, the term "parameter" is chosen because "feature" is sometimes used in the psychological literature to mean "feature + value". For example, Miller & Johnson-Laird (1976) state:

"We take "feature" to mean any particular value of an attribute that a percept must have in order to satisfy a test based on the perceptual predicates." (p137)

Certain issues have emerged as substantive in the preceding discussion, and I must briefly sketch a stance with respect to each. Regarding completeness of definition, I do not claim that the definition given of any lexical item is other than partial, both in the sense that the particular analysis is incomplete and also in the sense that lexical concepts are, as argued in chapter 1, subject to processes of instantiation. A specific position on the related decomposition issue will be developed (along the lines of Langacker 1982b, Langacker 1983 and, Leech 1981) in due course. See Miller & Johnson-Laird (1976) pp326-333 for an interesting discussion of decomposition and completeness of definition. On the issue of universality, I will have little to say, but any claims will be made explicitly at the appropriate point. Perhaps the most important associated assumption (at least from the philosophical point of view) and the most rarely explored, concerns the source of an interpretation for the notational atoms, and this brings us to the second lemma which I wish to argue.

3.2. Word meanings and distinctive features.

The classical answer to the question "what is word meaning?" is generally ascribed to Aristotle: in learning a concept one disregards particulars and extracts what is common to the class.¹⁶ In other words, a mature concept is a set of common features, shared by the entities in what has come to be known as the word's "extension". This solution has, until comparatively recently, dominated psychological and linguistic (not to mention lexicographic and anthropological) theorizing on the topic.

¹⁶ This section has benefitted from detailed comments from Karen Lyon and from discussions with participants in the Mental Lexicon workshop, to whom I am grateful. They should not, however, be held responsible for flaws; for these, I must be held to account. A paper based on this section was not read at the XXIII Linguistisches Kolloquium in Berlin, August 1988.

For example, Wundt 1894 modelled concept learning as a process with two phases: first generalization, and then abstraction of features of meaning. This paradigm was the focus of most research into concept formation in the behaviourist learning theory school up until the 1960's, when pioneering investigators such as Zaslav 1961 first began to argue that the concepts of natural language did not fit the classical mould. The classical concept acted as *telos* in Vygotsky 1962; children whose thought processes were still developing progressed through intermediate stages (eg complexes and chains) towards the achievement of these "scientific concepts". According to Clark 1973, children learn word meanings by inducing the common features and accreting them over a fairly long period to lexical entries.

In linguistics, this philosophy has emerged in the componential analysis of Katz and Fodor 1963 and its heirs (notably Bierwisch 1970, Leech 1969). A special case of the view is its subordination to the Saussurian notion of *systeme* (Saussure 1966). This, indeed, is the typical form of the classical view in linguistics: the features that represent the meaning of a word are just those which distinguish it from others in the relevant semantic domain (Leech, 1969), and from these "markers" (set up on the basis of distinctiveness) various semantic relations can be defined.

It was with the purpose of explicating the notion of "relevant" domain that the concept of "semantic field" was developed (Coseriu and Geckeler 1981, Miller and Johnson-Laird 1976, Trier 1931).

The same view has been held by others who emphasise that the meaning of a word derives from its relations with other words (Lyons 1963, Lyons 1981) and who therefore adopt the notational device of the "meaning postulate".

The issue, then, is the question of whether words are meaningful solely in relation to other words in the same language: a word means what it means because of what it does not mean. This view has been held by linguists such as Bierwisch and Lyons, and comes from the Saussurian tradition. Lyons, for example, puts it like this:

the sense of a lexical item [is]... identical with, the set of relations which hold between the item in question and other items in the same lexical system." (Lyons 1968)

Let me emphasise that the object here is not to criticise a particular linguistic theory or theorist. It would be too easy to take a quote out of context and represent that as the sum

of a theoretical edifice. Rather my approach is essentially interdisciplinary. I will look at the question of how successfully the linguistic abstraction can be transferred to the domain of cognitive structure. That is, to what extent can the concept of systemic contrast be successfully applied to a description of language as situated in the mind.

For many enterprises, this is not an important issue. For example, it is irrelevant to the lexicographer constructing a dictionary, for whom the systemic abstraction is probably a useful tool. However, it is timely to raise the issue not only because of its intrinsic interest but also because of the current rapid development of two areas of research which to a certain extent employ the abstraction as a background assumption.

1. Cognitive linguistics uses the notion of deriving meaning from location in a structured system. See, for example, Langacker 1982b.
2. In connectionist models, "meaning" derives from links in a very literal manner, and colleagues will sometimes talk informally as if this were the same thing as systemic contrast.¹⁷

Again, my intention is not to criticise the research strategy. Research programs may profit greatly from the implicit *rappports* with the work of de Saussure. Nevertheless, both these areas of research draw the lexicon into the cognitive domain, and ultimately they must confront the issues raised here.

There are four issues:

- Learnability
- Derivation of semantic relations
- Inherent interaction with context
- The existence of degrees of contrast

3.2.1. Learnability

¹⁷ It is of course simply wrong to make this equation; it is necessary to distinguish distributed links from local relations. Nonetheless, the metaphor is clearly influential.

A system must have at least two parts. If there is only one part, it has nothing to relate to, and it cannot be given an interpretation in terms of non-existent relations. What, then, of the very first word learned? How can the lexicon bootstrap itself?

One might propose that there are "hidden words" present in the "empty" lexicon of the infant language learner. Words such as EGO, OTHER, or some ontological concepts, such as "the object concept" are obvious candidates. Note, however, that now we are using extralingual relations, and we have broken faith with the strong definition of lexical meaning quoted above. This, of course, is not a problem for any real theory of language learning. The problem only arises if we try to force a certain linguistic abstraction into the cognitive domain.

Jackendoff 1983 makes a similar point (see above, cf however Fodor 1975). The problem is that it is difficult to see how the mesh of relations could be initiated.

3.2.2. Derivation of relations

An important advantage of the techniques of systemic contrast, such as componential analysis, is that they yield a representation of the system from which it is easy to derive lexical relations. The following table gives an illustrative analysis of four words using two markers:

	MAN	WOMAN	BOY	LAD
MALE	+	-	+	+
ADULT	+	+	-	-

An analysis like this lets you derive important lexical relations in a straightforward manner. For example, it is immediately clear that BOY and LAD are synonyms (in the terms of the analysis). You can also predict that a sentence like

(24) This man is a woman.

is contradictory (on one reading) because the two nouns have conflicting values for one feature.

However, it is possible to have systemic contrast without easy derivation of relations. This is demonstrated by studies which use multi-dimensional scaling (MDS). This technique requires subjects to rate a set of lexical items along a set of scales. For example:

LION —How fast is a lion (1-7)?
 —How round is a lion (1-7)?
 etc.

All the ratings for all the subjects for all the words are then boiled-down by a statistical technique which reduces the system to a set of locations in an n-dimensional space.

If mere distinctiveness is all that we require, then why not use the two-dimensional solutions to multi-dimensional scaling studies (eg. Fillenbaum and Rapaport 1971, Marx 1983, Smith, Shoben and Rips 1974)? If n-valued markers are permitted (Ladefoged 1971, Leech 1981), then it is in general possible to supply unique coordinates for each lexeme using just two dimensions: for example, Smith, Shoben and Rips 1974 classify a selection of animals along two dimensions which they label "size" and "predatoriness". Consider Figure 1. This is an example similar to their findings.

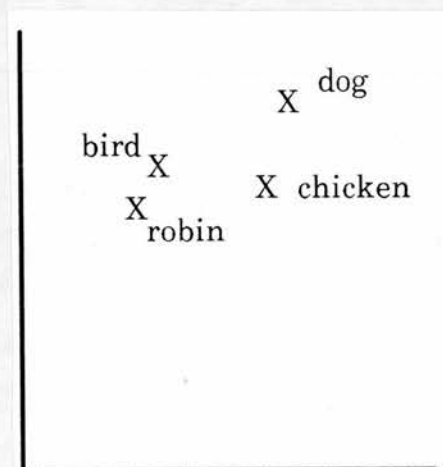


Figure 1.

We have systemic contrast, but where are the lexical relations? How, for example, could one derive the hyperonymy of BIRD to CHICKEN? Some configurations appear more promising. Perhaps, for example, we could interpret the configuration in Figure 2 as some kind of oppositeness.

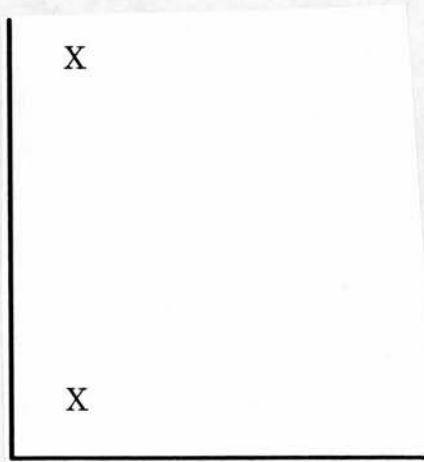


Figure 2.

Three difficulties come to mind. First, scaling the axes. How far apart must two items be before they count as opposites? Compare figures 3 and 4:

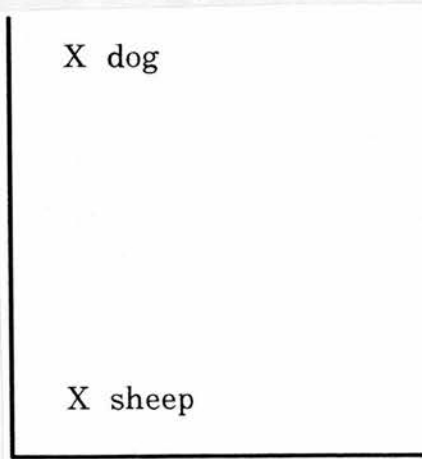


Figure 3

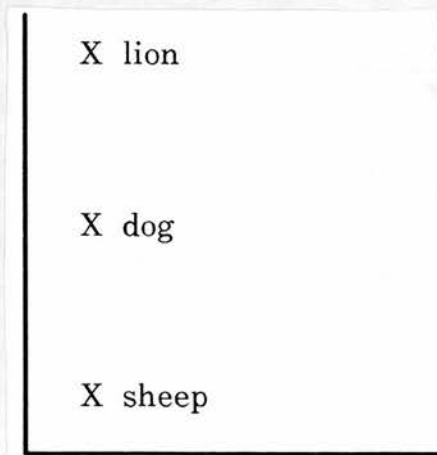


Figure 4

We would need to supply a similarity metric with which to measure semantic distance, and a threshold with which to evaluate semantic distances.

The second difficulty lies in the interpretation of axes. It should be noted that subjects did not give their ratings directly on these dimensions, but that these dimensions are a statistical summary of data from many subjects using several other dimensions. The axes can

usually be labelled roughly, but there is a serious difficulty in saying precisely what the nature of the oppositeness is.

And thirdly, a two-dimensional semantic space is too crude to articulate all the relations we might want. Specifically, it could not supply all the relations of oppositeness, for example, that we might need. There are other, perhaps more profound, difficulties with MDS, but my purpose is to demonstrate that it is not systemic contrast *per se* that supplies an analysis of lexical relations. That is, distinctiveness or contrast does not, when combined with rigorous application of a principle of economy, suffice to motivate features that can account for sense relations.

3.2.3. Interaction with context is inherent

It is well known that within different semantic fields, words take on different senses (Benveniste). I illustrate with DOG which has different co-hyponyms when it is a "pet" or a "guard".

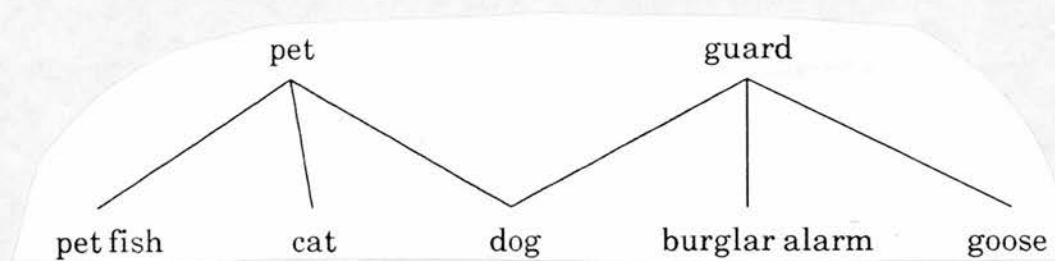


Figure 5.

It is obviously not good to have to postulate two lexical items DOG (one for each field) which would nevertheless be extensionally equivalent (though perhaps with a different typicality structure - see below) just to save the analysis. Assuming, instead, that these are two uses of a single lexical item, we are led to address the question of how to characterize the meaning of that lexical item in such a way as to allow it to interact with both semantic domains.

I will discuss context effects on word meaning in more detail in the next chapter, but I want to give a suggestion here of the dynamicity of this interaction.

Fillmore has remarked (Fillmore 1988) that EPISCOPALIANS can't be opposed to BOYS in the same way that GIRLS and MEN can. Imagine, however, the following scenario: a troop of boy scouts, washed out of their camp site by rain, and a coach load of churchgoers on a touring holiday, arrive at the Hotel Duomo, say, at the same moment. It would be possible for the manager to say something like:

(25) Put the episcopalians on the first floor and send the boys away.

Clearly, Fillmore has in mind the non-orthogonality of boys and episcopalians, but in our scenario the groups are functionally orthogonal, and so the opposition is plausible.

This dynamicity implies that the interaction is inherent in lexical expression. This notion of inherent interaction will be developed in later chapters.

There is an important point to be made here. When is an abstraction so abstract that it corrupts the description? That is, sometimes an abstraction abandons a factor that is crucial to understanding some phenomenon. If it does this, it is not helpful.

Gibson has discussed this in relation to visual perception. The paradigm in psychology was to study perception in terms of static "snapshots" of the visual field. Gibson and his followers have demonstrated that patterns of texture in the optic *flow* (that is, the continuously moving flow of reflectances on the retina) can specify invariances in the environment. In other words, motion turns out to be essential to understanding certain phenomena of visual perception.

Guillaume (Guillaume 1984) has discussed motion, too, but he discussed it in relation to word meaning and in relation to the work of de Saussure. Guillaume writes of the movement from potential to actual, and it is precisely this notion that I emphasise here. I will argue later that it is in the realisation of lexical concepts as concepts that contextual interaction inheres.

3.2.4. More and less contrast.

There is a marked subset of uses of words which are more contrastive. Some differences in the linguistic properties of these uses are described in the following paragraphs.

First, of course, these uses tend to be marked physically using stress in spoken language, and a bold typeface or underlining in written language. For example:

(26) Did you get Trixie a **dog** [and not a cat]?

Moreover, contrastive stress has syntactic implications. Consider, for example:

(27) That a man got the job pleased him.

Here the pronoun can be coreferential with the noun-phrase *a man* only if that noun-phrase is employed contrastively, the structure of the sentence ruling out coreference on any other reading. Marslen-Wilson, Tyler and Seidenberg 1978 report two similar cases of an interaction between contrastive stress and acceptability. They argue that a subordinate clause like:

(28) Although the drink is quite cheap **beer**,

is alright only if *beer* is stressed contrastively. For a non-contrastive reading, they suggest that it's much better to substitute a pronoun for *drink*, as in:

(29) Although it is quite cheap beer,

They also remark on the relative acceptability of subordinate clauses ending with an adjective which is/isn't given contrastive stress (getting it being better). For example:

(30) Although the new door is very expensive **steel**,

Their attention was drawn to this by "very long" (p236) times taken by their subjects to spot a target which followed the adjective when it appeared as required by their experimental procedure without contrastive stress. According to Zwicky and Sadock 1975, (31)

(31) Why not move to Chicago?

unambiguously has the illocutionary force of a suggestion unless there is contrastive stress on *not*, in which case the sentence is semantically ambiguous between a suggestion and a question (p25). This ambiguity is **semantic** because the crossed interpretation of (32)

(32) Why **not** move to Chicago and Maine?

is not possible. (The crossed interpretation is the one in which the speaker is questioning one move while suggesting the other; see chapter 7). Thus, distinctive meaning does exist, but it is distinct.

This evidence suggests that contrastive stress is not merely an emphatic device.

These phenomena belong to *parole* rather than *langue*. The question that must be asked, however, is how we make sense of the implication that other uses are less contrastive. I believe this goes to the heart of the matter: there is meaning other than the meaning that comes from the words not spoken. I explore the question no further at this point, because my purpose is merely to raise questions.

I have presented, then, arguments that a purely systemic theory of lexical meaning leaves questions unanswered about learning; does not necessarily explain sense relations and has difficulty with accounting for contextual variation in contrast sets. Furthermore, when meaning is distinctive this is typically linguistically marked usage. The burden of this discussion is that the lexicon can't interpret itself.

In sum, I have examined the linguistic abstraction of systemic contrast and found some puzzles that will have to be resolved before a cognitive transplant can succeed. It may be a useful abstraction for some tasks, but for cognitive modelling we need to introduce a dynamic to the Saussurian *etat de langue*. The lexicon must be designed with a need for interaction with context in mind. The view that word senses consist merely of relations with other words and that, hence, the vocabulary serves as its own interpretation is not a viable theory of lexical meaning.

Before leaving this topic, it's worth noting that contrastive usage can be selected by the situation rather than the explicit choice of one of the participants in a discourse. Thus, for example, the utterance:

(33) I'm sorry, I can't come tonight ... I'm visiting friends.

can be rather awkward when addressed to a relatively new acquaintance, since it can carry the implication that the addressee is not a friend of the speaker (ie. that *friends* is intended contrastively). It can be salvaged by depending *other*. The contrastive interpretation here appears to be triggered by its own social significance. That is, it seems to be that just because the distinction is crucial at a given stage of a relationship, the participants become sensitive to it. This point bears further investigation.

Chapter 4.

Lexical concepts as prototypes.

This chapter examines prototype theory as a theory of lexical concepts.¹⁸ Although prototype theory offers useful constraints on a characterization of concepts, it fails as a theory of lexical concepts because it shares with the classical view the assumption that lexical concepts are concepts. To preview, the argument will hinge on the dynamic, constructive and knowledge dependent nature of concepts on the one hand, and the requirement that lexical concepts should interface successfully both with other lexical concepts and with different and changing bodies of knowledge on the other. Essentially, GOE (Goodness Of Exemplar) structure is reasoned to be a property of the instantiations of lexical concepts as concepts. First, however, let's consider prototypes.

It is by now a cliché that, contrary to the classical prediction, concepts are "fuzzy". The principal observation is that category membership is not a uniform, all or none, binary matter. Firstly categories tend to have vague boundaries as a consequence of the properties of objects being samples from continua. For example, it can be unclear whether a given object is a cup or a vase (Labov 1972), and at these margins categorization decisions tend to be more difficult (Jackendoff 1983, Smith, Shoben and Rips 1974). Secondly, it appears that in practise **necessary** features of meaning can be difficult to isolate (Wittgenstein 1953), except in the case of technical vocabulary (see below). Thus, categories possess a family resemblance structure, with some members being more closely related (similar) than others. And thirdly, categories seem to be structured around some centre (a good or ideal exemplar) with members rated according to their typicality (centrality or goodness). According to Rosch and Mervis 1975, an item is central in a category to the extent that the total cue validity of its attributes is high (ie. the "prototype" is maximally like members of the category and minimally like non-members) and, hence, family resemblance leads to the formation of a prototype. Alternatively, for some categories like colour terms, the focal

¹⁸ A paper based on this chapter has been published as Dunbar and Myers 1988.

points are given (neurophysiologically) and universal, with vocabulary in such domains being assimilated during acquisition to these focii. The core meaning of a category consists of these good examples ""surrounded" by other category members of decreasing similarity" (Rosch 1973:112).

This internal structure of categories has been shown to affect performance on a variety of cognitive tasks. Subjects, for example, agree in their rating of exemplars for "goodness" of category membership, and these ratings correlate highly with attribute overlap (ie. highly rated exemplars tend to be attributed with more common properties of the category in question). Furthermore, people tend to agree as to the order in which to complete hedged predications: for example,

(34) Pink is a sort of red.

would be a more frequent completion than

(35) Red is a sort of pink.

This indicates a shared conception of landmark colours. The final finding I'll mention is an effect of category structure on the time subjects take to make categorization judgements. They tend to respond quickly when either the exemplar is **not** a member of the category and the two are highly dissimilar or when it is a member and the two are very alike, but to make rather slower responses otherwise. This can manifest itself in terms, for example, of faster positive verifications for typical than atypical exemplars.

An early attempt to model this phenomenon was made by Smith et al (1974) who employed the process of "Feature Comparison", and argued that categorization was a judgement of similarity between the properties of the exemplar and those of the concept. They used Lakoff's work on hedges to justify a distinction between what were termed "characteristic" and "defining" features. The features of any given concept would be ranked according to definingness, and this ranking would be utilized by a two stage comparison procedure. At the first stage, all features would be considered in a global, overall assessment of similarity of two concepts. If they had sufficiently similar (or dissimilar) dimensions and values on those dimensions to exceed a criterion, then the decision could be made straight away - this being likely in the case of affirmative decisions about "good" exemplars, since they will tend to have greater similarity to the category. If the value were too low, however, then a second stage would follow in which defining features only were

compared. It would be the need to undertake the second stage that would slow decisions for less typical members.¹⁹ The argument of these authors is, then, that typicality is a function of characteristic features shared. That defining features will not exhibit GOE structure is a characteristic theme of Smith's publications through Osherson and Smith (1981) to Smith and Osherson (1984) (see below).

Collins and Loftus 1975 extend the basic network theory of Quillian in various ways, and include adaptations designed to handle the prototypical structure of categories. Although the original model did allow differences in the criteriality of links (this had not been realized by many critics), Collins and Loftus elaborate the mechanism. They argue that matching processes involved in categorization judgements consist of the collection of evidence (positive and negative) until some threshold for acceptance is reached. Their contention is that this process is essentially the same as that described by Smith et al (1974) - typicality effects on categorization tasks are a function of similarity - although they go on to emphasise that subjects have other strategies available in this task, such as searching for a counterexample (eg. a hen that is not a bird). The collective view of these papers is sometimes termed the "probabilistic" view of prototypes (Medin and Smith 1984).

An essentially similar view is suggested by Jackendoff (1983), who proposes that "preference rules" underly categorization judgements. Jackendoff argues that, apart from a small number of necessary "well-formedness rules", the meaning of a word consists of a set of weighted heuristics (preference rules) representing typicality constraints on concepts. The closer to the default values of the more preference rules, the "better" the exemplar and the more "stable" the categorization judgement (with extra weight, of course, being given to preference rules with higher weights). Like Neisser 1976, he suggests that the overall schema or "set of preference rules" can be used to supply default values. Jackendoff proposes that this gives an account of the stereotypical character of images in the limiting case of projection of [TOKEN, TYPE[Y]]; what gets projected is a skeletal #thing# daubed with the default attributes of TYPE[Y]. Although Jackendoff is critical of Smith et al's (1974) notion of definingness, his proposal to weight the heuristics is equivalent. Further, he fails

¹⁹ It is not clear from their paper what would occur in the second stage that does not occur in the first, because they are not explicit about the averaging mechanism of that global comparison.



to offer a detailed account of the speed of categorization facts described above.²⁰ (We shall return to the issue of "necessary" or defining conditions below). Jackendoff's work is nevertheless interesting, particularly since he surveys a range of other psychological phenomena which such rules might underly and draws out some tentative links with recent work on the function of the neocortex.

I will argue that although goodness-of-exemplar ratings indirectly reflect the meaning of a word, there are some good reasons for doubting whether the semantic part of a lexical entry, the lexical concept, should be characterized in such terms. Four arguments will be provided in support of this contention: (i) prototypes cannot be combined; (ii) some properties inherent in prototypes are not inherent in lexical items; (iii) prototypes are not coherent in isolation from cognitive models; and (iv) prototypes are not constant (i.e. their inferential relations vary). In sum, prototypes do not display those properties that would enable lexical concepts to play a role in a natural language understanding system. Let us examine each of these deficiencies in turn.

4.1. Prototypes can't be combined compositionally.

One way of formalising the concept of prototype that has been adopted in the literature is Zadeh's "fuzzy logic" (Zadeh 1965). The key attraction of fuzzy logic is that it replaces the characteristic functions of set theory, which are binary (an entity is either in the set or out of it; the predicate is either true or false of any particular entity; the "middle" is excluded), with characteristic functions which assign values anywhere on a continuum between 0 and 1. Such functions offer an appealing analogy with distributions of "goodness of exemplar" (GOE) ratings (Roth and Mervis 1983). Thus, it has been argued, degree of membership can be formalized in terms of the relative truth of the predication.

Discussion centres around the translation rules of fuzzy logic, and so one has to begin by emphasising that translation rules in any logic can be formulated independently of the model theory, and are not fixed once and for all time. Thus, one of the manoeuvres made to avoid criticisms has been to suggest alternative translation rules.

²⁰ He talks of the relative stability of competing categorizations. He may have in mind a parallel distributed processing (PDP) model of such phenomena.

Translation rules function in a logic to indicate how, given the interpretations of its constituents, the interpretation of some construction may be determined. Thus, for example, in a standard predicate logic the set of objects which are truly *red squares* is defined as being the intersection of the set of objects which are red with the set of objects which are square. This is specified by an appropriate translation rule.

One contention of critics is that the corresponding function in fuzzy logic, telling you how relatively true it is that some object is a *red square* from information about how red and how square it is cannot be defined. Osherson and Smith (1981) examine fuzzy intersection in relation to the concept *striped apple*, which they take to be intersective. The obvious translation rule in fuzzy logic (the "min" rule) entails that the "goodness" of a complex concept cannot exceed its goodness as an example of either constituent concept. But, argue Osherson and Smith (1981), a *striped apple* will be a better exemplar of the category *striped apple* than of *apple*, and so this translation rule is not adequate.

Oden (1977) proposed as an alternative, a multiplicative rule, which he attributes to Gougen (1969), whereby the GOE rating of a complex concept will be the product of the GOE ratings of its constituents.²¹ To test this he asked subjects to judge the truth of pairs of sentences e.g. *robins are birds* and *beds are furniture*. He systematically varied the degree of rated truth of the individual statements in such pairs and attempted to predict the truth of their conjunction. What he found was an interaction between the ratings of the constituent statements. This could be better modelled by the multiplicative rule than by the minimum rule. Although the multiplicative rule is a better match to subjects' intuitions about complex statements, it too has serious shortcomings. For example, an object cannot have a higher rating with respect to how good a member of the conjunction of A and B it is than its rating for exemplariness in either of the constituents. Indeed its GOE rating must be lower, unless both of its constituent ratings are 1. Hence this rule, like the minimum rule, cannot allow any striped apple to be a better striped apple than it is an *apple*. Nor, to take a less exotic example (Osherson and Smith 1982), can it allow a *pet fish* to be a better example of a *pet fish* than it is an example of a *pet* or of a *fish*, and this is clearly not the case. A *goldfish* is a good example of a *pet fish*, but it isn't

prototypical of either *pets* or *fish*.

Osherson and Smith (1982) sketch a proof that, in fact, no simple function of the constituents can adequately characterize their conjunction. This is done by imagining a cube and a ball which metamorphose in such a way that each gradually approaches the other's shape. At the half-way point, the object should be equally as good a *ball* as a *cube*. If the goodness of something's membership in a conjunction is a simple function of its goodnesses in the constituent categories, then this entity should be a *round ball* to the same degree that it is a *round cube* - a conclusion that violates one's intuitions. Thus, no simple function of constituent goodnesses can predict correctly the goodness of an exemplar in a complex category.

By way of response to these arguments, Zadeh 1982 discusses the "normalization" of complex concepts (raising the maximum value of the set to 1 by dividing all values by the pre-normalization maximum). This is said to be brought about through focussing on an intersection "by giving it a label" (p291). This makes it technically feasible for a member of this set to have a higher degree of membership of it than of either of the constituent sets. Like ranking (discussed below), this result is achieved by considering the "absolute" goodness of members of a complex category to be a function of their relative goodness with respect to one another independently of their goodness in the constituent categories, while deriving that relative goodness from goodness in constituent categories. Both work (to oversimplify) by assuming that in any category there will be a "best" member, and stipulating that that member's rating for goodness should be 1. The crucial difficulty with normalization, pointed out by Osherson and Smith (1982), is that, if it is based on the "min" rule it only takes account of one half of the conjunct. For example, two red squares will be equally good as *red squares* so long as their values for the poorer of these properties are equal. Thus, if they are equally poor shades of red, even if one is a worse square than the other, they will be equally good *red squares* (so long as both are better "squares" than "reds"). This is counter-intuitive. Osherson and Smith (1982) also demonstrate that if a multiplicative rule underpinned normalization, then if the best *red flower* were a perfect flower it could not be a perfect red thing, because its goodness as a *red flower* would have to be less than 1 in order to prevent the product of the constituent goodnesses of any *red*

²¹ As Oden (1977) notes, both rules have the merit of yielding the same results as the rules of standard logic when truth is idealized.

flower equalling their normalized product, and hence the goodness of the complex concept. But, these authors ask, what contingency is there among properties that impels no entity to be perfect in more than a single respect?

Jones 1982 agrees that a flaw with the "min" rule is that it uses information from only one of the constituent concepts. Taking as an example the concept *bird*, he advocates defining the conceptual domain not as the set of all readily envisionable birds, but as the ranked set of all possible birds. Degree of membership is then defined as the inversion of the rank of an exemplar, with better members having lower ranks. The translation rule can then be that the goodness of a representative of a complex concept is the inverted rank of the negative product of its goodness in each of the constituent concepts. This rule is demonstrated in his paper to give reasonable results, and works because the relative order of exemplars within a complex category determines their goodness. The problem with the scheme, as Osherson and Smith (1982) point out, is the invocation of uncountably infinite domains - if the domain is uncountably infinite (and Osherson and Smith demonstrate that it must be), then there won't be enough integers (there's only a countably infinite supply of integers) to rank them all. Changing the domain to the extant set of exemplars would be objectionable since it would suggest that entities one had never encountered would determine one's concepts. For any other specifiable domain (see Osherson and Smith (1982) for a challenge that amounts to this claim), it will remain impossible to provide an ordering, since for any two exemplars you might order, there will always be an intermediate entity.

Even if ranks could be assigned, there are instances of concept conjunction that Osherson and Smith (1982) assert would defeat the scheme. Jones' rule could not permit an entity which is *prima* in two categories to be a better exemplar of their conjunction. But, although Paul Newman [was] both the best actor and the most handsome entity, he [was] more highly rated as a *handsome actor*. It's a bit like the best athlete was a decathlete, who regularly scored more points than anyone else ever had, yet never broke a world record in any individual event. Paragons, it would seem, rise above the canons of logic.

Cohen and Murphy (1984) describe concept combination as:

"a process of combining representations according to certain generative rules within a domain of knowledge." (p51)

They take from the knowledge-representation literature in AI the notion that sub-concepts

may specialize a super-concept by role modification (nouns are treated as frame-like, having slots called "roles"). They emphasise that the process is "asymmetric [and] knowledge-dependent" (p51), where by knowledge-dependent they mean, for example, that which role gets modified will depend on context, different contexts promoting some roles for modification more strongly than others. They argue that domain knowledge also modifies, for example, some of the role values of *pet* when it is used in *pet fish*. Thus, knowledge of the world influences the typicality structure of complex concepts which consequently, they continue, will not be predictable on the basis of the constituent concepts alone. They also note that the effect of a role value specification on typicality can depend on the setting of other roles. For example, the effect of baldness on a man's typicality depends on his age. Effects of correlational structure such as this would, then, have to be considered in evaluating typicality with respect to complex concepts, but this would entail further use of domain knowledge.

A crucial problem facing any approach to conceptual combination is that the two terms in the conjunction may not be interchanged without altering the typicality rating of an exemplar in the complex concept. Although the min rule is undoubtedly wrong to consider only one of the constituents in determining exemplariness in the complex concept, it is not the case either that each contributes equally, as would be predicted by the multiplicative rule. As Hampton (1987) has recently shown, for example, a constituent contributes more to the determination of exemplariness in a complex concept when it is playing the role of modifier than when it is playing the role of the head. His study used expressions of the type: noun plus relative clause modifier, for example *Games which are sports*. He found that a given instance (eg. *chess*) might be rated more highly in relation to this form of the putative conjunction than the other (ie. *Sports which are games*) Moreover, whenever there is a "head-modifier" relationship between the terms, the head is typically the one that determines the superordinate category²². Thus, for example, a *cow catcher* is a type of catcher, and not a sort of cow; a *milk bottle* is a kind of bottle, and not a dairy product; and a *pancake* is pottable but not a pot. Such non-commutativity confounds the attempt to model the combination of prototypes by means of fuzzy logic.

²² Hampton (*op. cit.*) has produced tentative evidence that entities may be treated subjectively as members of a complex category even if they are regarded by the same subjects as non-members one of the constituent categories.

Thus there appear to be two difficulties confronting attempts to model the combination of concepts by means of fuzzy logic: first, the knowledge-laden nature of combination, and second its asymmetry. However, even if one could find a way to combine prototypes compositionally, there are other reasons for not adopting them as a lexical representation.

4.2. Lexical indeterminacy of types.

My second reason to doubt that lexical concepts have a prototype structure is motivated by the finding that some properties inherent in prototypes are not inherent in lexical items.

Nunberg's first and obvious point is that many words can be used to refer to more than one sort of referent. He argues that this is a very general process that is not adequately handled by trying to list all the different senses. He proposes instead to list only a few senses from which other uses may be derived by general pragmatic principles. The principles may be summarised in the following way: the designatum, or basic sense of a word, can be used to refer to an entity if the cue validity of the referential function (RF) mapping from the designatum to the entity is sufficiently high to clearly and unambiguously pick out the entity in question from any alternatives. Cue validity is the exactness with which cues (in this case, referential functions) pick out a given referent. Take for example, an exam script. This object may be considered with respect to two relations: firstly, it is related to the set of examination candidates, and secondly it is related to a set of markers. Now candidates produce only one script while markers handle more than one, and indeed, more than one marker may be involved with a single script. Therefore, the referential function from the exam script to the candidate will have a higher cue validity than the referential function from the exam script to the marker. Consequently, if a speaker says, while pointing to the exam script, *He has a good grasp of the subject, he* will refer to the candidate. (Given certain assumptions about the background knowledge of speaker and hearer).

Nunberg's argument continues as follows: Just as we can talk about relative cue validity of referential functions in cases of deferred ostension, such as the one we have just described, we can also use relative cue validity to determine which senses of a word are most basic. Consider the word form *Dylan*. This can be used to refer either to a singer, Bob Dylan, or to his work, the songs of Bob Dylan. Which of the two is more basic? In order to compute this, we must determine the cue validity of the referential function from singer to song and the inverse referential function from song to singer. The more basic sense of *Dylan* is

taken to be the one in the domain of the referential function with higher cue validity. Since the cue validity of the referential function from Dylan to his work is greater than the cue validity of its inverse, *Dylan* designates the man.

Nunberg next considers the cases in which we are most interested for the purposes of this paper, cases in which it is not clear which of the two referential functions has greater cue validity. That is, there are cases in which cue validity appears to be equally high in both directions. Of his several examples we will consider just one. It concerns the choice between a token or a type reading of a phrase, e.g. *three trees*. Either this refers to tokens (the trees themselves) or it refers to types (different species of trees). In this example cue validity is the same for the referential function from trees to species of trees as it is for the inverse function from species to trees. That is, each tree is of just one species, and each species picks out just some trees and not others.

If Nunberg's arguments about cue validity are correct, then his conclusion with respect to the type-token distinction is of great interest and relevance to our own argument. Rumelhart (1987, talk to PHLING Conference, Edinburgh) has recently observed that prototypes could as easily be termed types. But if that is the case, choosing prototypes as the characterization of lexical concepts would imply that the type sense is more basic, with the token sense being derived from it. However, this is precisely the asymmetry that Nunberg argues we cannot empirically justify. Therefore his data do not provide support for the view that lexical concepts are to be characterized in terms of prototypes.

4.3. Coherence of prototypes.

Prototype structure has usually been described with respect to some similarity metric, where similarity has been taken as both the criterion of category membership and the principle underlying conceptual coherence. Things belong together in a category if they are sufficiently similar to the prototype. The prototype reflects the central tendency of instances in relation to the dimensions along which similarity is computed. However, this position has been challenged by those who believe conceptual coherence is determined by the theories people have about their physical and social worlds, and in terms of their current goals. According to Murphy and Medin (1985), GOE ratings that we discussed earlier are "heavily constrained" by knowledge of the world, and conceptual coherence is derived from a theory that explains why things are the way they are.

As Miller (1978:104) once wrote: "lexical knowledge ... is not isolated from [the] general conceptual system; the lexicon has a cognitive structure only because it is an integral part of everything a person knows and believes." Earlier, he and Johnson-Laird (Miller and Johnson-Laird 1976) had argued that semantic fields have a conceptual core that derives from the implicit theories shared by members of a linguistic community. In a similar vein, Lakoff (1987) observes that our categories are not determined with respect to properties of the real world, but in terms of "interactional properties" related to our exchanges with that world. And this is so even in the case of basic level categories. We "structure and make sense of our experience" in terms of cognitive models (roughly, what others have called variously images, frames and scripts). Cognitive models are defined "relative to idealized circumstances". Consequently, idealized cognitive models are what give rise to the prototype structure of our categorizations, not the lexical concepts that cue them.

4.4. Instability of prototypes.

Another doubt as to whether lexical concepts are prototypes arises from recent studies which have provided evidence that typicality structure is not a fixed property of a word, and that its focus can be shifted by context.

Roth and Shoben (1983) base a study on a finding of Garrod and Sanford (1977) that the speed of anaphora resolution depends on the semantic relatedness of anaphor and antecedent, as if resolution involved category membership verification (when the antecedent is more general than the anaphor). Roth and Shoben examined whether reading times were affected by a single GOE structure in differing contexts, or whether context could affect speed of resolution by changing GOE structure. The finding was that resolution for typical exemplars (typical for the category in isolation) was slower when the context was biased to make that exemplar a bad fit than when the context was neutral. Thus *robin* is a poor exemplar in the context *The bird walked across the barnyard*.

Barsalou and Sewell (1984, reported in Barsalou (1985)) found that GOE structure can change dramatically when people take different perspectives on a concept. Their subjects rated exemplars (e.g. *robin*, *ostrich*, *swan*) for typicality as *birds*, but assigned them different degrees of centrality according to the cultural perspective ("American", "African", or "Chinese") they had been asked to assume. These results held for both taxonomic and goal-derived categories. Thus, people can flexibly generate different concepts in different

contexts.

Barsalou (1985) also found that the way gradedness is determined varies with context. Before rating artificially constructed exemplars (fictional people) as members of certain categories, subjects performed one of two discrimination tasks which involved selecting the "best" candidate for some job (employment). For half the subjects the fictional people were divided into "PE teachers" who jogged with varying degrees of regularity, and "current affairs teachers" who read newspapers with differing frequency. For the rest of the subjects these joggers and readers were known as "Q" and "Z" programmers, a designation unrelated to the defining features of the categories. Additionally, all joggers shared up to three "characteristic pastimes"; newspaper readers shared a different three. Barsalou found that the basis for GOE ratings was principally the category ideal when this ideal was related to the name of the category (i.e. better exemplars indulged more frequently in the defining activity), but that GOE structure was a function solely of central tendency when the defining activity was not related to occupation.

What these various findings illustrate is that prototypes are unstable; they vary with context.

4.5. Summary.

Four arguments have been presented indicating that lexical concepts are not idealized summary representations corresponding to prototypes. Prototypes as defined appear to be unstable, hard to combine, too specific and not independently coherent.

Chapter 5.

Lexical Concepts

The previous chapter has cast doubt on the viability of characterising lexical concepts as prototypes. The requirement that in the process of comprehension these prototypes must be combined appears to be difficult to satisfy within the constraints of compositionality. Furthermore, the prototype associated with a label is not stable.

To address the question of what lexical concepts are, let me first ask: if prototypes are not the semantic portions of lexical entries, then where does GOE structure come from? Murphy and Medin 1985 argue convincingly that typicality structure derives from the (more or less naive) conceptual theories people have about the world. Relevant arguments were given in chapter 4. GOE structure, they say, is "heavily constrained" (p304) by knowledge of the world: conceptual coherence is derived from a theory that **explains** why things are the way they are. For example, birds can fly **because** they have wings. A concept would be represented as a set of correlated attributes "plus underlying principles that determine which correlations are noticed" (Murphy and Medin, 1985 p298, Table 1). With the aid of Lawrence Barsalou, Murphy and Medin develop an example showing that, given an appropriate scenario, a superficially odd category such as "prime numbers-or-apples" could be coherent.

Schank, Collins and Hunter 1986 develop an analogous position with respect to concept formation, rejecting knowledge-blind induction. Jackendoff's (1983) expression of a similar view was discussed in chapter 3. He argues that categorisation is a function of judgement.

These views are close to Carey 1984 who argues that the oddities of definition which Piaget observed in young children are not a consequence of any inadequacy of representational format, but rather are due to an insufficiency of domain specific information.

There are also correspondences with the discussion in Miller and Johnson-Laird 1976 of the "cores" of semantic fields. For them, lexical concepts are sections of theories. That is, meaning is given in conceptual terms. Conceptual cores are shared within a linguistic community, but the conceptual periphery can differ between individuals, particularly (it is claimed) as a function of age and education. Miller and Johnson-Laird is a large and detailed study, and perhaps as a consequence of scale, ambiguities in the account arise sometimes, with the authors appearing to say incompatible things at different points. Specifically, the role of decision tables, and their relation to conceptual cores is not clear. Decision tables serve processes of identification and recognition, coordinating the functional-perceptual schemas of a group of lexical concepts organized by some conceptual core, which provides a relative location for each lexical concept in a given field. The decision tables, it is argued at one place, interrelate lexical concepts through shared functional or perceptual conditions. However, slightly earlier (p696) it is argued that the connotative part of lexical concepts includes relations to other concepts.

A second ambiguity concerns the relation between lexical concepts and cores. On p300 it is suggested that decision tables and theories can be revised separately and independently, but the subsequent characterization of lexical concepts as including functional-perceptual schemas, knowledge and relations to other concepts suggests that the lexical concept is a **part** of the theory and that therefore any revision to the lexical concept must *de facto* be a revision of the theory. Perhaps, however, they have in mind a distinction between local revision and global reorganization. It would seem to be fair to conclude that Miller and Johnson-Laird assert that lexical meanings are concepts and that such concepts exist only within some (naive) theoretical framework. Miller 1978 states the following:

"lexical knowledge ...is not isolated from [the] general conceptual system ; the lexicon has a cognitive structure only because it is an integral part of everything a person knows and believes."(1978,p83)

Barsalou 1986 notes also the ability of people to construct *ad hoc* categories as required (particularly goal-derived ones). He proposes that "category representations are temporary constructs in working memory" (1986, p651). These representations will reflect context-independent (CI) information, which is stably represented in memory, context-dependent information (CD, see Barsalou 1982 for details of the CI-CD distinction) which is added. Recently occurring information also gets added; this is required to account for data he has showing that GOE structure of concepts changes within individual subjects, performing in

identical contexts but at times separated by relatively short intervals (a week). The direction for research that he stresses is the study of "how explanation guides the construction of temporary representations in working memory" (1986, p651).

The role of background information in the form of "theories" is crucial in determining the GOE structure of a concept, then. And instability can be attributed to flexibility in the choice of theory.

Now we can look at lexical concepts. A characterisation is needed that allows the lexical concept to interface with (different) theories, but which also allows it to combine with other lexical concepts.

Barsalou has partitioned concepts into three sets of properties, and has proposed that these properties be constructed into concepts under the rein of something like the "theories" of Murphy and Medin (1985). I propose to follow Barsalou in treating concepts (as traditionally defined) as, largely, "fictions created by theorists out of convenience" (1986, p652). To say that the meaning of a word is a concept is to adopt the position of a lexicographer; a perfectly respectable stance, but one tangential to the investigation of the comprehension of language.

I propose, however, that, rather than having a set of context-independent properties in the lexical entry, lexical concepts consist of unvalued parameters. These parameters cue information from currently active theories. Lexical concepts can be combined before cueing, so that the information cued is determined by a composite lexical concept. In other words, combination affects the way information is cued.

Two considerations favour this view.

Firstly, context-independent properties would have to present the same difficulties as default values. Any default values and GOE structure associated with CI information has to be derived from some theory because, as we have seen, prototype structure is merely a symptom of theoretical organization. Unless we are convinced that CI properties have fixed values, lack graded structure and remain constant under combination, we cannot unhinge them from theories. The values for the parameters will be provided by some currently active theory. Once the parameters of a lexical concept and the structure of a (group of)

theories have been combined, it may be possible to discern a concept. However, the description of "reference with" in chapter 2 is a reminder that a concept need not be generated directly.

Secondly, one argument given against the notion of "necessary conditions" is that, for example, a "bird" remains a "bird" even if its wings are removed.²³ (This argument is used by Collins and Loftus (1975) to criticize the notion of defining features in Smith et al (1974)). In other words, it is difficult to think of an aspect of the meaning of a term that cannot be optional in some context. Cohen 1979 argues that examples of such "cancellation" (as he terms it) are frequent: *stone lion* and *plastic flower* being but two he gives. His proposal, then, is that rather than have two words *lion*, the linguist does better to allow an adjective like *stone* to cancel a feature such as [+ANIMATE].²⁴ It's significant that this kind of modification does not lead to any sense of "anomaly" (unlike examples such as "sincere cobble" and "mauve idea"), even when the entity portrayed is thoroughly novel, like a "vegetarian lion". Let me make the following assumption: the bizarreness reaction characteristic of anomaly occurs due to difficulty encountered in combining concepts. If this assumption is correct, then the examples also suggest that anomaly is not born of a lack of fit between the words and the preferred scenarios of background knowledge. That is, it is not a clash of values but the absence of an appropriate parameter that creates the difficulty. This suggests that modification sets a value on a parameter of meaning, and if a suitable parameter is available (whatever its usual value) the modification is transparent and does not give rise to anomaly.

²³ I am aware that Barsalou's CI properties do not have epistemological necessity associated with them. That's why I depend here not on the existence of the examples but on the lack of anomaly surrounding their use.

²⁴ Actually, of course, it's a <feature, value> pair. Note that here our set-theoretic metalanguage begins to fail us, as might have been predicted on the basis of the discussion in Ch1. For example, is the set of "stone lions" a subset of the set of "lions"? I don't think so. This, incidentally, is further evidence that combinations need not inherit necessary properties of their constituents. It reinforces the argument that concepts such as "lion" with necessary conditions like "isa biological entity" are conceptual rather than lexical; they arise only when a lexical concept and theory are joined. That is, the lexical concept "lion" does not by itself have these properties, and when it is interfaced with a different theory under adjectival constraint (eg. "stone") the properties may be absent.

The relationship between this alternative and Barsalou's view can be viewed in terms of a distinction between type emergence and type coercion.²⁵ Under type emergence, a property emerges in the course of processing, whether it is tame or exotic. Parameter setting would, for example, be a case of type emergence. On the other hand, if values are set in the lexicon (like Miller and Johnson-Laird's definitional properties, or Barsalou's CI properties), then the type must be coerced into taking a new value (or a value outside the default range, for defaults). This is closely connected with the distinction between selecting and creating meanings (Clark and Gerrig 1983); do the processes of language perception select among candidate lexical meanings, or can they construct new meanings "on the fly" so to speak?

This distinction is also applicable in the domain of syntax, particularly in connection with "recategorization". Words sometimes crop up in unusual syntactic contexts. For example:

(36) Who did you sneeze before seeing?

SNEEZE is recategorized as a transitive verb here. There are two ways of looking at this. Either SNEEZE is an intransitive verb which is changed into a transitive one, or else it is unspecified for syntactic category, and this information is assigned by context. The "selection restriction" approach gives SNEEZE a categorization which would rule out this string (for reasons of ungrammaticality). For example, in Unification Grammar SNEEZE would have too few parameters to absorb an object nounphrase. In Cohort theory (Marslen-Wilson and Tyler 1980) it is assumed that the cohort would have to be depleted on the first pass because the syntactic information about the word to be identified would exclude SNEEZE; a new cohort would have to be formed with syntactic censorship abandoned before SNEEZE could be recognised.²⁶

The difficulty lies in explaining why SNEEZE is given the particular interpretation it is - the syntactic information is used in deciding what that interpretation should be. In the framework of type coercion, SNEEZE is converted into a transitive verb, and hence type coercion turns selection restrictions into transfer features (Weinreich 1966). Note that a type emergence account would imply that subcategorization information could not be used

²⁵ I am grateful to Ewan Klein for sketching type coercion for me; this passage would not have emerged without that sketch.

²⁶ A more recent formulation of the role of context in the cohort model (Marslen-Wilson 1986), still fails to address the role of context in interpretation.

predictively by the HSPM. Other examples are described in the next chapter, in which linguistic criteria for determining which account is appropriate in each case are discussed.

Returning to word meaning, we can ask "how specific are the meanings in lexical entries?" For example, does a noun like CHEESE have a mass interpretation in the lexicon, or is the mass-count distinction an emergent property of nounphrases, with lexical entries vague in that respect? In exploring these issues and developing this account, we shall turn shortly to a more thorough discussion of the mass-count distinction in English. Particular attention will be paid to the tension between the apparent existence of shadow lexicons (potential lexicons which contain for whole classes of words counterparts which have undergone a uniform semantic transformation) on the one hand, and phenomena of lexicalization on the other.

5.1. Lexical concepts, theories and lexical flexibility.

This section indicates in general terms how the characterisation of lexical concepts and the relationship with theories that I have described can deal with various phenomena of lexical flexibility and precision. Specifically, I discuss diachronic change, metaphor, classical concepts and conceptual interaction.

5.1.1. Diachronic change.

Splitting concepts between lexicon and theory clearly gives us a handle on certain diachronic phenomena. It is frequently asserted that the meanings of words change as the theoretical understanding people have available in relation to the entities denoted develops. This suggests that as the conceptual system changes then so will the meaning of words. Cohen (1962) describes in some detail the history of the debate on this matter, particularly as it affects diachronic change (see also Cohen,1979). Thus philosophers have argued that word meaning is in 'flux' (Robinson,1953 ; Sadock,1979) and that it has fluidity (Shibles,1977). Vygotsky (1962) gives this notion perhaps its clearest articulation:

"The relation of thought to word is not a thing but a process, a continual movement back and forth from thought to word and from word to thought. In that process the relation of thought to word undergoes changes"p125.

This is easy to model here simply as a change in the theory. It follows inevitably that any concepts later generated with that theory will be different from what they could have been before. Piaget and Inhelder 1968 provide striking evidence for similar phenomena within the developing child. In one experiment, he showed non-conservers a tilted beaker of water and asked them some days later to draw what they had seen. Their drawings reflected a lack of understanding about the way fluids behave. However, when he asked them on a subsequent occasion - after they had achieved conservation of volume for liquid (a Piagetian 'milestone') - to reproduce what they had been shown several months before, he found that their memories had 'improved'. These results are interpreted to mean that memories are interpreted in the light of the child's **current** knowledge of the world.

5.1.2. Metaphor

Some types of metaphor may be amenable to analysis in terms of a change in domain. Fraser 1979 presented a group of subjects with metaphors and invited them to interpret these out of context. Although none found any difficulty in supplying an interpretation, the interpretations varied a great deal from subject to subject. This suggests (if we assume that metaphor is an efficient way of communicating accurately) that, normally, contextual considerations constrain the process of understanding which interprets a metaphorical expression. (This isn't to say such considerations aren't applied to "literal" utterances). Indeed, speech act theorists, such as Searle 1979, take this to be axiomatic, arguing that metaphorical utterances are understood precisely because a literal interpretation would be nonsense - and so the speaker makes use of contextual information in order to construe the utterance informatively on the second pass. This approach which characterises metaphors as garden paths does not, however, coincide with the available psychological evidence.

Rumelhart 1979 reports an experiment by Gibbs in relation to the comprehension of indirect requests. Gibbs tested the three stage "speech act" model of Clark & Lucy and found that it accounted for the data only if test sentences were presented out of context. Tourganeau and Sternberg 1982 report an experiment by Harris (1976) in which no differences were found in the processing times for literal and metaphorical sentences. And, finally, Ortony et al. 1978 found a difference but that this was diminished by extended context.

This evidence suggests that a single process handles both 'literal' and 'metaphorical' language, with deeper understanding of utterances emerging only over time (cf Rumelhart's notion (1979) that they are fitted to schemata, which are similar to Schank's scripts (Schank and Abelson 1977), as particular instances of those general abstract concepts and understood in terms of them). The point I wish to emphasise here, however, is that this evolution of meaning appears to occur in the same way for both literal and metaphorical utterances. The more recent finding of Janus and Bever 1985 has to be taken into account too, however. They found that when reading times were monitored phrase by phrase rather than just sentence by sentence they were able to measure increased durations for sentential metaphors.

It seems likely that we can distinguish two sorts of metaphor: major and minor. That is, there is a distinction between elaborate "literary" and "scientific" metaphors on the one hand (Kittay and Lehrer 1981), and minor metaphors which involve merely the assimilation of a lexical concept to a novel domain. Here is an example of a minor metaphor:

(37) Wishing her pupils a good morning, the dragon demanded their homework.

In this example, the concept of "teacher" acts as a theory and provides an interpretation for all the shared parameters of "dragon", so that the "dragon" looks more or less like a "teacher". FIERCE, let us assume, is not a parameter of teacher and so this is added. However, the ferocity is the ferocity of a lady teacher. This will be different from the ferocity of a dragon or, more subtly, of a male teacher. That is, the interpretation of the added parameter is related to the topic.

This could, in a simplistic way, be formalized as the attachment of a new node to a theory which for simplicity's sake I'll assume is a fragment of semantic net. The node is the vehicle, and its site of attachment is the topic. By changing to different fragments of network, different metaphors can be generated. I call this concept "plastic inheritance": nodes can change their location in the net and so change their inherited default properties. This is reminiscent of the practice in lexicography of differentiating senses of a lexical item by relating each to a different field (Benveniste).

Of course, even minor metaphor is more complex than this. A more detailed account of plastic inheritance will be developed in chapter 9.

One might argue that the account relies on a kludge: the specificity of the parameter can be crucial. If, for example, TEMPERAMENT had been substituted for FIERCE, this presumably would have been a parameter of both "teacher" and "dragon". This dependence on specificity sneaks in default ranges. Even if parameters are characterized as free variables, information about type can act as a default. There are really two issues here. One concerns the nature of metaphor in general. A general theory of metaphor would say something like "the ground consists of the necessary features of the topic plus the characteristic features of the ground." I leave the specifics of this statement for the theory of metaphor. However, this version of the statement will illustrate my point. If the parameter was TEMPERAMENT, it would probably not be a context-independent feature of "teacher". On the other hand, it would be a characteristic property of "dragon". It would therefore be included. Let me repeat that it is for a theory of metaphor to indicate which classes of parameter are likely to be used and how this affects the quality of the metaphorical expression. The second issue is central to my account. Lexical concepts have no semantic value. Semantic value is only generated in the context of some naive theory which interprets the parameters. The default ranges implied by the differentiation of parameters are therefore different in kind to the defaults of prototype theory. Rather than being part of the material of a semantic representation, they are tools with which some representation can be constructed.

A small difficulty is that there is no sharp distinction between the two types of metaphor. Take, for example:

(38) The cup of the hill lay before us.

It's obvious that we are being told something about the shape of a feature on the hill, but how do we decide that it's not function that should be contributed ie. the cup of the hill might be the part that yielded water, for example. In other words, a typical metaphor will demand that the hearer deploy a certain inferential facility in addition to simple inheritance of default values.

While lexical meaning is flexible, lexical use can be very precise. Speakers can convey their ideas with great precision. This is paradoxical. How can vague words express precise ideas? The next two sections discuss mechanisms which allow this.

5.1.3. Classical concepts

My general approach can also deal with the occurrence of classical concepts, because we can allow that more precise theories can yield more sharply defined concepts. In other words, prototype structure is a contingent rather than a necessary property of the representation of background knowledge.

Medin and Smith 1984 discuss possible roles for classical concepts. One can imagine that classical concepts, like prototypes, might be derived from conceptual theories. They might, for example, be the product of rational (or rationalizing) thought about particular concepts, based on knowledge of the world, pretheoretic notions of what is plausible, and so on. I believe that this is what Vygotsky 1962 meant by the phrase *scientific concepts*. Such concepts are much closer to "ascribed meaning" than "lexical meaning" (see above).

Armstrong, Gleitman and Gleitman 1983 present evidence that even concepts which can be given a clear and complete definition (eg, "odd number", "woman") display a prototypical structure in the experimental paradigms alluded to in section 2.1 above. That is, not only are subjects willing to rate numbers such as 561 as "less odd" than numbers like 3, but they also take less time to verify that a statement describing the latter as an odd number was true than for the former. The experimenters' argument goes that, because the concept "odd number" is definitional, and yet there are exemplary odd numbers, then exemplariness is not directly concerned with category membership: their argument against accepting these results as confirming and extending the prototype model hinges on rejecting the possibility that people who could use numbers in arithmetic could be vague (conceptually) about their "oddness". To explain why people sometimes deal in prototypes and sometimes in classical concepts, we have to take into account the variations in attention and consideration that people give to ideas, and accept that on some occasions they bring more detailed knowledge to bear than on others (either because they have it, or because they can tolerate it). Note that more knowledge doesn't necessarily mean less conceptual overlap (Murphy and Wright 1984).

One of the difficulties with the classical view as a theory of lexical meaning stems from the fact that words must frequently be used to describe novel entities which only partially resemble familiar ones. That is, people put the vocabulary they command to the best use they can in describing their world and communicating their messages. As Slobin 1982

states,

"Language evokes ideas ... linguistic expression is a highly selective and conventionally schematic map. At the heart of language is the tacit convention that most of the message can be left unsaid, because of mutual understanding."(p132)

And hence the deployment of lexical items in the composition of this drawing can be a process of approximation - it is not expected to be exact; the hearers' processors expect to have to join the dots and shade the bounded areas. Thus, the word *game* is used both indoors and out, to describe the behaviour of teams and individuals motivated with as much variety of intention as exists among human beings. These are just a couple of the dimensions (or parameters) along which any particular *game* may take up a position. It is this chain of resemblances that leads to the "family resemblance" structure of categories which Wittgenstein 1953 described. And it is this family resemblance structure which makes it difficult to define concepts classically. Typically, classically defined categories are defined by human stipulation and can be found confined to some specific domain of rational human activity, such as the Law, a science and so on. Often in science such definitions are working definitions, acknowledged not to necessarily correspond with reality but to provide a necessary basis for rational enquiry. For example, one might want to define *literacy* as:

"being able to read page-X of book-Y and answer questions-Z correctly"

for the purpose of comparing teaching methods, while still accepting that there are better though vaguer definitions of the term.

This need to use words may well account for some of the overextensions in children's early language²⁷ (Anglin 1983, Nelson 1983). For example, my own daughter (Romy) employed the form /wiŋ/ (*swing*) to label a drawing in a book of children playing on a slide (at 18mths). By 22mths, however, /ʃut/ (*chute*) was in her vocabulary and was used appropriately. How should this change be characterized? Is it that adding a new item narrows the meaning of /wiŋ/, or is it rather that there is no longer any need to call upon /wiŋ/ to enter the breach? This requires further research. One test might be corrigibility; one would expect a child with an overextended lexical entry to be less willing (or able) to be corrected than one pragmatically solving a problem.

²⁷ But probably not for underextensions. See Anglin (1983).

Classical concepts, I have argued, represent theoretically motivated precisizations of lexical concepts. Their logical character and special inferential properties derive from the tight structure of the theory used to actualise the lexical concept.

5.1.4. Contextual combination

Nida 1975 has suggested that, in general, the larger a text is, the more precise the meaning becomes. This suggests a second mechanism for precisization: the actualised concept is more precisely defined by a text which places more constraints on its interpretation. In short, it is the process of combination which introduces precision. More generally, this approach also offers an account of contextual interaction, to which I have already alluded and which I will discuss in more detail in the next chapter.

The obvious example is adjective-noun combination. In combination, lexical concepts partially overlay one another in a manner governed by the syntactic structure of the text. Thus, for example, an adjective will overlay a noun, masking some parameter(s). For example:

(39) a red cup

In this example, the adjective masks the parameter COLOUR. Masking will constrain the way information can flow into the semantic representation from background knowledge. On the one hand, this makes the actualised concept more precise, and on the other it creates the conceptual adaption evident in semantic interaction. Conceptual adaption arises as background knowledge flows. For example, if the current theory specifies that the default material out of which cups are made is china, then that expectation hones the flow of information about colour. Thus, there is a process of mutual constraint and entropy reduction.

There are two further effects that arise from the way modification draws focus onto a property. First, the property tends to be given greater prominence. Barsalou (1982), in research discussed elsewhere in this thesis, suggests that this only happens for a subset of properties which he terms context-dependent properties. And second, there is a reduction in categorial fuzziness in relation to that parameter — the size of the borderline is reduced.

Within this framework, contrastive stress can be characterised as a hedge, masking the lexical concept with an implicit relationship or set of relationships. For example:

(40) a vase

This might be contrasted with *a cup* or, alternatively, *a bowl*. If it is the former, then the lack of a handle might be sufficient to render *a vase* a felicitous description of the object depicted in figure 1a. On the other hand, opposed to *a bowl*, figure 1b would be a good *vase*; a poorer one if opposed to *a cup*. The opposed category influences the instantiation of defaults. Obviously, this *gedanken* experiment might be checked empirically, using Labov 1972's techniques.

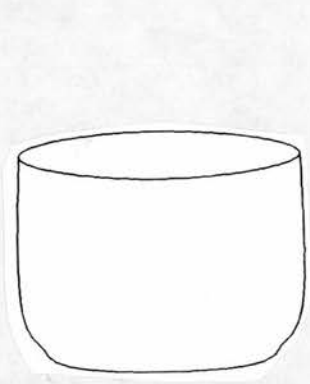


Figure 1a.

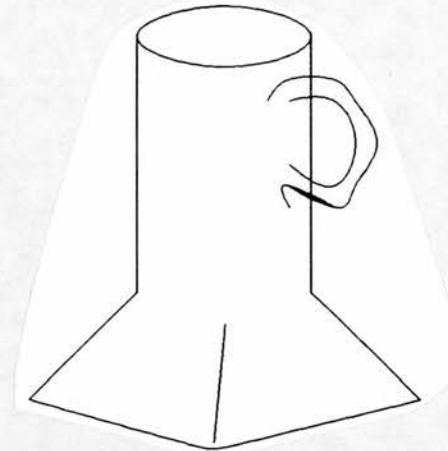


Figure 1b.

Paradoxically, contrast can also be used to redeem a borderline instance. For example, one might say that figure 1b depicted an unusual *cup*, only to be corrected "it's a vase". This type of use is connected with "pre-emption", discussed in chapter 2.

There are three problems that I have not solved that will arise with this approach. The first two are general, however. First, there is the problem of explaining why one theory rather than another is active at any given time. Second, I have not explained how an adjective knows which parameter(s) to overlay. We can't rely on there being a transparently appropriate parameter such as COLOUR available always. The process of "docking" modifier and head will typically have to be steered by background knowledge. This brings me to the third point - timing. I have not said enough about when information will start to flow from the different knowledge structures or about when and how semantic information from two items in a given syntactic structure will interact. I will have little to say about

these matters, aside from some general remarks in chapter 8.

Adjectival modification is discussed in more detail in Dunbar (in preparation, a) and Franks (in preparation).

5.1.5. Internal lexical structure

A final issue is the question of relationships among semantic properties. Many investigators (see, for example, Boas 1988) have emphasised the need for any representational scheme to supply information about the way the properties and parts of an entity are organised in relation to one another. I will deal with this question in detail in chapters 9 and 10.

Part 3.

From Lexicon to Phrase.

This part of the thesis examines the distinction between coercion and emergence. Chapter 6 develops the general ideas in relation to a variety of linguistic examples, some semantic and some syntactic. It shows that the distinction is relevant to these examples and introduces some criteria that might be used to test the distinction. Chapter 7 applies these and other tests to the mass—count distinction in English. Chapter 8 puts forward a theoretical account of the data that has been observed. In this part of the thesis, I focus especially on dependency relations and their influence on the way lexical concepts combine.

Chapter 6.

Dependency and Lexical Vagueness.

So far, I have begun to characterise a psychological, a cognitive account of meaning, and have examined the role of lexical concepts in that system. Various properties of lexical meaning have been described, such as the requirements for composition and interaction with knowledge, and I have begun to characterise lexical concepts in such a way as to take account of these considerations. The fundamental issue of concern in this thesis is, stated plainly, what information belongs in a lexical entry? So far I have been lead to the conclusion that, semantically, lexical entries contain unvalued parameters which act like sophisticated valves controlling the flow of information into some ultimate representation; the lexical concept is one tool in this constructive activity.

In examining the combination of concepts, the psychological literature has been handicapped by two difficulties, as I observed above. On the one hand, asymmetries in the process need to be accounted for, and on the other, the process must be successfully interfaced with knowledge of the world. In a later chapter, I will develop the ideas mentioned at the end of chapter 2 regarding knowledge, but in this chapter and the next, I intend to provide an account of one aspect of the asymmetry of combination - dependency.

I've already mentioned, in connection with nominal compounds, that we must expect the syntactic structure of language to contribute to the manner in which concepts are combined. I wish to spell out in some detail exactly what issues dependency relations raise for a psychology of the lexicon.²⁸

The other purpose of this chapter is to sharpen the coercion-emergence distinction by reference to specific examples. As well as clarifying the distinction, this chapter uncovers some criteria which can be used to test the distinction.

²⁸ Dependency, in general, is the notion that one constituent requires another.

In the next chapter I will examine a particular example of lexical combination to illustrate the relationship between combination and the coercion-emergence distinction. The example will be tested against the criteria listed here. To keep the analysis "clean" at this stage, I will choose an instance in which the role of knowledge of the world is simplified.²⁹

6.1. Dependency and type emergence

Briefly, there are two main roles of syntactic information in conveying meaning³⁰:

1. It makes relationships among the constituents of sentences explicit (both within and between the major phrasal constituents), providing information about function-argument structure that is central to the calculation of truth-conditions. This can be signalled by word order and inflection in different languages.

2. Prominence relations are signalled and thus discourse is given structure (sometimes called "given-new structure") through manipulation of focus. This is done through conventional deviations from canonical form, described in transformational terms by rules such as extraposition, clefting, pseudo-clefting and so on.

Both types of device may be used to suggest one perspective on some thing or event rather than another, and we can operationally characterise this as producing differences in the information about the entity that is made available. I have discussed in particular the amount of information made available in chapter 2, with reference to examples such as dative shift.

It is this role that I'll concentrate on here: the mechanism by which meaning can be manipulated by such factors as the presence or absence of linguistic context, the internal structure of linguistic context and so on. In other words, the way the flow of information can be manipulated by the way words combine and the properties of the words they combine with. An important issue will be whether when context manipulates meaning it coerces a

²⁹ Many thanks to Jim Miller and Ewan Klein for their comments on an early (1986) draft of this chapter.

³⁰ The function of conveying cultural solidarity by indicating a shared tradition will be set to one side.

change in an underlying meaning, or simply makes an adjustment in the way information flows.

There are various properties which, reflectively, we might be inclined to associate with some particular lexical item. In chapter 1, I identified a view of lexical meaning which sought to go beyond this lexicographic perspective, and the following sections examine the issues involved in more detail. In particular, how can one argue for associating some property with a lexical item as opposed to the phrase(s) in which it appears?

By examining these issues in relation to cases exemplifying the central phenomena, I hope to clarify such questions. I begin by illustrating contextual influence on lexical meaning, and examining its relation to the syntactic organization of language. Then, I consider how one should represent the flexibility of the valency of some English verbs. Finally, I consider the implications of alternative views as to the locus of certain aspectual properties.

Let's examine first the combinatorial relations between a verb and its arguments - here, meaning and dependency are closely entwined. Although most of the psychological literature on concepts has focussed on nouns, and occasionally verbs, a key issue in traditional grammar and modern valency theories is the nature of the semantic relation between a verb and its arguments. There are two complementary ways of approaching this question. On the one hand, to what extent can arguments tailor the meaning of the head of the construction? And on the other, to what extent are the number, category and role of its arguments inherent properties of the verb?

6.1.1. Semantic tailoring and dependency

In combining, concepts exert influence on one another, as we have seen. I turn now to further evidence of such effects, and examine their relation to dependence. Allerton (1982) employs the term "semantic tailoring" in this connection, which conveys the conceptual fine-tuning involved. Allerton argues that nouns tend to be stable in meaning in comparison with verbs and adjectives which are semantically tailored by the noun. In particular, the direct object of a transitive verb and the subject of an intransitive tailors the verb, standing in the relation of specifier to the verbal core (Keenan 1978, Smith 1980).³¹ Compare:

(41) The child/nose/tap runs.

(42) Mother stripped the girl/bed/tree/wallpaper/engine.

A striking parallel is available in the work in computational linguistics of DeJong and Waltz 1983. They argue that if the selection restrictions of direct object and verb fail to match, the object should be taken as referring literally and used to generate candidate processes. The most contextually suitable of these is then aligned with the verb actually used and - roughly speaking - their meanings are unified. DeJong and Waltz do, however, present as a caveat an example in which a verb can be the literal constituent attended by a metaphorical noun phrase, and it is trivial to demonstrate that any of the surface grammatical relations can be metaphorical. Nevertheless, it would appear that an interesting tendency has been described.

Similarly, although according to Allerton (1982, p19) an adjective is subordinate to a noun (whereas an object is semi-subordinate to a verb), it is worth recalling that Half, Ortony and Anderson 1976 found that the sense of RED changes according to the identity of the noun it combines with. Once more, however, we need the caution that the transfer of properties may be two way: recall that Barclay et al. 1974 demonstrated variation in the quality of a given property as a cue for a noun according to the nature of the verb; for example, *heavy* served as a better reminder of the word PIANO when the verb was LIFT. Furthermore, Barsalou (1982) demonstrates that the verb phrase in a sentence can affect the availability of properties of the grammatical subject.

A related approach has been employed in the Japanese MU Machine Translation project to aid the accurate translation of *Waga* verbs (basic verbs, functioning like MAKE, DO etc. in English) (Sakamoto, Ishikawa and Satoh 1986). In this project, the case of the arguments to the verb is used to restrict the possible meanings of the verb. The cases are in turn linked by co-occurrence restrictions to particular groups of the semantic markers which are used to classify concepts. For example, the verb:



³¹ I leave aside the question, raised by Allerton, of whether or not the direct object of a transitive verb is a "deep subject" of a corresponding intransitive.

is given a reading like HIT, STRIKE, COLLIDE when the "subject" is an object or physical phenomenon while the "object" is an object or place (yielding *light hits a slope*, for example). On the other hand, it is read as UNDERTAKE, BE ENGAGED IN, DEAL WITH when the "subject" is an object with will and the "object" is an action.

The question of whether a distinction in meaning arises through coercion or emergence is clearly germane here, and these examples indicate the generality of the problem, as well as its relation to dependence.

6.1.2. Transitivity and dependency

Next I tackle transitivity, which as a topic has long attracted students of meaning because it inhabits the area between syntax and semantics. Is the transitivity of a verb predictable from its meaning, syntactically inherent or an emergent property of a string (Kilby 1984)? Why are some verbs transitive, some intransitive and others flexible? Consider the flexible ones; why are some of the "ergative" type while others are not? There is not time to tackle all these issues, but in the case of flexible verbs the question can be put: is one pattern basic with the other derived by a coercive process or is the valency an emergent property of the sentence as a whole?

We have already seen that the problem of contrasting type emergence and type coercion arises with respect to examples like:

(43) Who did you sneeze?

Here a normally intransitive verb is used transitively. Ultimately such cases may be an interesting test, for if the manipulation of valency is ever an instance of coercion it must be here. However, since the purpose for the moment is partly to trawl the linguistic literature for advice, we will be better served by an examination of analogous cases which are part of conventional usage and which have therefore received the attention of professional students of the language.

Lipka (Lipka 1972) argues that subcategorizing verbs as bluntly transitive or intransitive is uninformative, pointing out in particular the productivity of rules adding a component CAUSE to the meaning of the verb (converting one-place verbs into two-place ones, for example) and deleting objects (reducing the number of arguments, of course). Clearly,

however, he views these processes as changing the underlying valency of the verb; the metaphor of "transformation" hangs in the background. Lipka cites Jespersen, Lyons and Poutsma along these lines, quoting the last-named on the "floating nature of transitivity" (Poutsma 1926 pp45-90), but this tradition has ignored the possibility of type-emergence.

Lyons 1968, in relation to object-deletion, notes that it has been suggested that *eat* in sentences like:

(44) We never eat at 5 o'clock.

is a pseudo-intransitive or "absolute" usage of a basically transitive verb. He also regards ergative verbs as being basically intransitive verbs which can be transformed by a causation "operation" introducing the causer as subject. For example:

(45) The wheel turns → John turns the wheel.

He also points out that in the past productive morphological processes have lexicalised similar transitives (eg. FELL, ROUGHEN, STERILIZE), and argues that, by analogy, the ergative verbs should be seen as undergoing a zero-derivation. In short, Lyons employs the morphological analogy as evidence that the process is coercive and, further, that the causative is the derived pole of that process.

Superficially, Halliday, to whom Lipka also refers, seems to lean more towards considering transitivity as an emergent property, since in systemic grammar the "transitivity system" is a system of choices (or "contrasts") available in, say, English, in respect of the number, role and type of arguments & "circumstances" (cf. *circumstances*), aspect and so forth, with choices being made at the level of the clause. In the jargon, the "entry condition" for the transitivity system indicates that transitivity is a property of units of the "rank" of clause (Berry 1975). The transitivity system can interact with the "voice system" leading to elision of elements. It is the case, however, that each new choice (or specification) within the system narrows the range of verbs that are permitted; in short, verbs are classified according to transitivity.

The alternative that is not in the end addressed, then, is whether it might in fact be the case, for example, that the presence of a direct object triggers the causative interpretation of ergative verbs, rather than insertion of the hypothesised causative zero-morpheme.

One writer who does address such questions is DJ. Allerton (Allerton 1982). He offers four possible accounts of surface differences in the valency of a single form.³²

a) There may be two verbs with the same form. For example, compare:

(46) They hailed the leader all day.

(47) It hailed all day.

b) A single verb may have two valencies, thereby exhibiting polysemy or "class cleavage", the latter term being due to Bloomfield.

c) A single verb may have a single (compound) valency which admits alternative instantiations.

d) Finally, a single verb may have a single (simple) valency which can be transformed by syntactic rule into another.

We will not be further concerned with the first of these; the third and fourth may be compared to type emergence and type coercion respectively, while the second I shall provisionally term "lexicalization", although this assumes more than Allerton does.

To which does he assign object-deleting and ergative verbs? And how does he defend the assignment? Object-deletion is given the third solution (compound valency): the object is bracketted in the valency frame and thereby signalled as optional (with the proviso that for some of these verbs an ellided object must be contextually recoverable). Very little argument is provided for the application of this solution to verbs whose objects need not be contextually recoverable (such as READ, WRITE, PAINT), beyond the claim that the meaning of the intransitive does not specifically exclude the understood object, which will in turn remain easy to "activate" (p70). It should be noted that Allerton has no qualms about the observation that some object-deleting verbs tend to have "absolute" readings when used intransitively (eg. *he has eaten, John believes, Sally drinks*), arguing that such hyponymic restrictions in verb meaning are "context-governed" (p71), with the ordinary

³² I retain Allerton's ordering since he derives terminology from it. He mentions the fifth possibility of a lexical transformation, but dismisses this, arguing that it must in practice reduce to either (a) or (d).

readings remaining potentially available.

In the case of the ergatives, Allerton regards the two surface forms as being too discrepant and the semantic relation too irregular for a compound valency solution to be tractable. He considers type coercion (ie. the fourth account) but argues against this first because it is not clear which of the two should be regarded as most basic (cf Lyons above) and, second, because the rule is not completely general, failing in the case of verbs like PROPEL and DRIFT (p75). His preferred solution for ergatives is, then, "multiple valency", the second account.

Although there is not time to elaborate the point here, I believe that it may be possible to overcome Allerton's objection to a type-emergence account of ergatives. We would require a means of deriving the variation in the semantic relations from the inherent meaning of the lexical items involved, together with a schema encompassing both possibilities. This may be possible if some suggestions made by Kilby (1985, ch2) are followed up in the spirit of Gibson and Langacker. Pressing for such a result would have two obvious benefits:

- 1) Removing multiple valency removes a local ambiguity and therefore simplifies parsing.
- 2) In order to derive the requisite schema, it appears that we will need to develop our understanding of the temporal contour of the interaction of lexical representations and knowledge of the world, a question of deep theoretical importance.

It should be noted that such questions have been of concern in both a-intelligence and psychology, as well as linguistics. Boden, for example, asks just which slots conceptual dependency schemata should include - how fine-grained should the representation be? In a-intelligence, however, workers have shifted the boundary of debate, tending to proliferate slots to enrich the representation. This is to be expected, given their emphasis on inference. In psychology some relatively recent work has examined differences in the way nouns and verbs are organised. In particular, Huttenlocher and Lui (1979) sought to explain findings like greater release from proactive inhibition for nouns (compared to verbs) partly in terms of the dissipation of semantic connections to concepts associated with the slots of verbs (see Dunbar 1988).

Both these approaches have focussed on the dependence of the verb on its arguments for full instantiation of its meaning. This is the other side of the dependency coin; in traditional grammar, the emphasis is placed on the dependence of the arguments on the presence of the verb, without which they could not appear. It is interesting that in his Space Grammar, Langacker chose to draw a contrast between autonomy and dependence in which it is the dependent item that requires the presence of slot-fillers. That is, in Space Grammar, dependence is semantic incompleteness.

6.1.3. Aspect and emergence

Let us push the next example further in order to sharpen the contrast between the mechanisms and to draw in the issue of ontological constraint. The contrast between emergence and coercion arises again in relation to aspectual distinctions found among verbs of motion in English: to what extent can a single verb accommodate more than one meaning?

In relation to this example, we can also address the more specific question: can a rule be formulated to the effect that there is a certain degree of difference of meaning such that no two meanings that far apart can be alternative, contextually determined, readings of the same word, and if they are expressed by the same word-form, then they are homonyms? We can profit here by considering briefly the ontology of concepts, because it is ontological barriers that we find are said to be insurmountable. Keil (1979) has claimed, for example, that ontological categories are arranged in a hierarchy subject to a constraint called the "M-constraint". The M-constraint simply states that the hierarchy must be a hierarchy: no instance of an ontological category may also be an instance of an ontological category on a different branch of the tree. The same assumption is implicit in Jackendoff (1983):

"Which major ontological category is expressed by a particular major phrasal constituent depends on the semantics of the head." (p68).

Jackendoff identifies two readings of word-forms associated with motion verbs. Take for example GO; compare:

(48) The lorry went from Ardlui to Tyndrum.

(49) The road went from Ardlui to Tyndrum.

He postulates two semantic predicates corresponding to these readings - GO and GO_{Ext} - but is unable to decide whether to equate these. Note that technically, within his system, they can't be identical because state-hood is a necessary condition on the GO_{Ext} predicate, while being an event is a necessary corollary of GO. Nevertheless, Jackendoff hedges on this, pointing out that the distinction between the two uses is largely determined by linguistic context. The examples given above illustrate this: the difference between *road* and *lorry* appears to tailor the meaning of the verb. In the second example, the subject has "... [low] motility" (p173) and we are hence inclined to assign a "state" reading, whereas the first example is interpreted as denoting an "event". We might also point out that the extensive interpretation appears to be accessible with all motion verbs, including novel ones. For example:

(50) The path darted about among the trees.

(51) From the helicopter we see how the city drifts into nearby villages.

(52) The hedge sinclair³³ around the cottage.

The problem for Jackendoff's theory as it stands is that if GO and GO_{Ext} are distinct predicates, then there will have to be a shadow lexical entry for every motion verb, whose necessary condition is the predicate GO_{Ext} rather than GO. Furthermore, there would have to be a meta-lexical rule predicting the existence of this shadow in order to account for the potential ambiguity of any novel denominal verb, for example. The key to this difficulty is that the semantics of the head of a major phrasal constituent are the fount of its ontological category, but ontological category is part of a necessary condition on motion verbs and cannot, as things stand, be evaded. And it is precisely because their ontological categories are different that GO and GO_{Ext} cannot be equated.

³³ Since this verb is (I believe) a new denominal, I should explain the intended meaning. A "sinclair" is a low, battery-powered, slim, three-wheeled bicycle. To *sinclair* is, therefore, to move with a low profile. Thus, for example, prone soldiers might sinclair across a field on their hands and knees in order to make small targets of themselves.

Let's press for a better analysis. Jackendoff (1983) rightly points out that Miller and Johnson-Laird (1976) cannot handle extensive interpretation in their account of aspect, but we will find that the process of amending their analysis so that it does will give us the clue to solving Jackendoff's problem. The relevant predicate from their account is TRAVEL:

TRAVEL(x): Something x "travels" from time t_0 to time t_m if, for each t_i such that $t_0 < t_i < t_m$, there is a place y_i such that $R_{t_i}(AT(x, y_i))$ and:
 (i) $R_{t_{i+1}}(\text{not } AT(x, y_i))$

(Where R is an existential quantifier over moments of time). An alternative and convenient notation is:

$$R_{t_i}(AT(x, y_i))$$

$$R_{t_i}^1(AT(x, z))$$

$$F_{t_{i+1}}(y, z)$$

(Where F is a predicate meaning that the locations "y" and "z" are distinct). If we add the following clause:

POSS(R1=R2)

and amend the subscripts of the quantifiers accordingly, we provide a definition which allows for both extensive and eventive interpretations. Note that this also handles cases which Jackendoff does not. In particular, time travel where "R2" is earlier than "R1". However, we should perhaps take note of our intuition that such cases are at best atypical; perhaps we could place our clause within the scope of a Jackendovian PREFERENCE operator and amend it slightly :

PREF(R1 later than R2)

If this does represent the beginning of a hybrid solution, then clearly the answer to Jackendoff's difficulty is to make the event-state distinction a preference rule rather than a necessary condition. Ontological category would then still depend on the semantics of the lexical head, only now these semantics would be subject to contextual instantiation. We have a "type emergence" approach in which semantic extension is not necessarily hampered by ontological considerations.

There is, however, an alternative analysis. Although sentences like "The road goes..." refer to what are states in the real world, they are conceptualised as events. That is, event-hood is a necessary condition of GO, there is no GO_{Ext} , and when an entity with low motility (like a road) is assigned the trajectory of a motion verb (cf Langacker 1982), then it is

profiled as being in motion. This could be realised as a kind of metonymy - each section of the road is successively profiled; what is moved (conceptually) is the location of the profile (rather than the road). Under this analysis, GO could be viewed as a sort of transfer feature (see below), making it a case of type coercion. This alternative does allow that ontological boundaries limit semantic extension. (There is a clear link between this and linguistic relativity, which is considered further in Dunbar 1988).

Although there has not been time to resolve this dilemma, it has illustrated the issues in some detail, showing that both approaches can be developed. The distinction between coercion and emergence should now be clear enough.

6.2. Vagueness and type emergence

Before proceeding, I will relate the distinction between coercion and emergence to the traditional philosophical distinction between ambiguity and vagueness. When a form is associated with more than one category, it can be either vague or ambiguous with respect to the distinction between them. A good example of vagueness would be the relation between *cup* and the colours "red" and "blue". If I say there is a cup on my desk, then the proposition expressed is simply vague as to whether the cup is red or blue. On the other hand *bank* is ambiguous between a geographical feature and a financial institution.

The relationship between the two distinctions is as follows. First, if a term is vague in some respect (say, colour) then its membership of that type must emerge through, for example, modification by an adjective. Thus, type emergence can issue from vagueness through a process I'll term "precization" (after Pinkal 1983). Type coercion occurs when a new specification overrides a previous one. Thus, for example, (assuming for the moment that *custard* has "yellow" as a specific property) *red custard* exhibits type coercion. Frequently, when a term that is vague in relation to some distinction is regularly and consistently used with a particular value, a specific sense of the term is lexicalised which corresponds to that usage. For example, CANE was originally a "stick", vague as to function. However, it now has a conventional reading in which the function is specified as "aid for walking" (roughly). This conventional specification of additional meaning I term lexicalization. Lexicalization is a pre-condition of psychological ambiguity.

In the next chapter, I will take a detailed look at the mass-count distinction. There are two goals: to press the emergence-coercion distinction hard, and to examine the role of closed class "grammatical" morphemes as a special case of linguistic context. Naturally, it is to be hoped that a satisfying account of the syntax and semantics of the mass-count distinction will emerge. These results will be integrated with those of chapter 4 in a semi-formal model of the interpretation of noun-phrases to be presented in chapter 9.

Chapter 7.

The mass-count distinction in English.

This chapter focusses on the mass-count distinction as expressed in English. As will become clear, this phenomenon illustrates the combination of two lexical items standing in some relation of dependency to one another. The phenomenon also relates to the questions raised in chapter 2 concerning the use of referring expressions to place alternative construals on an entity. It is a good example to tackle because it is relatively knowledge-free in the sense that, at a coarse level of analysis, the semantic alternation involved is constant across lexical fields. This allows the focus on the key issues to remain sharp.

I begin this chapter by discussing the nature of the mass-count distinction. Then I apply tests and criteria, some of which were catalogued in the last chapter, to the mass-count distinction. By doing this I plan to demonstrate that the distinction arises through the interaction of determiners with lexical concepts. I concentrate at first on simple noun-phrases, but later the discussion will decalage to encompass transitivity, aspect and classifier constructions.

7.1. What is it?

In English, there are two sorts of noun (or two ways of using a noun, depending how you look at it): mass nouns and count nouns. Syntactically, the two are associated with complementary sets of determiners. And semantically they are associated with different sorts of referent; by and large the referents of mass nouns tend to exist in the world as globs of liquid or stuff, while counts tend to have "built-in modes, however arbitrary, of dividing their reference...." (Quine 1960:91).

In many grammars of English, this distinction is taken as a syntactic subcategorization of nouns (see also Ziff 1964)). Indeed, as Crangle has noted, entries in published dictionaries implicitly encode a specification in this respect. (Crangle 1984). For example:

Banana A tree... cultivated largely in tropical and subtropical climates.... (OED)

This definition implicitly conveys the information that it's a count noun.³⁴ But is this a property of nouns or of the noun-phrases in which they are used? Study of this will allow us to begin to discover the extent to which the significance of a use of a linguistic sign is partly an emergent property of its combination with its linguistic context. First, however, I wish to make the case for regarding the distinction as semantic rather than merely formal.

For the distinction to be semantic, there must be a systematic difference in meaning between the categories. Firstly, I know of no philosopher having the intuition that there is no difference in meaning; that is, at a reflective level, analysts tend to agree that there is a semantic difference, and there is a rough consensus as to the nature of the difference (although there is disagreement in relation to important details - see below). Secondly, as Allen (1981) has shown, some nouns do not show equal affinity for both mass and count noun phrases. It could be argued that this is a case of syntactically conditioned, even arbitrary, complementary distribution if it wasn't that these patterns correlate with semantic properties. For example, certain superordinate terms (eg. *furniture* and *equipment*) resist appearance in a count noun phrase rather firmly, but such words have in common that their referent is a perceptually heterogeneous set of entities. Thirdly, an argument can be made on the basis of co-predicability patterns (cf Keil 1979, Nunberg 1979, but see also Keil 1979) that the mass-count distinction is sensitive to meaning. For example, count terms but not mass terms can be COUNTED

(53) *Do have three (of) some cheese.

Nunberg (1978) argues that for a case of polysemy to be established (one lexeme with more than one meaning), it is necessary (but not sufficient) that the different meanings be extensionally distinct, but in the case of the mass-count distinction this requirement appears not to be met. To be precise, *an apple* is *apple* but *apple* is not necessarily *an apple*. Thus, at least in the cases felicitously described by count usage, both mass and count terms are applicable. Let me give an example. If I arrive at customs with an edam, it is the case that I have *a cheese*. Nevertheless, should an exciseman enquire whether I have *some cheese* among my things, the answer must be that I do (unless I am a determined

³⁴ Sinclair 1987 argues that dictionaries should do this more systematically.

smuggler). The semantic difference that I have established here is that mass terms lack one entailment that count terms possess.

However, we need not rely on an argument from an extensional difference because (as indicated in earlier chapters) I do not intend to give an extensional account of reference. My position is that the difference is cognitive: these two readings of a noun differ in the way that they cause the noun to interact with knowledge of the world; they differ because they highlight different information by offering alternative perspectives on (or construals of) an entity. Shortly the nature of this difference will be discussed in more detail. However, this general view does avoid some pitfalls of a realist approach.

Some researchers have advocated a realist approach (Cheng, cited in Bunt 1981); let us examine the most obvious difficulties it faces. I have already noted that a single entity can be referred to in either fashion. Assume for the moment, however, that this cannot happen, and that a particular entity is classified as either count or mass. First, there are cross-linguistic and diachronic changes in classification. For example: the French *un meuble* contrasts with English **a furniture*. An example of diachronic change is *dust*, which is typically portrayed as a mass term in contemporary English, but which Jespersen 1914 (p118) claims would once have been more frequent in the form: *a dust*. It is not credible to assert that the entity has changed its essential nature. Secondly, even within English, as McCawley (1975, reported in Gathercole 1985) has pointed out, we get pairs such as *noodles-spaghetti*, *garlic-onions*, and *rice-beans* whose members are ontologically alike yet are differentiated with respect to the mass—count distinction.

Finally, note that abstract nouns such as *honesty* typically fall into the mass-noun class ("The distinction between material and immaterial mass-words seems linguistically irrelevant... [they] take the same quantifiers... [and] may leave the category of mass-words and take the plural form just as other mass-words (kindnesses)." (Jespersen, 1949 p438)). It is hard to see what could be the basis of an ontological class that could span both concrete and abstract nouns, while excluding count nouns.

Against the observation of dual applicability, the realist might argue that, indeed, two different things are being referred to, perhaps calling them the object and the substance (Parsons 1970). Thus, for instance, different speech communities may be focussing on different entities (rather than focussing on different aspects of a single entity). There is a

basic insight shared by both views (ie. that objects are not referentially atomic), but I believe the realist version is untenable. First, it still doesn't handle the problem of abstract nouns (Pelletier (1979), for instance, proposes simply omitting abstract nouns from the field of investigation for this reason). And second, it seems obvious that these "entities" have no existence independently of one another - destroy the gold and you destroy the ring; remove the ring and you remove the gold. In other words, the only coherent way to interpret these "objects" is as aspects of a single entity. I characterise the distinction, then, as an alternation of cognitive perspective.

7.2. Is it lexical?

Having established first that there is a distinction, and having then gone into some detail to characterize it, we are in a position to consider the fundamental question of whether the distinction is lexical or phrasal. I begin by trying to apply some of the criteria that we have seen applied to the phenomena discussed in the previous chapter.

Lyons used the presence of an analogous morphological rule to argue for both coercion and the direction of coercion. From a psychological perspective, we must draw a distinction between diachronic change and synchronic process: although change must, at the time it occurs, involve a psychological dynamic, that dynamic may ossify over a (relatively long) period so that its path is no longer active. Therefore, the argument from morphology can only be suggestive. Furthermore, there are derivational processes in both directions. To derive mass from count, there is the suffix *-age* (eg. *bag* — *baggage*); and in the opposite direction the agentive suffix *-er* can derive a count noun from a mass noun (eg. *milk* — *milker*);

Allerton (*op. cit.*) argued that similarity of the surface forms favoured compound valency (which I have claimed is analogous to valency). Here he means that the occupants of case roles in the event denoted by some verb are realized in the same grammatical relation. This is a difficult analogy to pursue, but mass-count alternation does not affect the basic structure of the NP as far as the order of head and modifier (*some red cheese*; *a red boy*) and so forth goes. In short, the distinction does not create surface complexity, and so this cannot be put against a type-emergence account.

Allerton's other argument concerned the simplicity of the relation. However, like similarity (see chapter 4), simplicity is relative to background knowledge and the only basis we have for assessing it (quantifying it is out of the question) is our intuitive judgement. Nunberg (1979) has introduced the more workable concept of cue validity in this connection. When more than one reading can be given to a term, we naturally want to know whether one sense is more basic, while others are derived from it, and cue validity is a possible criterion for assessing this. Nunberg argues, as I explained in chapter 4, that a term can only designate something which will function efficiently as a cue for derived readings.

Referential functions as described by Nunberg are similar to the construal rules (CR's) discussed by Miller (1978). It is perhaps important to note that this type of rule can only suggest hypotheses to be tested about the interpretation - it cannot algorithmically yield an interpretation. To be more specific, if there is more than one CR or RF that could be applied on a given occasion, there is no way of deciding which should be applied other than trying them out in turn. If there was a way, then the CR's would be redundant.³⁵

The cases in which we are most interested for present purposes, are those in which it is not clear which of the two referential functions has greater cue validity. That is, there are cases in which cue validity appears to be equally high in both directions, and which Nunberg consequently classifies as polysemous. Might it be that this symmetry obtains between the mass and the count senses of nouns? Nunberg mentions that it might, but does not deal with the mass-count distinction at any length.

There are seven cases to consider:

Mass	Count	For example...
basic	idiomatic	<i>iron</i>
basic	lexicalized	<i>cheese</i>
basic	novel	<i>cream</i>
basic	basic	<i>cake</i>
novel	basic	<i>mushroom</i>

³⁵ Lipka 1987 discusses related issues.

lexicalized	basic	<i>cabbage</i>
idiomatic	basic	<i>chicken</i>

Here the term "basic" indicates the unmarked, most usual reading; other terms describe the status of readings derived from this which can be contrasted in terms of the mass-count distinction. Broadly speaking, the referential function involved in deriving a count reading is "unit-of", while derived mass readings are typically "pulp-of".

First, the derivation of count readings. In novel cases, the size of the unit is not necessarily fixed and has to be inferred from background knowledge. Whether the context isolates a unique unit will probably vary from noun to noun and, indeed, from situation to situation. It is instructive to consider the German phrase *ein Brot* which literally translated reads *a bread*. If we assume that the intended interpretation of this nounphrase is as a "natural unit", we can infer its denotation by asking "what is the natural unit of bread?". The answer suggests three possibilities: a roll, a slice or a loaf. In fact, the correct translation would be *a loaf*, but because the phrase is not lexicalized in English we would need to consult either the situational context or a native speaker to determine this. However, even in English, certain units will be lexicalized for certain substances. Which units are lexicalized can depend, of course, either on the laws of nature (such as when *a water* is a body of water)³⁶ or by convention. We can safely assume that the cue validity of some word for a given reading is increased when that reading is lexicalized: if context licenses more than one unit, then preference can be given to lexicalized candidates (unless the speaker has "marked" the phrase). Given the account of lexicalization sketched roughly above, the availability of a lexicalized reading will be relative to the set of currently active theories,

³⁶ Note that the extent of such waters can vary greatly, and that this concept leaves open questions like "where does one water begin and another end?" Such questions arise in philosophy. Pelletier 1979 takes a strict interpretation of Quine's definition of sortals and suggests, for example, that the noun *thing* is non-sortal because its denotation fails to provide criteria for division of reference. (This is one of a group of arguments he puts forward for distinguishing the sortal v. non-sortal distinction from the mass-count distinction). However, the absence of precise criteria for individuation is somewhat beside the point, and will not trouble the less realist concept of reference adopted here. The essence of sortals is countability, and count terms profile an entity as countable, requiring only that individuation be possible in principle on at least a *post hoc* basis. Without this qualification one would be forced to question whether any novel readings of basically mass concepts were sortal.

and so will vary from context to context.

I will assume that symmetrical cue validity is precluded by idiomatization.

The pulping or grinding that derives mass readings from countable concepts seems to be a straightforward mapping, although there may be some doubt as to how many tokens a given heap of pulp could reconstitute. This is not a serious difficulty, however, and matters will be even simpler where lexicalization has occurred.

In fact, the Universal Grinder usually has a filter constructed by background knowledge draped over its outlet. This, for example, will assure the removal of pips, stalks and outer leaves from foodstuffs. At least, such filtering is invariably incorporated in cases of lexicalization. The other major RF deriving mass readings involves conceptual re-orientation, but I will discuss this more fully below.

The minor (less general) RF's display a gradient from lexicalization to idiomatization which leads one to ask whether the structure of a referential function can in itself be responsible for idiomatization. This certainly seems to be so for the RF "artefact typically made-of" which sometimes sets up count senses of mass concepts, particularly when technology moves on leaving the lexicon of the language looking like an industrial museum. Examples include: *iron*; *rubber*; *nylon(s)*; *paper* and *card*. I tabulate here some RF's that might connect a sortal with a mass reading.

Referential Function	For example:
meat-of	chicken; lamb
artefact made-from a part-of	mink; coney; cat [(gut?)] [esp. fur]
substance used to-make	box [?]; pot
produce-from	Wensleydale; port
characteristic performance-of	woman [abusive]; Pele; Dylan
characteristic contents-of	trough; pipe; pot [twice?]

Examination of this kind of data may allow us to make more precise the intuition that an increase in semantic distance increases the likelihood of lexical split (Bybee 1985). For now, however, I observe that a variety of RF's are available, and that whether two readings are related polysemously will depend on which RF relates them, on which of the possible

readings are lexicalized and also on background knowledge and the situational context.

Closely related to simplicity is the notion of productivity, which Lipka adduced (see above). It is an important aspect of the data (noted by most who have considered the topic) that the distinction is not absolute. Frequently, nouns take on different senses according to which determiner has been employed. That is, the distinction is relative to the choice of determiner. This has led many to suggest that the distinction is one between mass and count usage of nouns rather than between mass and count nouns. (see eg Pelletier 1979).

This is highly productive, and we have already seen the construal schemata involved (although they will be discussed in more detail below). Pelletier (*op. cit.*) has argued, in relation to his Universal Grinder, that this alternation is possible for any concrete noun. The small set of exceptions, as I have noted, characteristically have a heterogeneous collection as their denotation. Pelletier quotes Gleason to the effect that if alternative readings sound odd it's just a matter of securing the right context to render them plausible. This is reminiscent of the discussion of *ad hoc* categories in chapter 4.

One feature of productive rules is semantic predictability (ie. compositionality). In their discussion of aspect, Miller and Johnson-Laird (1976) argue that verbs should be subclassified for aspectual distinctions because of an interaction between different verbs and formal markers of aspect:

"The effect of the progressive... depends on the time schema of the verb it is applied to." (p443)

By and large there are no such interactions between nouns and, for example, different determiners. Here are some examples:

"count" concept	a boy	some boy
"mass" concept	a jam	some jam
"count"	an onion	some onion
"mass"	a garlic	some garlic

The meaning of each determiner is obviously the same no matter which noun it appears with. (This criterion corresponds roughly to Zwicky and Sadock's test of inconstancy under substitution and is also used by Cartwright 1975).

Obviously, we must take into account lexicalization - the conventionalization of a sense or form that was initially derived by some active process (Bybee 1985, Langacker 1982b). One assumes that there is a culturally conventional concept to which the phrase *a cheese*, for example, alludes, and which does not require construction *de novo* so much as recollection. (Details of this will be discussed in a later chapter). A frequent corollary of lexicalization is idiomatization - a radical drift in meaning - and this is found in cases such as *an iron* [the domestic appliance] and *a glass* [the drinking vessel] (see Close 1975). Nevertheless, even in such instances the corresponding "mass" noun (*iron, glass*) can be used in a count term (*an iron, a glass*) denoting a unit of the substance if the situational context licenses it (for example, think of a foundry). Indeed, these new uses may also be lexicalized - although it's a bit archaic *a glass* can denote a telescope, for example. This underlines the productivity of the alternation, and it would seem that if mass and count readings are distinguished lexically, it is necessary to provide a shadow lexical entry for every noun. Given the existence of other (orthogonal) sorts of polysemy (such as type-token, see chapter 4), there will be a rapid proliferation of shadow lexicons unless we follow Nunberg in confining each family of designata to a single lexical item.³⁷

One hypothesis (the Shorthand Expression hypothesis; the phrase is Clark 1983's) is that the derivation of novel readings depends on the elision of material from an "underlying" partitive. Thus, *a wine* is merely the surface realization of, for example, *a glass of wine*; and a sentence like:

(54) John ordered a large coke, but the waitress spilled some.

has the words *of it* elided from its tail. That is, there would be a single reading from which a novel reading could be derived. While clearly ellipsis is possible, and is probably the most favoured reading of the sentence above, it depends upon the presence of an antecedent and so cannot underly all cases. In particular, the choice of classifier is context-dependent (Bunt 1981), different classifiers being appropriate for different substances. Furthermore, in different situations, different classifiers may be required for the same substance (Ware 1975): *a beer* may refer to "a pint", "a keg" or "a barrel" depending on whether it's being ordered for lunch, for a party or for the cellar of a pub. It could be argued that the

³⁷ Nunberg (1981) doesn't take a specific stand on the mass-count distinction, suggesting only that it may be a case of polysemy.

underlying string is a *CLASSIFIER of substance*, with the classifier being instantiated by context, but in the absence of specific evidence for the proposed abstract "archimorpheme", this will turn out to be a notational variant of the account to be proposed here. For cases of true ellipsis (ie. those involving anaphora), a separate mechanism will be proposed below (although I will not be concerned with anaphora *per se*).

Zwicky and Sadock (1975) catalogue a range of tests of the ambiguity-vagueness distinction. Two of their semantic criteria have already been discussed: "plausibility of systematic ambiguity" (being heightened by the formal marking of the meaning distinction in some other language) corresponds to the discussion of morphology. "Plausibility of lack of specification" corresponds to the earlier discussion of simplicity. Some of the tests, such as the criteria for establishing the occurrence of "the intersection of patterns", and tests for changes in transformational potential, are not relevant (except to the extent that they cover the same ground as Allerton's point about similarity of form). They deal at length, however, with some special cases of tests of transformational potential - tests for the applicability of rules which require identity of *sense*. These tests are pertinent and deserve close examination. Put at its simplest, these tests require that a single occurrence of a given form should be capable of being given both readings in a single utterance without being physically repeated.³⁸

Using conjunction reduction we can first establish that mass and count terms are of broadly the same type:

(55) Sam wrapped some cheese and an apple for Mrs Goggins.

cf.

(56) *Sam wrapped quickly and some cheese for Mrs Goggins.

Having done this initial screening, we can apply some more stringent tests based on the material shared ie. recovered in one conjunct from appearance in the other. Here, identity

³⁸ For the purpose of these tests, I define the following symbols: *c and ?c. They indicate intuitive uninterpretability of the sentences they precede, on a **crossed** reading of the antecedent/residue, where a crossed reading is one in which the contrasted interpretations are both given. Note that although these tests are not restricted to one theory of grammar (TG), their interpretation is heavily theory-dependent, as we shall see.

of sense at a more specific level is required. A wide range of tests is available, the main differences among them residing in the syntactic properties of the material elided. However, all deletion tests (which I consider first) give similar results.

First, some tests in which the ellided material does not constitute a syntactic constituent: VP deletion and gapping: (some examples have preceding context to clarify the intended readings)

VP deletion:

(Fred has an iron to mend, while Derek has a sack of iron to sell).

(57) *c Fred had insured his iron, but Derek hasn't.

(Danny loves cheese and Robin often buys an edam).

(58) *c Danny might share his cheese with you and Robin always will.

(Stan's favourite dish is a plateful of fried mushrooms, while Ollie prefers a (single) stuffed mushroom).

(59) *c Ollie would finish his mushroom but Stan might not.

Gapping:

(60) *c Brad read his paper in the shop and Scott on the train.

(61) *c Leslie ate her cheese at home and Sally on the train.

(62) *c Derek spilled his cream in a cafe and Simon in the manse.

As I have indicated, my intuitions suggest that neither construction permits crossed readings, whether the marked reading is idiomatic, merely lexicalized or neither. This suggests

that the mass-count distinction is an ambiguity rather than a vagueness.

In the following tests, the shared material is a constituent:

Conjunction reduction:

(63) *c Jane unplugged and Lucy smelted her own iron.

(I asked Jim for a pint of No.3 and Dave for some beer.)

(64) ?c Jim, with dignity, and Dave, brusquely, refused to buy the beer I requested.

(65) ?c Sam and Fred each tossed down their cheese.

(I sold Bill a double-decker and Tam a kilo of scrap.)

(66) ?c Bill and Bob (each) sold their bus for a profit.

Again, idiomatization is an obstacle to reduction. In other cases, however, judgements are, perhaps, less clear. The next example involves "extraction":

Extraposition:

(67) *c It was the glass they hated that Jack and Alan (each) stared through.

(68) *c It is the cake their mother makes that every child prefers.

(69) *c It is only the most succulent tomato that the chef will serve or the customer will eat.

The following constructions employ still more exotic relations between the two readings of

the shared material:³⁹

Same thing:

(70) *c Susan threw her rubber in the air and Joan did the same thing.

(71) *c Jonah spat out his fish and Peter did the same thing.

(72) *c Dex admired his own table and Bobby did the same thing.

Vice-versa:

(73) *c Tom gave Jim the tin he owed him, and vice-versa.

(74) *c Captain Hook gave Smee the coffee he wanted, and vice-versa.

(75) *c Tony spilled Richard's shampoo, and vice-versa.

The evidence of identity of sense tests examined so far points strongly in favour, then, of ambiguity. However, these tests have all involved deletion and Zwicky and Sadock (*op. cit.*) point out that constructions in which a proform remains as a relic of the ellided material don't inhibit crossed interpretations so strongly. They suggest that in some cases it may be what they classify as the perceptual difficulty of assigning two readings to a single lexical token that impedes such interpretations.

³⁹ Semantically, these constructions are polar opposites. They seem to utilise a propositional pro-form, and so represent an intermediate stage between the previous tests involving deletion and those involving pro-formation with a successively narrower domain which follow.

One such test involves the pro-form *so does* which is a VP placeholder. For consistency, I will adapt the examples used for VP deletion above.

(Fred has an iron to mend, while Derek has a sack of iron to sell).

(76) *c Fred had insured his iron, but Derek never does /and so had Derek.

(Danny loves cheese and Robin often buys an edam).

(77) *c Danny shares his cheese with you and so does Robin.

(Stan's favourite dish is a plateful of fried mushrooms, while Ollie prefers a (single) stuffed mushroom).

(78) *c Ollie always finishes his mushroom and so does Stan.

(Ben has a glass of water in the morning and Stephen has lots of water in the evening).

(79) *c Ben drinks his water too quickly and so does Stephen.

This still points towards ambiguity, and any sense that these examples are "better" than those of deletion involves a fine judgement. It might be that it is the absence of a pronominal in particular that is continuing to obstruct the crossed readings, and so I consider next a test of precisely this hypothesis.

A related criterion, more applicable to NP's, is that if a word is vague in relation to some semantic distinction, then it must be possible to erect an anaphor-antecedent relation between the two readings. (Bolinger, 1973 makes a similar point; the test is also implicit in Jespersen, 1914 p99n and has been advocated by Frei (1961, reported in Coseiru & Geckeler, 1981; Nunberg 1979 also uses it). First, as one might expect, this is not possible where idiomatization has occurred:

(80) *c John read a paper, while Derek drew on some.

- (81) *c Simon makes horseshoes with iron, while Nigel presses his clothes with one.

In such cases, it's hard to escape the sense of punning (cf. Nunberg (*op. cit.*)). In the more general case in which idiomatization has not occurred, instances in which the form of the anaphor marks the type of reading can be distinguished from those in which it does not. There are two strategies which can be followed, because now there is a choice of either using pronominals which do disambiguate/specify the appropriate reading, or of using other pronominals which do not achieve this.

Let's consider first the latter strategy, not using the pro-form to clarify the reading. Pragmatic considerations will prevent a crossed reading if anaphora involving identity of reference is employed. One way to avoid the identification of anaphor with antecedent is to use possessives, which facilitate other readings; consider the following examples:⁴⁰

- (82) *c I won't want your cheese when I've got Samantha's.

- (83) *c The tea in this pot will be cold when mine arrives.

Neither crossed reading seems to me to be alright here.

When it comes to pronominals which do mark the mass-count distinction, it is of course again important to avoid an identity of reference reading, and such readings are prone to arise through the postulation of an underlying partitive.⁴¹ This is demonstrated in the following example:

- (84) He bought a large cheese, and Margie sent some [of it] to Cairo.

A way to avoid the difficulty about possible underlying partitives is to construct the context so that, if there is an underlying partitive, its kernel can't be identified with the antecedent. For example:

(My local Gourmet Club has a tasting of one type of food each month).

⁴⁰ In fact, the examples involving deletion rely on this device too.

⁴¹ This option was obviously also impossible in the examples involving deletion, since only one surface form was available.

- (85) Last month I took along a stilton, but the Secretary said there was already plenty.

Here for the first time, a mass reading of an anaphor founded upon a count antecedent does no violence to my (admittedly somewhat numb) intuitions. The partitive reading is not possible because the two quantities of cheese are explicitly disjoint. The only *it* that there could be *plenty of* would be the mass term *stilton*. It cannot be an implicit plural because of the presence of singular *was*. In contrast, the analogous construction for a count anaphor is downright ungrammatical:

- (86) * Last month I took along some stilton, but I was told there were already plenty.

- (87) * Last month I took along some tea, but the Secretary said there were already plenty.

Another way to yield an explicitly count anaphor is to use a spatial superlative to modify the anaphor. Thus:

- (88) *c Each year there is a competition to find the tastiest cheese in the county, and Jamie's biggest always tastes best.

- (89) ?c There was a cheese competition and Dod entered the roundest one.

The result, then, is that although pronominal anaphora is slightly more favourable than other tests of identity of sense, such criteria tend to indicate that the mass-count distinction is an ambiguity. This conflicts with the other criteria discussed, and I will briefly consider which should be given more weight.

Part of the difficulty here lies in the semantic relationship between mass and count readings, which is similar to privative oppositeness. As with privative opposites, what appear to be contradictions on the surface turn out to be interpretable:

- (90) That cheese is not a cheese.

This semantic difference was noted above, albeit in different terms. It means, however, that it is always possible to interpret such sentences by assigning a mass interpretation throughout, and so problems with interpretation will not be manifest. Zwicky and Sadock

(1975) suggest that tests based on identity of sense must be void in relation to privative opposites because, they argue, it will not be possible to disambiguate the readings and determine whether a crossed interpretation works. The privative oppositeness creates severe difficulties in making interpretability judgements and, more particularly, in supporting one's judgement with disambiguated examples. I have endeavoured to contextualize the trickier examples, and in the case of pronominal anaphora a formal way of specifying the second reading was found. Nevertheless, the applicability of identity of sense criteria is in some doubt.

Secondly, it should be pointed out that some of Nunberg's instances of polysemy fail some of these tests too. Take the use-mention distinction, for example:

(91) ?c Lager looks like beer and so does beef.

cf.

(92) Lager looks like beer and *beef* looks like *beer*.

Thirdly, and crucially, it is not clear whether such criteria demand "identity of sense" at a conceptual level or at the level of a lexical concept - they are only relevant if it is the latter. It may be, for example, that this is the difference between deletion and anaphora: deletion requires two propositions to intersect around an instantiated concept, while anaphora allows them to share a lexical concept. One might speculate, then, that if crossed interpretation is alright for anaphora but not for coordination reduction, these constructions differ semantically in the number of positions they make available for the conceptual realization of entities denoted. This, however, goes beyond theories of the mental modelling of anaphora, such as Johnson-Laird 1983. In particular, it indicates that an anaphor doesn't point back to a conceptual representation (analogous to a token) of a previously introduced entity, but to some intermediate representation from which a new token can be generated. Coordination reduction, on the other hand, would oblige two propositions to share a single concept. Zwicky and Sadock (*op. cit.*) point out this difference between the classes of construction and attest to its generality. They suggest that it may be due to the perceptual difficulty of having to give two readings to one form. To paraphrase, according to the perceptual hypothesis, the crossed reading is neckered. Either analysis is, of course, speculative, and more extensive work involving experimentation would be required to take it any further. Nevertheless, the conceptual hypothesis forces consideration of the different levels at which

a word can be said to be vague, for I have assumed that vagueness in the lexical concept does not entail vagueness at the conceptual level.

The difficulty is that theoretical understanding of these phenomena is not sufficiently advanced to settle such questions at present. In particular, it would be useful to be able to place a theoretically motivated bound on the basis of identity of anaphor and antecedent. Zwicky and Sadock observe that they must be equally sincere (ie. both sincere or both insincere).⁴² It is furthermore obvious that phonological identity is insufficient and we can speculate that referential identity represents an ideal, but it would be a great advantage to have a theory underpinning such data.

One hypothesis would be that differences in meaning that impinge upon the "scope of predication" (Langacker 1982b) of the antecedent, or which are "relevant" (in the sense of Bybee 1985) to the denotation of the antecedent, will rule out identity of sense, at least in relation to the parameter of meaning they affect. However, I'm not sure that there are any questions such a formulation doesn't beg.

I have suggested that pronominalization is based on lexical concepts rather than instantiated ones, and evidence for this comes from the nature of the relation between anaphor and antecedent after instantiation of both in cases of lazy pronominalization. Classic examples are like this:

(93) Every man who gets a paycheck beats his wife with it.

and involve a token-token indeterminacy: different *paycheck* tokens and different *wives* for the different *men*. This doesn't take us too far in itself because, assuming for the moment the prototype theory of concepts, a concept may be vague in relation to different tokens - it represents a type. That is, this form of indeterminacy doesn't let us differentiate concepts from lexical concepts. However, some of Nunberg's cases provide the evidence required, because they imply a type-token distinction between anaphor and antecedent which, as we saw in an earlier chapter, precludes prototypes at least. For example:

⁴² They base this on some observations of surprising perlocutionary identity. Sincerity should be taken very broadly in this context.

- (94) That flower [token] didn't grow in your garden because it [type] requires lots of sunshine.

Before leaving the identity of sense criteria, there is one more class of example that should be examined. In this example context is used to clarify the sense of both antecedent and anaphor, but now the pronoun itself does not mark the distinction.

(There was only one left. I asked the detective if he wanted it).

- (95) He rolled me the cream, and it whitened my coffee.

In this case *the cream* strongly suggests a small catering pack, but the *it* that whitens can only be the mass of cream. An instance involving deletion is:

- (96) It was my mouse that Robert squashed and David scraped up.

Indeed, these are rather special examples, because they involve identity of reference in the traditional sense that the same real world object is referred to. The account of reference I have adopted, however, has no difficulty with this, because it allows that the same entity may be picked out differently by different phrases: the two referents are two distinct cognitive perspectives on a single entity.

Overall, then, a variety of criteria have been examined, and some have given clearer indications than others. I have argued that criteria like productivity favour treating the distinction as lexically vague. On the other hand, identity of sense criteria motivate a diagnosis of ambiguity, particularly in cases involving ellipsis as opposed to anaphora. However, I have suggested that the identity of sense criteria compare noun-phrases rather than nouns. This implies that noun-phrases that are indeterminate with respect to the mass-count distinction are to that extent ambiguous, whereas nouns should be regarded as vague. At the very least, there is no strong evidence against the proposition that the mass-count distinction is an emergent property of phrases.

Indeed, if we assume that noun-phrases correspond to concepts, and assume for the moment that a concept, as opposed to a lexical concept, can be characterised as a prototype, it follows that noun-phrases must be specified as either mass or count, at least implicitly. If, for example, the concept is a prototype of a basic level concrete noun, it will make the distinction because the prototypes of such nouns encode information about shape,

and anything construed as having a characteristic shape is going to be countable.

In addition, I have argued that this polysemy among readings of nouns may hold only for a subset of the possible referential functions relating them. That is, some mass-count transformations may involve such a large cognitive step that lexicalization is inevitable.

Chapter 8.

A phrasal approach.

So far I have argued that evidence derived from ambiguity tests implies that nounphrases are ambiguous while nouns are vague with respect to the mass-count distinction. This means that the raw noun is parameterized (in the sense of chapter 6) for mass-count-hood and suggests that the closed class words, the determiners, specify the classification of a given noun-phrase.

There is a consensus along similar lines in the philosophical literature. For example, Cartwright 1975 says that the distinction is not lexical but rather a property of:

"...the grammatical structure of the contexts in which [the words] occur."

Bunt 1985 states that "nouns *per se* cannot be classified as count or mass" (p11).

The evidence of lexicalisation and idiomatisation shows that sometimes the determiner acts like a flag, indicating either one familiar concept or another. For example,

(97) some iron

picks out the metal while

(98) an iron

typically picks out the household appliance.

On the other hand, phenomena such as productivity and re-lexicalization demonstrate that sometimes the determiner acts as a tool in the construction of a new meaning, or concept.

Having argued that being "mass" or "count" is an emergent property of phrases involving nouns rather than of nouns themselves, there falls an obligation to provide an account of the process of emergence.

Carlson (1983) considers alternative ways of providing a semantics for a phrase structure tree. One of these alternatives ("propagating nodes") will be a convenient analogy to employ. The idea is in fact very simple: at certain nodes in the phrase structure tree of a syntactic analysis features will gather together, and so if the rules of interpretation are directed towards these "propagating nodes", these rules can take account of the properties of several constituents when prescribing the correct interpretation of the phrase.

There are two reasons to focus on the closed class words that appear in noun-phrases. First, there is a *prima facie* correlation between sets of these words and readings. Often in the literature (as I mentioned at the outset) the correlation is described as a syntactic index of the mass (or count) status of a noun: a noun will occur with one set of items if it is mass; the other if it is count. I propose instead to take the view that the meaning of the closed class words determines the reading. (A comprehensive table of acceptable combinations is provided on pp127-128 of (Close 1975)). For present purposes I will adopt the idealization that an NP corresponds to a use of a noun and denotes a concept, while the noun itself corresponds to just a lexical concept.

The second reason for focussing on these closed class words is syntactic. Their dependency relation with the head noun is, according to Allerton (1982) the same as the semi-subordinative relation between a verb and its direct object, the potency of whose modulating effect was demonstrated above. Allerton cites the following criteria in support of this analysis: determiner and noun agree in number; only a few nouns can appear without a determiner. Thus the determiner is essential to phrasality and plays a "vital ancillary role" (Allerton 1982:13).

Before analysing the relevant aspects of the meaning of these words, a couple of general remarks on the distinction between open and closed class words. This is not an exhaustive list of differences. For specific criteria and discussion of related distinctions see eg. Matthews 1981. The purpose of these remarks is merely to locate what follows in terms of some broader issues.

First, the role of closed class words in comprehension is controversial. Bradley 1983 has argued that they may be recognised by different mechanisms (adducing evidence concerning differences in frequency effects) and speculates that they may have a crucial role in the perception of the syntactic organization of language. This is supported by their linguistic behaviour (often they delimit

constituents and sometimes they are effectively case-markers, for example) and by evidence from aphasia and acquisition data (Bates et al 1983). However, Bard 1987 points out that closed class words are not acoustically salient, as one might have predicted a perceptual anchor would be. Furthermore, Gordon and Caramazza (1985) present evidence that some of Bradley's results may have been due to the narrow frequency range of the closed class words in Bradley's materials.

The second observation is that such words are semantically subdued relative to open class words. Here I draw heavily on Bybee 1985 who argues that the formal realization of semantic distinctions ranges along a spectrum from lexical expression to derivational and inflectional morphology, to clitics and finally syntactic expression. She provides evidence that, for example, the evolution from derivational to inflectional form involves both phonological and semantic lenition. While Bybee does not deal with closed class words specifically, it makes sense to introduce closed class words at some point on this continuum (roughly between derivational and lexical).⁴³ One prediction that can be borne out is that open class items are sometimes semantically subdued. Indeed, this is precisely what was argued in chapter 2 in relation to semantic illusions and also the use of marked constructions to achieve foregrounding. Note that foregrounding *Moses* reduced the number of illusions by about half in Erickson and Mattson 1981's experiment, and that background items (and given information generally) tend to be phonologically reduced. This is one of the key phenomena I set out to account for.

There are two ways in which the subdued semantics of closed and open class words can be interrelated. First, the former represents a conventionalization of instances of the latter. And second, closed class words that enter into certain constructions with lexemes can be viewed as instruments of the precization or enrichment of the meaning of the lexemes. Specifically, I intend to adopt the position that the determiner (or quantifier etc.) **actualizes** the head noun of a noun-phrase. Hewson 1972 provides an analysis of definiteness and indefiniteness carried out within the framework of the writings of Guillaume. The characterization of the general relationship between the meaning of nouns and the meaning of

⁴³ I am grateful to Lorraine Tyler for pointing out the parallel with Bybee's work. In fact, closed class words assume roles that range from relative fullness (*he went to Lisbon*) to the inflectional (*he read the message to Tony*). Moreover, productivity and "relevance" are said to be characteristic of inflections.

articles that is presented there is an insight fundamental to the analysis here.

Hewson attributes to the definite and indefinite articles a kinetic value and "[this] movement which is the sole content of the articles operate[s] as a clarifier of the notion to which it is attached" p75. For example, the notion of the indefinite article is oriented from the general and universal towards the particular and singular. Hewson's claim is that:

the bare noun, calling into play as it does the great extensivity of notion belonging to the potential significate, presents the limitless, formless, continue entity; when this vague representation must be limited, reduced, clarified, defined, refined, restricted, the article comes into play. (p94)

He argues, further, that demonstratives, possessives, quantifiers and the plural morpheme all stand in a similar relationship to the noun. His position, moreover, is that the mass-count distinction is rooted in precisely this relationship.⁴⁴

In English, the noun without article does not represent a mere idea, a total abstraction. It may represent, in fact, a concrete reality, but a reality without clarifying exterior form, a mass-word or continue. Add an article and the concept is given form and becomes a thing-word or class-word; the threshold [between use and non-use of an article] in English lies between the presentation of the notion as a formless, non-numerical entity, and its presentation as a separate singular entity, member of a class and necessarily having form a class word is merely the mass word defined and given form by the use of the article. (p77)

The central theme that it is important to take from Hewson 1972 is the process of "notional ideation" (p65) (the "operation de discernement") whereby a particular notion is discerned among the many that present themselves. The determiners and the plural morpheme can be viewed, then, as affecting notional ideation in a manner that can resolve cases of polysemy such as those under discussion.⁴⁵ (Weinreich 1966) discusses a similar approach, and what he said will be examined in some detail since it is close to the account that will ultimately

⁴⁴ In English - Hewson claims that languages differ in respect of the extensivity of their nouns. For example, French needs articles even for mass words because of the greater extensivity of the noun in modern French. By extensivity, Hewson appears to mean abstractness; the more extensive a concept is, the less it is tied to a particular concrete reality.

⁴⁵ The approach to be taken here viewed as being in the spirit not only of notional ideation, but also of Weinreich's (1966) transfer features, since it affects the semantic interpretation of the noun with which it is in construction.

be adopted here.

First of all, Weinreich points out the flexibility of adult language: we cannot simply outlaw certain article-noun combinations, because there are readings on which examples like *a water* or *some pillow* are acceptable.

Secondly, he argues against the idea of creating new space in the lexicon for *water*#2 etc. on the grounds of economy and points out that in general, one sense is more basic. Regarding economy, Nunberg (1979) argues that the size of the lexical list is not of general theoretical importance. Regarding markedness, Weinreich's approach in deriving one sense from the others differs from that of Nunberg who argues that the different senses of a polysemous word have equal cue validity. Here I take a middle course that has already been outlined: whether there are distinct lexical entries will depend on whether specific senses have been lexicalized. For some words there will be two entries, for others just one. However, the possibility for further creation of readings always remains. This can be demonstrated in some relatively rare instances where more than one derived sense is idiomatized. For example:

iron: golf club; domestic appliance; equestrian equipment

pot: liquor; earthenware; marijuana(?)

These examples show that this is true for the derivation of both mass and count readings.

Thirdly, Weinreich considers marking determiners for count status and marking the noun in any nounphrase as being in concord for this feature, but argues that even though determiners such as *the*, *this*, *any* and *my* are neutral with respect to the mass-count distinction, some nouns appear to be more readily interpreted as mass than count and vice versa when in construction with members of this subset, for example:

the blood, mass;

the flood, count

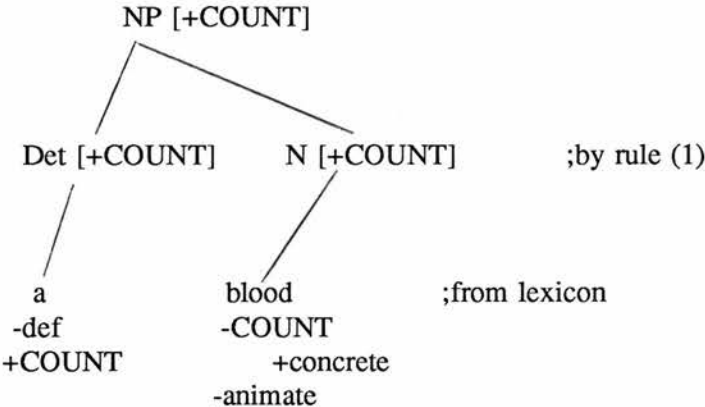
On this point, we should bear in mind Crain & Steedman (1985)'s counsel to be wary of judging what interpretation would be assigned to a potentially ambiguous piece of data in a hypothetical neutral context. The presupposition carried by the definite article - that the nounphrase can be matched with some previously mentioned entity - can only exacerbate any such problem with these examples: it's less clear that we prefer *this blood* to be mass.

This point is significant, and is discussed more fully later on.

Weinreich’s eventual solution is to introduce the feature COUNT as a property of the NP as well as marking it in the lexical entries for both determiners and nouns. The rewrite rule for a nounphrase becomes:

$$NP [@COUNT] \rightarrow Det[@COUNT] + N[@COUNT]$$

This will lead to double-valuing of the feature COUNT on those lexical items that are marked for this feature in the lexicon, which we might feel is rather untidy.⁴⁶ For example:



And in cases such as the above in which the values are contradictory the so-called Semantic Calculator will have to construe an interpretation. Weinreich gives no detailed explanation of how it would do this. Nevertheless, he is satisfied that as a first approximation, "this accounts for both the interpretability and the oddity of 'a blood'"(p437).

The analysis which follows is an attempt to overcome difficulties with the rules proposed by Weinreich. This will be a syntactico-semantic analysis. Firstly, some grammatical features are introduced and evidence for postulating these features is marshalled.

⁴⁶ This feeling of untidiness can be expressed in more precise terms. The extension of N achieved by unification in this example is not a legitimate (partial) function from features to values (Gazdar et al. 1985).

8.1. Delimitation

First I isolate a feature that will play an important role: DELIMITATION. Let me emphasise that this feature does not exhaust the meaning of the words it is associated with: it deals only with only one aspect of a plurifunctional group of items (Karmiloff-Smith 1979) This feature is essentially associated with the category of determiner (and predeterminer) but its value has consequences for the interpretation of phrases which dominate it. Its role in the noun-phrase of which it is a constituent will be the principal concern for the moment, however. The notion is whether or not the determiner places a boundary on the quantity of the referent of the noun. Examples of determiners which do are: *all* and *a*.

The notion of boundedness is linked in the literature of traditional linguistics to the mass-count distinction. For example, consider Close's (1975) observation that "concrete objects having shape" (p109) are prominent among those typically used as count nouns. And again, Jespersen 1914:

"... we have a great many words which represent 'uncountables', that is, which do not call up the idea of any definite thing, having a certain shape or precise limits (p114)."

However, it would not be correct to say that DELIMITATION entails a count reading or vice-versa. For example, the non-numerical quantifier *all* appears in some noun phrases (such as *all bread*) which are given a mass interpretation. Conversely, the phrase *some girls* must receive a count interpretation, even though *some* is [-DELIMITATION]. This analysis of *some* follows Jespersen (1949) who describes it as follows.⁴⁷

"*Some* is the pronoun of unspecified quantity, denoting an unknown or unspecified amount (before mass-words), number (before plurals), person or thing (before singular countables)." (p608)

It is as if the plural is undelimited at a higher level:

"Plurals are semantically related to mass-words. Both of these in themselves denote an indefinite quantity, the difference being that mass-words denote uncountable, plurals countable quantities.... [The plural] denotes all members of the genus

⁴⁷ This quote and the next one from Jespersen's grammar are in fact due to Neils Hailsund, who prepared the final volume for publication (writing some of the chapters himself) following Jespersen's death.

rather than the genus as a whole."(p442) (Jespersen 1949)

Many linguists have commented on the semantic similarities between plural and mass nouns. On this analysis, I can offer the following generalisation: some plurals share with typical mass nouns the property of being undelimited in their main profile. In the case of mass nouns, it is a lack of delimitation at the level of the individual. For plurals, the lack of delimitation occurs at the level of the ensemble (Bunt 1985).

Zemach 1979 enumerates four ontological frameworks (ie. fundamental ways of categorizing entities in the world). He argues that all are philosophically adequate and, indeed, that they are not mutually exclusive. The ontologies have in common that they characterize the domain of existence as the four dimensions of space and time. They differ regarding the number of dimensions along which entities are bounded. For example, the ontology that characterizes everything as bounded in space and time will describe material things as "lazy processes", whereas the one that characterises them as bound in space but continuous in time leads to common sense conclusions such as "the definition of a pin specifies that whatever is a pin must have a certain characteristic *spatial* shape, but it says nothing about the kind of career a pin should have...." (p67). The thrust of Zemach's paper is that each of the four ontologies is useful for different kinds of entity. What should be noted for our present purposes is the crucial role of boundedness - it lies at the very heart of Zemach's system of ontological distinctions. If we accept this system as an account (perhaps only partial) of the fundamental matrices of cognition within which human beings must conceptualize their world, then we will possess an explanation of the feature DELIMITATION — it occurs in the grammar because it is a cognitive primitive (primitive in the sense of Carey (1982))

The role of ontological knowledge in lexical acquisition has become a hot issue in recent years (see, for example, Keil 1984) and both Carey (Carey 1983) and Keil (Keil 1979, Keil 1983) indicate that they regard "boundedness" as a strong candidate for the status of ontological primitive. Keil has argued that children use such knowledge to infer further properties about entities to which some predicate has been applied. The inferences are based on the premise that an entity of which a given predicate has been used must belong to the ontological category which that predicate spans, and are legitimated by the powerful tendency for ontological categories to form a strict hierarchy. Carey (1983), however, uses the example of the cross-classifiability of entities along the dimensions of "boundedness" and

"materiality" ⁴⁸ to demonstrate that the ontological categories are not confined to a strict hierarchy. She argues instead that ontological constraints on the induction of word meaning be explained in terms of the scientific theories held (by the child) at the time of induction: predicates are appropriate to an entity if they enter the same theoretical domain. The scientific theories would consist of core and peripheral concepts expressing the relevant phenomena and causal laws (or mechanisms), the core concepts being, in Carey's view, the best candidates for representation at the ontological level.

Research, then, on assessing the role of ontological knowledge in the development of word meaning is still at an early stage. Nevertheless, it does not seem to contradict the hypothesis that the feature DELIMITATION is, in fact, the grammatical reflex of an ontological primitive. If this hypothesis is indeed correct, then it will come as no surprise, as suggested by the work of both Zemach (1979) and Carey (1983), to find that DELIMITATION is applicable to more than one type of entity.

There is a close relationship between delimitation and definiteness. It is worth noting that in the singular, definite articles are quite neutral with respect to the mass-count distinction. It appears that the reading should be "inherited" from the member of a shared discourse set which is being referred to (cf Hawkins 1978). Jackendoff's (Jackendoff 1978) observation that stuff can be referred to by the pronoun "it" in the same way as countable entities is relevant in this connection.

Hawkins (1978) has analysed definiteness in terms of two aspects: location and inclusion. Roughly, location requires of definite reference that it refer to an entity which is either present in some shared discourse set (or, as Langacker (1982b) puts it, "in the shared epistemic space of speaker and hearer" (p48)) or else explicitly linked with an entity in some such set. In this connection, it should be observed that possessive pronouns are like definite articles in as much as when the head noun is singular the noun phrase is indeterminate with regard to the mass-count distinction. Following Hawkins (ibid.) we will adopt the notion that the two serve similar functions as what might be termed relational matching filters. In procedural terms, both presuppose that the head noun can be linked with some

⁴⁸ For example, a *shadow* is immaterial but bounded, while *gold* is concrete but unbounded; conversely, a *table* is both concrete and bounded while the *sky* is neither.

entity in a shared discourse set and that in the case of the definite article the relation this entity has with another (that must be explicitly contained in the relevant discourse set) may be one of identity, while with the possessive it is nominally a relation of possession (in fact the range of relations which possessive pronouns may cover is extremely wide). The hearer is "instructed" to match the referent of the definite noun phrase with the related entity, using the description in the rest of the noun phrase to filter out candidates. Frequently such reference is used to add properties to the description of the entity picked out, and meeting this end may come into tension with obeying the Gricean maxim of Quantity.

Inclusiveness refers to the requirement (especially relevant in the case of plural and mass reference) that the entire set of objects satisfying a definite description are to be taken as having been referred to. It is important to note that what is necessarily included is the relevant entity or set of entities, rather than all those known about. That is, inclusiveness is always relative to some shared discourse set.

What, then, is the relationship between DELIMITATION and Hawkins' "inclusion"? It seems reasonable to suggest that it is because definite reference is inclusive in the sense described above that it is [+DELIMITATION]. That is, because the meaning of the definite article requires the whole of the referent, the referent's linguistic representation must be [+DELIMITATION].

8.2. Number in English

In this section, the grammatical category of number is analysed. This is done with specific reference to English in which the principal distinction is between singular and plural. This will be marked by the binary feature SINGULAR.

It is possible, however, to go one step further in the analysis of the plural morpheme, and to identify two syntactico-semantic component features. The first is INDIVIDUATION: when the noun is plural the noun phrase must be interpreted as countable. The second is MULTIPLICITY - both are binary.

The association between plural and semantic individuation is observed by (Dekeyser et al. 1979) who describe collective nouns as "individuated" (p125) when paired with a plural verb. For example:

(99) the family are deeply shocked

emphasises that each member is shocked, while:

(100) The family is deeply shocked.

focuses more on the collective shock of the family. That is, the singular verb portrays the people in the family more strongly in terms of their role as family members.

Palmer 1976 notes the following interesting case of a language in which the "plural" morpheme is clearly a marker of individuation:

"In Tigre there are many mass nouns which have a singulative (individuating) form made by a suffix, eg *neheb* "bees", but *nehbat* "a bee". But the singulative form is the form used with all numerals - not merely 'one' [eg] *hatte nehbat* "one bee", *satas nehbat* "three bees" etc. What seems to be important is not plurality, but individuation."(p125)

That is, in Tigre, the morpheme is neutral with respect to MULTIPLICITY — it is a sort of Trubetzkoyan archi-plural.

While noting the individuating function of the plural, Jespersen (1914) also points out a few exceptions such as *brains*. In the case of *brains* and similar examples what we seem to have is neutralization of the INDIVIDUATION. One might speculate that the availability and neutralizability of these components of plurality are universals of some kind.

8.3. Relating the primitives to the mass-count distinction

In the next chapter I associate combinations of these properties with particular determiners. I wish, however, to first show how these features relate to the mass-count distinction drawn in traditional grammar.

First of all, the mass-count distinction as drawn up by semanticists is not directly derived from these features. There is therefore no need for rules relating combinations of the features to mass readings on the one hand or count readings on the other. Rather, the categories "mass" and "count" in linguistic semantics are abstractions which summarise metalinguistic reflection on the meaning of noun-phrases; noun-phrases whose meaning has been actualised in some context. In other words, mass readings and count readings are

lexicographic entities. Underlying these readings is the determiner system and, in particular, the properties of delimitation and individuation. These properties are to be seen as a mask on the meaning of the bare noun, constraining the interpretation which can be evolved. The "mass" or "count" reading which the lexicographer is aware of involves further elaboration and interaction with background knowledge.

At the same time, I would argue that there is no possibility of actualisation without some specification of these features. In the next chapter, I will argue that the type of an entity is provided by these features. I contend that there is no semantic realisation without typing. In particular, the bare noun has no meaning in isolation; it must be construed in the context of some phrasal unit (ie. a noun-phrase). This means that it is futile to attempt to isolate the essential meaning of a noun by studying its denotation when it appears as a bare noun. To be interpreted, it has to be built into a phrase, using a null determiner if necessary. There is no privileged access to the pure noun meaning.

Philosophers, reflective people, have worked hard on the denotations of bare nouns and their relation to the mass-count distinction. They have drawn a broad distinction between nominal and predicative usage. Roughly, nominal uses are generic, while predicative uses are particular. Quine (Quine 1960), for example, assumes that these uses are associated with different syntactic positions, but in fact they are conditioned by a number of syntactic and semantic features of the context (see Jackendoff 1983, for example). Bunt 1985 argues that this distinction can be applied to count nouns as well and that, therefore, it is orthogonal to the mass-count distinction. For this reason, I take it no further, and follow Bunt in confining my discussion to predicative uses.

Consider INDIVIDUATION. An entity which is individuated must have structural parts, even if the structure relating those parts is only "inside as opposed to outside". This follows because individuation distinguishes entities; to distinguish entities we must be able to identify a place where a different entity begins. This place is necessarily distinguished structurally from other parts of the entity. Individuation does not entail that the entity is a structural part-of some larger entity.

[+ DELIMITATION] does not entail that the entity has structural parts. This is because delimitation may bound the totality of some substance: in the case of a single sprawling object it is not necessarily the case that we can point to some part of the entity and identify

that part as its boundary. Note that when DELIMITATION is applied to a plural it applies to the set of entities.

[- DELIMITATION] presents an entity as not having structurally differentiated parts. In other words, it portrays an entity as being homogenous. There is an important philosophical issue of how to handle homogeneity. The difficulty concerns the minimal parts hypothesis. This hypothesis claims that only parts of a certain size of a homogenous substance can themselves be described properly as instances of that substance. Mereological accounts find this a serious problem. There seems to be no way of sifting out the two types of part - those which the predicate is true of, and those of which it is not true. For elements, for example, sub-atomic parts are sub-minimal; for compounds, the molecular level sets the limit. For colloids and mixtures, it is difficult to be precise about what constitutes a minimal part: is the shard of carrot that is the sole relic of a bowl of broth accurately portrayed as *broth*? Only metonymically, I suspect.

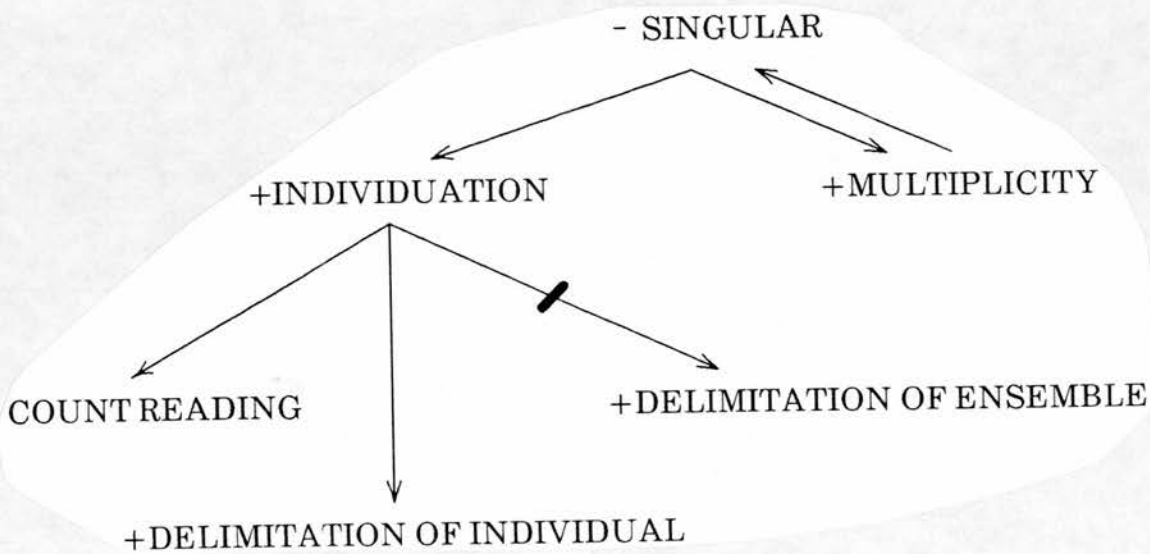
Bunt (*op. cit.*) points out that to simply invoke a subset relation which includes only parts bigger than the minimal parts is not adequate. A general relation of this type could not be specified, because the scale of the minimal part varies from substance to substance. The introduction of specific subset relations (one for each substance) will fail because no relation can evade including the minimal parts and, hence, sub-minimal parts will exist in every subset derived by the relation.

One problem is the realist perspective which has been assumed in the literature. From the cognitive standpoint, for example, there isn't a lot of difference between an element and a compound. Every day talk of *cheese* is not concerned with its microstructure. A possibility would be to talk instead of homogeneity limited by the acuity of the naked eye, and divisibility restricted to division that can be performed manually. There would be three problems with this approach. First, there are manual operations that fail to preserve mereological transitivity. For example, the preparation of ultramarine for use as paint involves grinding it to a fine powder. However, if it is ground too finely its blue colour is lost. Second, some types of homogeneity are not objectively available for inspection by the naked eye. For example, the theoretical homogeneity of substances produced by the Universal Grinder can only be imagined. Third, some types of homogeneity are functional or social rather than perceptual. That is, the homogeneity is defined in relation to some form of interaction.

I adopt a subset relation that doesn't gate out non-minimal parts, but instead excludes parts which do not satisfy the properties satisfied by the whole. Specifically, \subseteq_a retains the affordances (Gibson 1979) of the superset. This will retain the key entailments. Transitivity is assured because this is a general relation applicable to any substance. Parsons 1970 uses a similar approach, but the use of affordances differentiates this approach from Parsons', and solves a difficulty concerning the relationship between the words WOOD and FURNITURE in a universe in which they are co-extensional. Wood and furniture may coincide, but only furniture affords comfort, say. That is, wood and furniture will always be differentiated by their respective affordances.

A difficulty with Parsons' approach is its apparent circularity. Quantities may exclude sub-minimal parts, but they are informally defined: Parsons simply stipulates this property. It is a desirable property, for sure, but the key issue is how we attain it, and stipulation finesses this question.

The relationships among the features discussed in this chapter is described in the following diagram.



Part 4.

Theoretical Development.

The final part builds on the theoretical approach already outlined. Chapter 9 provides a detailed model in Space Grammar of the approach advanced in chapter 8. This is implemented in Unification Categorical Grammar. The theory is then extended to other linguistic phenomena before the chapter concludes with an integration of the results of Part 3 with those of Part 2. Chapter 10 proposes some psychological experiments and describes how these could contribute to developing the theory. Chapter 11 advances some final conclusions.

Chapter 9.

Modelling and clarification of the analysis.

This chapter is concerned first to model the analysis just presented of the mass-count distinction, and then to integrate this with the general account of lexical concepts sketched in chapter 2. I will characterize the general framework within which I intend to model, and then present the model along with some examples of its application. The model is then generalised to a broader range of data. The second part of the chapter draws the account back into the sphere of chapter 2 before presenting an integrated account of lexical combination. I conclude with some remarks on lexicalization.

9.1. Space Grammar

Ronald Langacker has over the last ten years or so been developing a representational system he calls Space Grammar, which is one of the "cognitive grammars" now emerging which seem to spring almost directly from the work of Uriel Weinreich (whose work also provided some of the impetus for generative semantics). Space Grammar is congenial because it explicitly contracts the domain of idealization - more factors are included in the domain of the model than are included, for example, in Chomsky's conception of competence (Chomsky 1965). I believe Chomsky's general approach to competence is correct. However, for certain enterprises additional variables have to be taken into account.

Let me focus on three factors Langacker incorporates. First, the grammar is (in principle) embedded in knowledge of the world. Langacker explicitly describes a view of words, for example, as embeddings in theories of the world (although he uses the term "domain"). Secondly (as I mentioned briefly in chapter 2), provision is made for differences in the relative prominence of information. And thirdly, Langacker allows for a continuum between syntax and lexicon which amounts to a distinction between forms "generated" or "sanctioned" by schemata (Langacker 1982b:26) on the one hand, and forms having "unit status" on the other. This will be a useful approach when it comes to considering lexicalization.

Space Grammar is a formalism in which language may be described, but it also expresses a philosophy about language. For Langacker, a grammar of a language will contain "content units" (roughly "lexical items") and "schemata" ("rules"). There are semantic, syntactic and phonological units, but a syntactic unit is simply a bipolar unit, one pole being phonological, the other semantic. Thus, syntagmatic combination of units is at once syntactic, semantic and phonological. The way in which units are combined is controlled by two parameters of the schema which permits the combination. The schema indicates which of the combining units is autonomous (units which are not autonomous are dependent) and which is profile determinant. To illustrate, a "constituency tree" (p39) is given below:

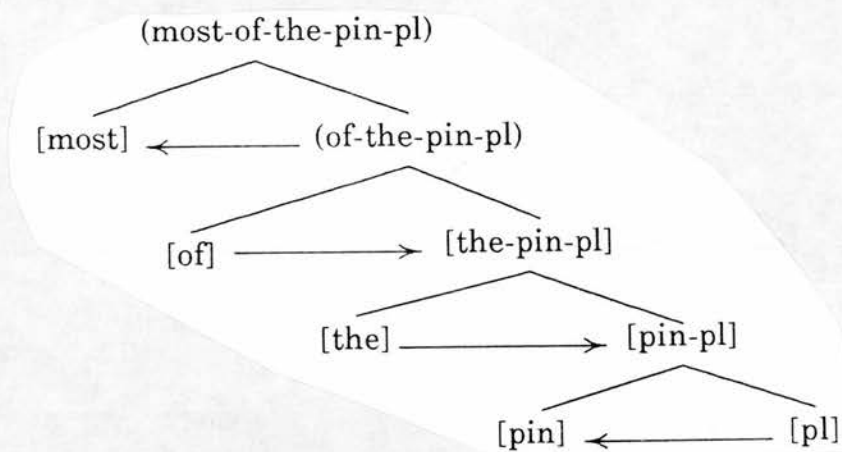


Figure 1.

Briefly, the autonomous item elaborates a gap in the structure of the dependent item (its "e-site") ie. provides the missing element. In the tree, the horizontal arrows indicate syntagmatic combination, and lead "from the dependent component to the autonomous one (p37)". Typically, the resultant composite structure will have the grammatical "profile" of the dependent item, and so the dependent item is characterized as "profile determinant" (underlined in the constituency trees). Note however, that the autonomous item can be profile determinant, as in the case of the final combination in the example given.

Combination of this sort is underpinned by the possibility of combining items without necessarily having a schema. The structures of the items are compared, and where they match they are joined. Precisely what constitutes a match is not, in fact, defined beyond the following statement:

"Correspondences are established between substructures of the components; and corresponding substructures (marked by lines of integration) are equated as referring to the same conceived entity. This overlap is what permits the components to unite in a coherent, integrated, composite structure."

The schemata only emerge as abstractions over series' of similar combinations.

Langacker makes a particular point of assigning some composite structures unit status in the interests of psychological reality. Units are delimited with square brackets, composite structures with parentheses. Thus, for example, [PIN-PL] is said not to require "constructive effort" (p24) and exists in the grammar as a unit at the same time as the schema [[COUNT N]-[PL]], "the elaborative relation between [them] ...[serving] only to categorise the [former] and to group it with other instantiations as manifestations of the same general pattern."(p36). This stance involves explicit rejection of economy of description as a goal in itself and provides a "continuum of lexicon, morphology and syntax."(p37). This is also connected to the important concept of "analysability" "the salience of the relation between a composite structure and its components."(p28). That is, the components of units will differ according to the degree to which they are traceable.

For Langacker, analysability provides the hinge ("pivot" p28) for his claim that "the meaning of a symbolic expression is not the semantic structure alone, but is rather the composite structure **in relation to** all the components out of which it is built." (p28). For example, the kinship semantics of "father" can be represented as [male-parent], but in the phrase "male parent" the same components are "active semantic components" and so the appropriate representation is [[male]-[parent]].(p29).

Langacker also makes crucial use of the notion of conceptual imagery (semantic structure "is simply conceptualization as conventionalized for linguistic purposes"(p38)). Taking a Whorfian line (different languages use different conventional imagery) he argues that five conceptual abilities with imagery are relevant (he mentions six, but repeats himself once - the third and the sixth are equivalent). To select two, first he points out that an image can be conceptualized at different levels of abstraction (analogous to different degrees of microscopic examination). Second, he notes that the relation between figure and ground can be reversed. The latter point is the essence of his analysis of the passive: what was figure becomes ground and vice-versa. Both of these conceptual abilities are relevant here.

The semantic pole of a symbolic unit is known as a predicate and designates an entity which may be either a thing or a relation. This entity is represented as a bounded region (figure) in some domain (ground), and the semantic value of the predicate lies in the relation between the two. For example, [UNCLE] is defined as follows:

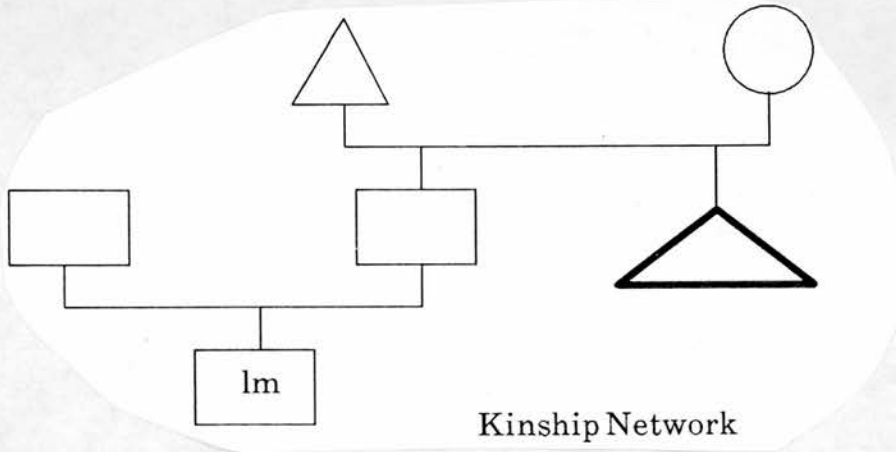
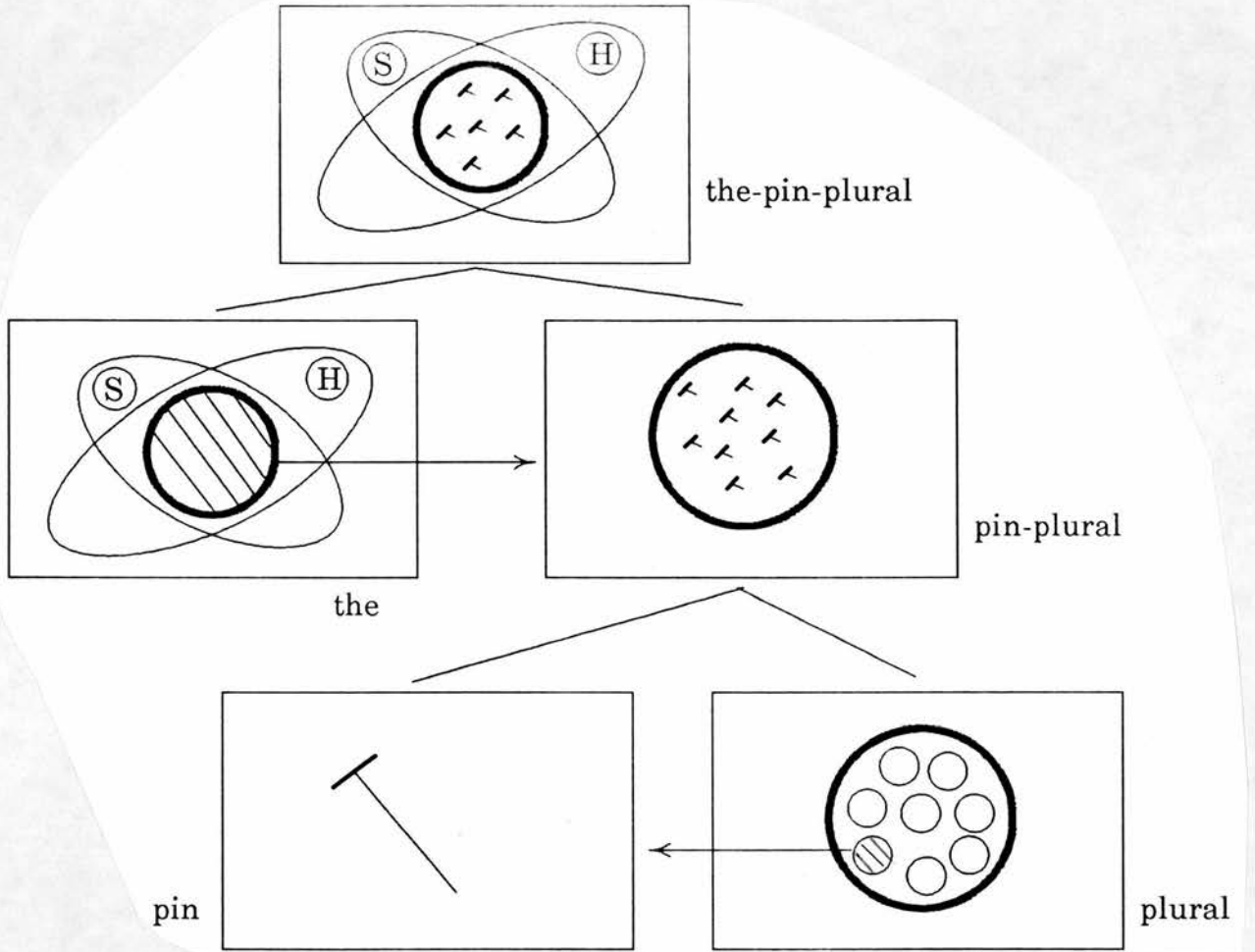


Figure 2.

In the case of relations, one of the entities is designated as trajector (tr), the other(s) as landmark (lm), this designation being "another layer of figure/ground organization within the relational profile"(p44). The subject of a sentence is a prototypical trajector. In an amplified semantic representation, [THE-PIN-PL] looks like this:

Figure 3.



In figure 3 the profile determinant is indicated by shading its outline more darkly. Note also that [THE] locates an object in the shared "epistemic space" of speaker and hearer. Langacker's concept of different types of space will be important.

Langacker (Langacker 1983) sets out to give a notional definition of the ostensibly syntactic category "noun". He presents a development of his own view that nouns denote a "bounded region in some domain". For example, a *beep* is bounded in both "pitch" and "time"; a *flash*, however, is bounded in time but not in the visual field - bounding need not occur in every domain with which a given noun interacts. "Scope of predication" is an important concept which Langacker describes as follows (p199):

"The scope of a predication (which is equivalent to its base) is the extension of its coverage within relevant domains.... I will speak of bounding only when an entity is fully included in the scope of predication (with respect to a particular domain), so that its outer limits are a specific matter of predication."

He emphasises that he refers to the way a person conceptualizes a thing, rather than the objective condition of the object itself. Considerable space is devoted to describe a concept of distance (a crude similarity metric cf. Osherson 1987) applicable to more than just physical space. This is intended to underly the optimization of nouns like *team*, for example. Thus, the members of a team are said to be construed as parts of a single object because of their abstract (specifically, co-operative) homogeneity. There are doubts about some of this (whether the similarity metric he specifies holds up empirically, or whether any of the specimen definitions given are satisfactory, for example), but it is what Langacker says about the relationship of bounding to domains that is important.

On mass nouns, Langacker expresses a view similar to the one I am taking here.

"Mass nouns are problematic for this characterization primarily because the bounding they imply is more subtle and less central to their meaning than with count nouns. I will suggest, in fact, that what distinguishes mass nouns from others is that any necessary bounding in primary domains lies outside their scope of predication.

[...] The source of this indefinite expansibility/ contractibility is the **effective homogeneity** of a mass ...[which] is **construed** as being effectively homogeneous for expressive purposes...."(pp209-210)

This is similar to the view expressed in the last chapter. He also suggests, as others have, that quantifiers can provide bounding and that bounding can be in what he terms "quality space" (p211) as well as physical space.

Important differences from the account I have sketched above are that Langacker states that quantifiers like *some* are involved in bounding, and appears to equate bounding with replicability (that is, what I have called individuation). We will see shortly how useful it is to distinguish these, but he does state his intention to look at these problems more closely in the future - this is but an interim account. It is also the case that Langacker associates bounding with the noun itself, whereas the minimal unit with which bounding can be associated on my account is the noun-phrase.

9.2. Unification Categorical Grammar

I will shortly put this theoretical apparatus to use, but first I wish to argue that a useful subset of Space Grammar can be expressed in terms of Unification Categorical Grammar (Zeevat, Klein and Calder 1987). Specifically, it will be convenient to employ the notation of UCG in what follows. I must emphasise that the fragment created here is not a fragment of UCG, but rather a fragment of Space Grammar. Nevertheless, I will draw out what I see as the correspondence between the two wherever I can. To begin with, aspects of the UCG notation can serve as a metaphor for features of Space Grammar.

First, the process of "instantiation" (ie. unification) gives us just the test for a match that is required. Second, "stripping" (which yields the result of combination in UCG) corresponds closely to the presentation of a new profile which inheres in the relationship of profile determinance. Typically, the functor will be the profile determinant and dependent item. There is even a convergence of terminology: that which Langacker calls the "elaboration site" corresponds to the thing which Zeevat et al term the "active symbol".

UCG provides several advantages: the syntax and rules of combination are more explicit; ordering can be represented; and a parser is readily available. For a brief introduction to UCG, see Zeevat et al (*op. cit.*), whose rules of combination are used here. I provide this translation as a useful result in itself.

I begin by presenting some categories for nouns and the determiners described above. The categories proposed incorporate the view that the determiner is both dependent and profile

determinant.

noun
N
noun'
O

Now some determiners:

a
NP/ *noun*: N: noun': post
[+SINGULAR, +DELIMITATION]noun'
O

the
NP/ *noun*: N: noun': post
[+DELIMITATION_{epistemic}]noun'
O

some
NP/ *noun*: N: noun': post
[-DELIMITATION]noun'
O

many
NP/ *noun*: N: noun': post
[+PLURAL, -DELIMITATION]noun'
O

few
NP/ *noun*: N: noun': post
[+PLURAL, -DELIMITATION]noun'
O

each
NP/ *noun*: N: noun': post
[+SINGULAR, +DELIMITATION]noun'
O

every

NP/ *noun*: N: noun': post

[+SINGULAR, +DELIMITATION]noun'

O

all

NP/ *noun*: N: noun': post

[+DELIMITATION]noun'

O

Note that the entries do not distinguish *many* from *few*, or *each* from *every*. This highlights the fact that these are only partial definitions; only the information relevant to the mass-count distinction is noted. The last category given means that the phonological form *all* is syntactically a function taking a following noun as an argument and yielding a noun-phrase with the meaning [+ DELIMITATION]noun'. In addition to the categories given above, a category is required for so-called "bare nouns" which occur in sentences like:

(101) Johnny ate bacon for breakfast.

This category will be called the "null determiner", and is broadly equivalent to *some*.

δ

NP/ *noun*: N: noun': post

[-DELIMITATION]noun'

O

Jespersen and Close independently note that a 'zero' article before a noun indicates either a generic use or reference to "an indefinite, unidentified amount" (Close, 1975 p131) - what Christophersen (reported in Jespersen, 1949) calls the "parti-generic" sense.

A couple of examples will serve to illustrate the formation of noun-phrases of the form "det N".

a boy

instantiation:

a
NP/ *boy*: N: *boy'*: post
[+SINGULAR, +DELIMITATION]*boy'*
O

stripping:

a boy
NP
[+SINGULAR, +DELIMITATION]*boy'*
O

some cheese

instantiation:

some
NP/ *cheese*: N: *cheese'*: post
[-DELIMITATION]*cheese'*
O

stripping:

some cheese
NP
[-DELIMITATION]*cheese'*
O

This captures in mechanical fashion the description of the previous chapter. An important point is the use I have made of the position in semantic structure reserved for the "index" in UCG; I have put my features here. This is not too surprising, since one of the distinctions marked by indices in UCG is the mass-count distinction. The additional claim that I would stress is that the "sort" of a construction \equiv its main profile.

9.3. Classifiers and the partitive

In order to establish the utility of the characterization presented above still further, and to integrate it more closely with Langacker's work, I consider here two constructions: the partitive ("Q of NP", where Q is some quantifier) and classifier constructions (which I confine

to "a Y of NP_{head=X}").⁴⁹ Examples would be (respectively):

(102) Some of the cake.

(103) A slice of the cake.

I begin with the classifiers, since they turn out to be a little more straightforward. The first observation to make is that there is a continuum of nouns which can instantiate Y in classifiers. At one end are "measures" like *pound* which can only occur freely as the object of a cognate (stative) verb, and which cannot be promoted by passivization.⁵⁰ Measures can't enter into compounding. Thus:

(104) He weighed a pound at birth.⁵¹

(105) *A pound was weighed by him at birth.

(106) She bought a pound of sugar.

(107) *She bought a sugar pound.

At the other pole, Y is construed as being constituted of the substance denoted by the kernel noun-phrase. Here compounding is possible⁵²:

(108) She ate a stick of carrot/ cube of sugar.

(109) She ate a carrot stick/ sugar cube.

Occupying the middle ground are nouns which can be construed as containers:

(110) She drank a cup of tea.

⁴⁹ These terms are arbitrary.

⁵⁰ I omit consideration of cases where Y is *lack* or *absence* because these seem to behave differently.

⁵¹ It isn't possible to construe this as an event occurring at the birth of the pound without figuratively promoting it on the scale described here.

⁵² On compounding, many other interpretations of such compound nominals are in principle feasible, of course (cf. Coulmas 1983).

(111) ??She drank a tea cup.⁵³

(112) She ate a slice of apple.

(113) ??She ate an apple slice.

Next, it is necessary to be able to construe the kernel as undelimited (ie. it has to be given either a plural or a mass reading).

(114) He bought a pound of sugar/ chestnuts/ cu* a chestnut.

(115) He bought a cup of cheese/ grapes/ cu* an apple.

(116) He bought a stick of sugar/ ants/ cu* an ant.

The cases of an indefinite kernel in these examples are excluded on a reading as a count noun (that's what the symbol *cu means) but the first could be given a mass reading (eg. it's a particularly big chestnut). I return to indefinite kernels shortly.

Thirdly, quantifiers are ungainly in the kernel, seeming redundant and repetitious:

(117) ?He bought a pound/ cup/ stick/ of some sugar/ ants/ apple.

The sense of awkwardness is particularly strong if *all* is the quantifier in the kernel, and numerals are downright bad, unless individuation is construed as being in quality space:

(118) *A pound of three cheeses.

(119) *A cup of three cheeses.

⁵³ These judgements indicate that the reading in which the compound is paraphrased by the classifier is not available. In both these cases a distinct reading is possible. Note also that many nouns can be construed at more than one point on the continuum. For example: *cup of china* versus *china cup*. Even measures can be promoted (by compounding with a cognate noun eg. *pound weight*, *pint pot*). Conversely, it is also fairly common for containers to have specialist readings as measures, notably in the fields of cooking and agriculture.

(120) ?A stick of three ants.

The exception is the case in which the classifier is viewed as constituted of the kernel. When quantifiers do appear in the kernel (superficially in the kernel, that is), they in fact have semantic scope over the entire construction:

(121) He put a pound of cheese in the bag.

—> the bag weighed 1lb

(122) He put a stick of three ants on the counter.

—> there was one stick

Compare the following:

(123) He put a pound of each cheese in the bag.

—> the bag weighed nlbs

(124) He put a cup of each cheese on the counter.

—> there were n cups

In general, quantifiers cannot appear in the kernel of a classifier. Whenever they appear to be acceptably ensconced therein either they in fact have wide scope or they suffer perceptual neglect (although I have no evidence for the second of these claims). A pragmatic filter (unspecified) will, then, gate out any cases of quantifiers in the kernel.

Indefinite kernels follow a complex pattern. In the case of measure classifiers they can be fine, being converted to a mass reading. In all cases, if the kernel is read as being a type or sort then the construction is legitimate. That is, when quality space is the domain of individuation, there is no obstacle to requisitioning a part of the whole to serve as the contents of a container or the material constituting some object. However, when physical space is the relevant domain, a constraint is placed on these two types of classifier that can be hard to satisfy: the Y must be construed as a structural part of the kernel, in the sense that a leg is a structural part of a table, while a crumb is not a structural part of a cake. Here are some relevant examples:

(125) A pound of a cheese.

(126) A cup of a cheese.

(127) A stick of a cheese.

Despite this requirement, however, definite kernels are fine (so long as they conform to the other constraints). That is, their delimitation in epistemic space doesn't preclude a lack of delimitation in the space relevant here, which one imagines to be three-dimensional space. Overall, the greater acceptability of kernels delimited in quality space and epistemic space suggests that the domain of classifier quantification and the requirement for a lack of delimitation is physical space, and that classifiers profile an entity in this domain. Parallel constructions like "a type of" and "a sort/kind of" presumably profile in quality space, and a corresponding shift in judgements can be predicted.

The closed class word *of* appears to play an important role. For example, it seems to be obligatory:

(128) *A cup cheese.

I believe *of* plays a necessary semantic role in coercing the undelimited profile on the kernel (cf. Langacker 1982b). When it combines with the kernel NP, *of* is dependent and profile determinant. For purposes of number agreement, for example, the classifier is the head of the construction as a whole. However, the view that this noun-phrase determines the profile of the whole is only partially correct; the dependent *of*-NP profiles an important layer of figure-ground structure. Nevertheless, one could argue that this figure-ground relation is implicit in the indices of the nouns (this will become clearer shortly), and allow the classifier to be an autonomous profile-determinant (which is exactly what Langacker 1982b claims the quantifier in a partitive is). In order to draw out this important relation, however, I adopt the following category for *of*:

of
NP/W₁: [I₁]np₁': pre/W₂: [I₂]np₂': post
F-G([I₁]np₁', [-DELIMITATION]np₂')
O

Where F-G is a relation between two entities such that the first is figure and the second ground. That is, *of* is profile determinant. An example will be instructive.

a cup of tea

First, the noun-phrases:

classifier:

a cup

NP

[+SINGULAR, +DELIMITATION]cup'

O

kernel:

δ *tea*

NP

[-DELIMITATION]tea'

O

Next, combine the kernel with *of*:

instantiation:

of

NP/ W₁: [I₁]np₁': pre/ δ *tea* [-DELIMITATION]tea': post

F-G ([I₁]np₁', [-DELIMITATION]tea')

O

stripping:

of δ *tea*

NP/ W₁: [I₁]np₁': pre

F-G ([I₁]np₁', [-DELIMITATION]tea')

O

Which leaves the final combination with the classifier:

instantiation:

of δ *tea*

NP/ *a cup*: [+SINGULAR,+DELIMITATION]cup': pre

F-G ([+SINGULAR,+DELIMITATION]cup',
[-DELIMITATION]tea')

O

stripping:

a cup of δ *tea*

NP

F-G ([+SINGULAR,+DELIMITATION]cup',
[-DELIMITATION]tea')

O

For some evidence in support of the assumptions about constituency that I have made, consider, for example:

(129) Of David's chocolates, John ate a whole box

The construction *of* NP can appear on its own in no more than a handful of archaic or lexicalized cases, such as:

(130) We ate of meat and bread.

or

(131) Nicola speaks of you often.

It is therefore appropriate to treat it as dependent.

The partitive has much in common with classifiers, and one has the intuition that there may be grounds to view quantifiers as degenerate nouns, one step less substantive than measures. Some evidence to support this emerges shortly, but it will also be important to bear the difference between the two in mind in evaluating the data that follows.

First, the quantifier in a partitive appears to be of the same category as the classifier in classifiers (cf. (129)):

(132) Of David's chocolates, John ate a whole box and Mary guzzled many.

Neither is in the semi-subordinative relation to the kernel that obtains between a determiner

and noun. Rather, both appear to be the head of their construction as gauged by number agreement, and in some sense determine the profile of the whole (I have already noted that Langacker analyses the quantifier as profile determinant in partitives). Classifiers are clearly endocentric, but the situation is not so clear with partitives. On the one hand, there is the evidence of examples like:

(133) David brought sufficient cheese that everyone could have some.

(134) Some came to the match, but most preferred to watch it on TV.

This example suggests that the word can be used as a pronoun or non-anaphoric substantive. On the other hand, some of these uses have an elliptical feel (as noted in the previous chapter) and there is no doubt that these substantives are semantically "defective" in the same sense that /h/ can be considered phonologically defective (Lass 1976). Lass's term "gesture" is highly appropriate; this isn't quite the full-blooded deixis of anaphora, but there remains a strong sense of the presence of another player in the wings. Semantically an argument is obligatory, but syntactically it is not (it can be understood from situational context, for example).

This breakdown of the parallel between syntax and semantics creates a dilemma. Before advancing a possible solution, I consider some similarities between partitives and classifier constructions.

One similarity is that both demand that the kernel noun-phrase be either plural or construable as mass. A few quantifiers have the stronger requirement that the kernel must be plural. For example:

(135) He bought many of *the sugar / the slaves

At first it looks as though this is a case of straightforward agreement between quantifier and kernel: their indices must match. Indeed, Zeevat et al (*op. cit.*) propose that indices can be used in just this fashion to aid anaphora resolution. Examples like the following refute this hypothesis, however:

(136) He bought each of the slaves / *the boy.

Singular *each* demands a plural kernel. (That *each* is singular can be demonstrated by examining number agreement on verbs). The same constraint applies in anaphora:

(137) When the boys / *boy arrive(s), each has / *have a bath.

That is, the antecedent must be plural, and cases where indices match are ungrammatical. The parallel with anaphora is evidence that the pronoun and the quantifier are of the same category.

A further similarity between the constructions is that quantifiers are again ungainly in the kernel. Additionally, quantifiers that cuckoo into the kernel have wide scope in partitives too:

(138) The process wastes much of every tree.

→ the number of portions of waste = the number of trees

I will assume that the same pragmatic filter screens both classifiers and partitives.

One difference is that the structural part interpretation of classifiers with indefinite kernels is not available for partitives (although possibly my lack of ingenuity is the only obstacle). Moreover, straightforward mereological part interpretations **are** feasible:

(139) Tracy ate some of a cheese.

In this respect, the quantifiers pattern in the same way as measures.⁵⁴ This is strong support for the hypothesis expressed above that the quantifiers represent a point on a continuum, a continuum on which measures occupy an intermediate position. Furthermore, quantifiers like measures resist compounding:

(140) *He bought a carrot-some.

⁵⁴ There is a divergence, however, in relation to passivization, which I cannot readily explain:

(141) He weighed some.

(142) Some were weighed by him.

(143) It weighed a pound.

One thing that must be noted is the ungrammaticality of:

(145) *Noddy ate every of the apples.

This can be accounted for in terms of "blocking" (Aronoff 1976, Romaines 1983), cf.:

(146) Noddy ate every one of the apples.

which has the intended reading of (145) That is, the underlying concept cannot be realised as *every* because the form *every one* has been lexicalized with that meaning. This alternation of form is good evidence that the partitive employs a substantivized determiner: in the case of *every*, the substantive element is realised formally and separately.

On the basis of these arguments, the quantifier in a partitive is given the following category:

generally:

$$\begin{array}{l} W_q \\ NP \subset W_2: NP: [I_2]noun_2': \uparrow \\ [I_q]noun_2' \\ O \end{array}$$

The symbol '⊂' means that the argument is syntactically covert and may, for example, be understood from the situational context. Similarly, '↑' indicates that the physical location of a token representing the covert argument cannot be predicted. I should add that these symbols have not been implemented computationally and go beyond the rules of combination provided in Zeevat et al (*op. cit.*). Space Grammar, too, would need to be augmented to handle this type of phenomenon.

Syntactically, the quantifier functions in the same way as an NP, and so combination with *of* is feasible. The differences between quantifiers will be marked primarily by differences in the profile 'I_q', and the agreement conditions observed above will be noted on the profile of the covert category 'I₂'. Thus, for example, the instantiation of this schema for *each* will

(144) *A pound was weighed by it.

To attribute this to a difference in case role begs the question really.

be:

each
NP ⊂ W₂: NP: [+INDIVIDUATION]noun₂' : ↑
[+SINGULAR, +DELIMITATION]noun₂'
O

And *each of the boys* will instantiate as:

of the boys
NP/ *each*: [+SINGULAR,+DELIMITATION]noun₂' : pre
F-G([+SINGULAR,+DELIMITATION]noun₂',
[-DELIMITATION, +INDIVIDUATION, +MULTIPLICITY]boy')
O

Clearly, I am simply assuming a mechanism that will associate 'noun₂' with 'boy'.

I turn now to one or two particular phenomena that occur in special cases. Naturally, I will try to account for these anomalies. First consider examples involving a bare noun in the kernel: the so-called "pseudo-partitives" (see eg. Lehrer 1986. By and large, quantifier pseudo-partitives are ungrammatical:

(147) *He ate some of sugar.

(148) *He shoved many of boys.

Or to put it another way, pseudo-partitives are only alright with classifiers. One wants to be able to handle this while retaining the assumption that both quantifier and classifier are syntactically noun-phrases. One reason - beyond elegance - for not seeking a syntactic bifurcation is that certain classifiers (specifically, those Allen 1977 terms fractional classifiers) share this property with quantifiers:

(149) *Three quarters of whale have been sold.

(150) ?A tenth of air is left.⁵⁵

The difficulty seems to be that the bare noun forces a "constituted of" reading, such that the classifier is constituted of the kernel. Because, in the case of quantifiers, no entity is denoted whose constitution any substance might compose, there is an obstacle to comprehension. The covert argument of the quantifier is too insubstantial to serve, and fractional classifiers are equally ethereal. This explanation would also account for the unacceptability of preposing the prepositional kernel of a pseudo-partitive (this is in contrast with partitives, see (159)).

(158) *Of men we invited a group.

(159) Of the men we invited a group.

(158) is bad, then, because of the pragmatic peculiarity of placing focus on the material of which an object is composed. This is a difficulty for other constructions, too:

⁵⁵ Compare the acceptability of:

(151) One tenth of all air is left.

There are some exceptions among quantifiers. For example:

(152) Lots of boys came to the party.

However, (153) suggests that the sequence *lots of* is lexicalized as a quantifier:

(153) *Lots boys came to the party.

More challenging are the following observations:

(154) Plenty of boys arrived.

(155) Plenty boys arrived.

This may be associated with the delimitation *plenty* predicates in what, for want of a better term, I'll call "goal space", since:

(156) Enough of boys take up computing, but not enough girls do.

(157) Sufficient girls but insufficient of mothers understand the offer.

(160) ?Of girls I saw a picture.

That is, it is not specific to the pseudo-partitive.⁵⁶ Furthermore, the obvious prediction that acceptability will be restored when context is more supportive is borne out by the following:

(161) Of sticks we made a chair.

The final piece of evidence is that where a filler for the covert argument slot can be understood from context, even the straightforward pseudo-partitive can be rehabilitated:

(Of what did you make those decorations you sold?)

(162) Some of wood, and many of glass.

Similarly, certain fractional classifiers have a conventional reading associating them with a specific container (or possibly measure), and these readings are fine with pseudo-partitives:

(163) I'd like a half of lager.

(164) I'd like a quarter of toffees.

Among other things this set of data shows that, in a very tangible sense, being constituted of some substance is not the same thing as having the denotation of that substance; such a denotation is insufficient. One issue I haven't addressed is why the δN noun-phrase forces this reading, and for the moment I will only suggest that it is probably due to some constraint that δ places on the interpretation of N.

Before going on to the other important anomaly, I wish to briefly consider the possibility of using the sort of figure-ground reversal with which Langacker 1982b characterizes the relation between active and passive sentences to capture the rough paraphrase of classifiers and certain prepositional phrases. For example, the difference between these:

(165) a cup of tea

⁵⁶ The last three examples and associated judgements have been taken from Lehrer 1986. For clarity, I should add that some quantifier pseudo-partitives fail because of number agreement, and cannot be resuscitated by context.

(166) some tea in a cup

is *prima facie* that in (165) *tea* is in the background and *cup* is in the foreground, while in (166) *tea* is in foreground, with *cup* in the background. Three considerations mitigate against this attractive hypothesis. First, classic tests of "bar level" differentiate classifiers from prepositional phrases (see Radford 1981):

(167) A cup of tea with a handle.

but

(168) *A cup with a handle of tea.

(169) *A cup of tea and with a handle.

(170) *A cup of sugar and in a bag.

By analogy with Radford's conclusions about students of physics, such data imply that the classifier denotes a single "property", while the prepositional phrase denotes two. Langacker 1982b argues that prepositional phrases denote a relation between a trajector (ie. a figure) and a landmark.⁵⁷ This is similar to Miller and Johnson-Laird's (1976) conclusion that a relation between an entity and a relatum is denoted, and fits with the assertion that the relatum is also profiled in a prepositional phrase, being indeed the second property referred to by Radford.

Further evidence for this comes from the differing strength of the implication that there is a cup involved - it is strongest in prepositional phrases and weakest in the pseudo-partitive. Compare:

(171) Tea in a cup with a handle

(172) ?A cup of the tea with a handle

⁵⁷ He also assigns the same analysis to quantifier partitives, but the data here demonstrate that whatever analyses are correct for classifiers or partitives and prepositional phrases, they are different from each other.

(173) A cup of tea with a handle

(174) A cup of tea in a mug

(175) ??Tea in a cup in a mug

(These findings are the same for other possible prepositional paraphrases of classifiers, involving, for example *from* and *out of*.) Another complication obstructing the establishment of a general schema is the sloppy nature of the identity of the kernel of a partitive and the figure of the corresponding prepositional construction:

(176) A cup of the tea_i

(177) The tea_j in a cup

Unfortunately, tea_i and tea_j are not strictly identical. On the assumption, however, that sloppy identity is a general problem not specific to this relation, we can afford to defer it. This allows the following schema to be postulated:

$$\begin{array}{l} [{}_{np}NP_k \text{ of } NP_1]: \mathbf{F-G} (NP_k, NP_1) \vdash \\ [{}_{np}NP_1 [{}_{pp}\text{Prep } NP_k]]: \mathbf{F-Lm} (NP_1, NP_k) \end{array}$$

Where (1) the symbol '⊢' is to be interpreted as meaning that an instantiation of the left-hand-side will sanction the appropriate instantiation of the right-hand-side and (2) the relation **F-Lm** holds between a figure (trajector) and landmark.

The final phenomenon I wish to address is the possibility of dropping the medial *of* from partitives. This seems to be available only for a few quantifiers (Langacker 1982b):

(178) All the cheese

(179) Both the boys

(180) *Some the cheese

(181) *Many the boys

Langacker (*op. cit.*) argues that *of* can be omitted when its "restricted subpart" meaning is

redundant, pointing out that *of* and *both* subsume the whole of the kernel. However, a slightly different meaning has been attributed to *of* here - I have characterized it as coercing an undelimited perspective (ie. as a sort of "universal spreader"). Nevertheless, the generalization can be captured as follows: *of* can be omitted when the profile produced by the quantifier is delimited. This also captures:

(182) Half the boys

which would trouble Langacker's account. There is, moreover, a consequence. Omission of *of* removes the source of figure-ground organization from the surface, and there is a tendency towards loss of separation of figure from ground. That is, the profile of the kernel tends to be assimilated to the profile of the whole construction. In the limit, the separate article of the kernel is dropped as well, as in:

(183) A half-day

(184) A ha'penny

It becomes difficult to refer to the kernel as opposed to the whole. Compare the following:

(185) We celebrated half of the day Denholm had his 21st birthday.

(186) ?We celebrated the half-day Denholm had his 21st birthday.

When *all* is the quantifier, a singular kernel cannot be construed as individuated, it is necessary to use the form in (188) instead:

(187) cu*All the cheese

(188) The whole cheese

(188) however, also makes isolation of the kernel difficult:

(189) ?Did you eat the whole apple Sally cooked.

There seem, then, to be (at least) four levels of semantic availability an entity may have. It may be figure, landmark, ground or it may be assimilated. This is an important result because it substantiates the discussion of foregrounding and prominence in chapter 2.

9.4. Aspect and the mass-count distinction

Looking at classifiers and partitives, then, has let me illustrate the way combination works at the level of entity profiles. One can see how entities are established in cognitive space using the profiles provided by closed class words. This approach can be applied to more than just determiner-noun combinations, however. I will now explore the influence of the direct object of a verb on the aspectual denotation of the verb phrase.

Aspect is a grammatical category, relevant to the temporal contour of events or processes denoted by a construction. An important point to bear in mind, is that, as with the mass-count distinction, any real entity can be construed in different ways cognitively. In relation to aspect, Sapir puts it like this:

"Logically and psychologically, nearly every activity can be thought of as either point-like or line-like in character, and there are, of course, many expressions in English which definitely point to the one or the other, but the treatment of these intuitions is fragmentary and illogical throughout." (Sapir 1951:117)

Classically, linguists and philosophers have differentiated verbs according to their temporal contour: for example, Vendler 1957 distinguished verbs of accomplishment (eg "run a mile") achievement (eg "recognize"; "find") activity (eg "run") and state (eg "love"; "know"). More recently, people have argued that these are categories of the verb phrase (Moens and Steedman 1986, Moens 1987), or of predications (Mourelatos 1978). Essentially, the argument is that the grammatical category of aspect interacts with the meaning of particular verbs, so that in different contexts the same verb can occur in different categories. This is parallel to the claim here that the mass-count distinction is not a property of lexical entries for nouns, but results from an interaction of lexical concepts with determiners and background knowledge.

Mourelatos 1978 and several authors in Hopper 1982 go further and explicitly draw a parallel between the mass-count distinction and the imperfective-perfective distinction.⁵⁸

⁵⁸ Bolinger (1972, p161) makes a similar point:

"The singularity of nouns is so typical that the count noun is viewed as basic, with plurals generated from them and with mass nouns treated as a kind of exceptional class. Verbs are typically mass or plural-like to begin with and are 'singularized' by being tied to a time and a place.

For example, Mourelatos argues that events are inherently countable and points out that nominalizations of event predications are "count quantified" (eg "Mary capsized the boat" => "There was a capsizing of the boat by Mary") whereas nominalizations of state and process predications are not (eg "Joseph was painting the stable" => "There was painting of the stable by Joseph"). Hopper writes:

"... an event must be bounded at its inception and conclusion in order for its limits in respect to adjacent events to be maintained without overlap and hence without ambiguity."(p6)

(190) He swallowed.

contains no limitation on the extent of his swallowing:

(191) We knew the poor man was nervous because he swallowed so.

Only contextually can we be sure that a single point action is involved:

(192) He swallowed hard when I looked at him a moment ago.

This is to say that fundamentally verbs are extensible, like mass nouns - a fact that is plainer when one notes that verbal nouns are normally mass nouns, and are difficult to singularize without a fair amount of context:

(193) He talks too much.

(194) There's too much talking.

(195) *One talking is enough.

He also notes that "the recognition that durative aspect in verbs is equivalent to mass aspect in nouns is credited to Sweet by Harold V. King in "Language & Learning" (1969)." Close (1975) makes the observation, of relevance here, that "often a word can be used as a mass noun when it refers to a substance, material or phenomenon in general, but as a count noun when it refers to one separate unit composed of that substance or to an occurrence of that phenomenon"(p111). Here we see the conceptual basis of the nominal vocabulary overlapping with that of the verbal. Among Close's examples are: "This is the age of science", "Physics is a science"; "I will come with pleasure", "It'll be a pleasure to see you".

That is, boundedness (cf. DELIMITATION) is seen to be an important element of the distinction. Lyons 1968 discusses aspect in Russian, and observes that the perfective is associated with completion and the imperfective is neutral with respect to whether there has been a conclusion to activity. Mourelatos gives the example of imperfective *lecit* "treat" the addition of the perfective prefix to which yields *vylecit* meaning "cure".

Crucially, Mourelatos observes that verbs whose objects are count nouns must be interpreted as events (ie. as occurring over a delimited period) while the verbal sisters of mass nouns are interpreted as processes. Thus we have (196) as opposed to (197)

(196) he played some Bach

(197) he played a Mozart sonata.

The first observation I wish to make is that the direct object is semi-subordinate to the verb (Allerton 1982). This, of course, is the same as the relationship between a determiner and a noun.

The semantic interaction can be examined fruitfully in terms of Space Grammar. I will not pursue the translation to UCG here, because it would not further the argument.

Roughly, verbs are characterized in Space Grammar as ordered sets of configurations of a trajector and one or more landmarks ie. a set of configurations of participants ordered over time. If this "trajectory" has a positive profile in both time and space, it is perfective. For example, if "John enters the room", the spatial relationship between them changes in the course of the "execution" of the trajectory. Imperfective processes by contrast, are "degenerate in regard to the non-temporal dimensions" (p270) of the trajectory and so describe "the mere perpetuation of a configuration through time without essential change." (p271). For example, if "John hates the room", then he stands in a constant affective relationship with it. Langacker says of auxiliaries that epistemic (modal) predicates and aspectual predicates intervene between the "ground" ("the position of the speaker and speech situation" p278, 1979) and the objective situation being talked about. Furthermore, "the lexical head organizes the objective scene and provides it with an initial temporal profile, and the aspectual predicates adjust this profile and determine the ultimate temporal extension of the objective situation as it is viewed by the speaker for purposes of linguistic expression." p287, 1979. "Each aspectual predicate in the sequence adjusts the profile it receives from the

predicate below it" (p283,1979).

Langacker 1982a discusses in detail the application of this approach to the aspectual predicates BE, PERF and ING. Note the difference there is in applying PERF to perfectives and imperfectives. In general, "PERF derives a state from a process by focussing on the point at which the trajectory is fully instantiated. In the case of perfective processes this can only be the end point, but ... an imperfective process ... defines a configuration stable through time, and for this reason any point within the imperfective process meets the condition for being singled out...." (p280-281). Thus we can begin to see how the various semantic representations can interact.

Within this framework, we can see the direct object as instantiating the trajectory denoted schematically by the verb and other aspectual predicates. For example, when someone plays music, they follow a trajectory through the music, and the precise nature of that trajectory will be influenced by the contours of the music. In particular, if the music has a limit, then so will the trajectory.

Other participants will represent landmarks on the trajectory. For example, if John goes to the zoo, his path will lead to an entity - the Zoo. In short, the other participants, like the direct object, will contribute to the instantiation of the relatively schematic trajectory denoted by the verb and any aspectual predicates.

The interaction depends also on the case role of an object. In particular, the object has the strongest influence when it is a range, result or patient. The influence is weaker for a source, goal or benefactive. Compare these examples:

(198) John will paint some pictures

(199) John will paint for some friends

This can be understood in terms of the manner in which the landmark interacts with the trajectory. If the trajectory passes across some plane of the object, rather as a torch beam traverses an object in its path, then the trajectory is moulded around that object and is instantiated by that object. In the same way that a trajector progressively conceals a patient, it progressively discloses a result. If an instantiating participant is bounded, then the trajectory inevitably inherits that bounding. On the other hand, if the trajectory merely

approaches a given landmark, for example, then the path followed by the trajectory is partially instantiated not by the landmark but by the location of the landmark instead.

Langacker's general theory of aspect allows us to predict that there will be no interaction when the spatial trajectory is degenerate. When the spatial trajectory is degenerate, there is nothing for a landmark to instantiate. Stative predicates have degenerate spatial trajectories according to Langacker (a state involves no change). The following examples are stative, and there is no interaction with the direct object:

(200) He resembled a matchstick.

(201) He resembled some cheese.

Some predicates presuppose that their direct object has some semantic property. For example, SEE presupposes that the direct object is visible.⁵⁹ Similarly, the verb ROLL presupposes that the direct object can be rolled. This entails that the object has shape and is therefore individuated. Thus, for example:

(203) He rolled some cheese.

suggests that the cheese was countable. This example shows that there is an additional interaction between a predicate and the landmark. Both place constraints on the trajectory, but both must also accommodate to it. Background knowledge also plays a role in determining, for example, whether the object derives its shape from the action, or has the shape from the outset. Consider, for example:

(204) He rolled some tobacco and some cheese.

Because I concentrate here on the delimitation of the noun-phrase, rather than its status as mass or count *per se*, it is possible to correctly predict that a verb whose direct object is plural and hence countable, but is nevertheless [-DELIMITATION] (eg. *many boys*), will

⁵⁹ Modal verbs can affect this. For instance, this example doesn't carry the usual presupposition.

(202) He couldn't see the ship.

not be read as referring to an event. This can be tested with reference to the "imperfective paradox" to which accomplishments are claimed to be vulnerable (Moens & Steedman, 1985). Thus, even if Seb Coe was running a mile, we are unable to conclude that he ran a mile - he may have been run over by a truck after sixty yards. However, if he was simply running, it is clear that he ran, whatever befell him. To return to the hypothesis in question,

(205) He was running some miles when a truck hit him

does seem to me to imply that he may have run *some miles* before the accident. In other words, the undelimited but countable direct object does not turn the predicate into an accomplishment. A reading along the lines of

(206) he was running four miles before breaking four minutes

smacks of stage-Indian.

This is strong evidence in favour of an analysis which distinguishes DELIMITATION and INDIVIDUATION, rather than using the traditional categories "mass" and "count" directly.

DeJong and Waltz 1983 describe a representational system for verbs called "Event Shape Diagrams". In it, a verb is decomposed as a series of overlapping temporal contours. For example, EAT would have contours for DESIRE, INGESTION, DIGESTION, and so forth. DESIRE would rise and then begin to fall away. INGESTION would cease just after DESIRE. This is similar to a decomposition of the trajectory associated with a verb, and is consonant with Langacker's general approach. I discussed in chapter 7 the way DeJong and Waltz modelled the interaction of the direct object with the Event Shape Diagram, and how they used this to give an account of certain kinds of metaphor. It is easy to imagine implementing the analysis of the interaction of delimited objects with verbal trajectories in terms of Event Shape Diagrams.⁶⁰ The aspectual alternatives would be related in the same way as alternative metaphorical readings, with the object contributing to the choice of an appropriate reading.

⁶⁰ There is also a strong resemblance to the "ease-in, ease-out" diagrams used in advanced computer animation.

9.5. Lexical concepts and theories

This section makes more explicit the model outlined in chapter 5, and integrates it with some of the results discussed in the last chapter. It includes a brief discussion of the nature of "theories" and the interface between "theories" and lexical concepts. First let me review some key points.

I have indicated that a lexical concept can be thought of as a set of parameters used to guide the flow of background knowledge into some semantic representation. This representation may or may not be a prototype. Background knowledge is seen as a naive theory of the world, and was conceptualised as a semantic net for illustrative purposes. Plastic inheritance was the idea that a given lexical concept could be attached to different theories and thus receive different interpretations.

The mass-count distinction arises from the combination of a noun and a determiner. The noun is parameterised with respect to its profile; a noun-phrase actualises a noun (possibly as a concept) with a profile. Some readings are lexicalised and do not have to be constructed on-line.

9.5.1. Lexical concepts

To take into account the question of representing relationships among the parameters, as discussed at the end of chapter 5, lexical concepts will be augmented with a configurational parameter. This parameter is called the GESTALT parameter. This is the parameter that features such as INDIVIDUATION and DELIMITATION instantiate. This parameter would also be used to represent information about the inherent orientation of an entity, for example. It is appropriate to link inherent orientation with the mass-count distinction because possession of an inherent orientation entails countability.

As indicated in the previous chapter, this parameter is realised as the profile of an entity in Space Grammar. This is realised as an index in the CUG formalism. Note that, on Langacker's account of semantic structure, the substructures representing entities, for example, are retained as embeddings in the overall structure of a sentence. What Space Grammar gives in addition is a representation of the relative prominence of these entities in the overall structure.

Here, then, is an example of a lexical concept:

cheese

GESTALT
COLOUR
TEXTURE
FLAVOUR

This would replace *cheese'* in a representation like:

a cheese
NP
[+SINGULAR, +INDIVIDUATION]*cheese'*
O

Of course, the notation means nothing without some interpretation (see chapter 3), and there are two parts to the interpretation of lexical concepts. First, there is the general framework which attributes a role to lexical concepts in language understanding. This was discussed in chapter 5. Second, there are the "theories" which provide the semantic material to actualise lexical concepts.

9.5.2. Theories

Theories are not clearly defined anywhere in the psychological literature to which I referred in chapter 5. To some extent, it remains to be seen how they differ from knowledge representation schemes in the a-intelligence literature, such as KRL (Bobrow and Winograd 1977) or KL-ONE (Brachman and Schmolze 1985). Nevertheless, it is possible to clarify one or two points. (These points are developed in Dunbar (in preparation b)).

First, rather than considering theories as individual semantic nets, they should be seen as collections (indeed, configurations) of knowledge structures. On different occasions, the set of knowledge structures that is active will vary. This will affect the availability of information for any semantic representation, and the interpretation assigned to lexical concepts will therefore be subject to the composition of the set of currently active theories.

In particular, we should distinguish local theories from global ones. A local theory is linked to a particular episode and represents a subject's current understanding of that episode. In many instances, the local theory is precisely the semantic representation which lexical concepts are helping to construct. Nevertheless, local theories as well as global ones will be used in assigning an interpretation to lexical concepts. Local theories could, obviously, provide a mechanism for the pre-emption discussed in chapter 2.

As indicated earlier, default values are supplied by theories. If the set of active theories changes, the default values are also liable to be labile. A concept of default set of defaults can be derived by postulating a default active configuration of theories. Context-independent properties can be understood in the same way: a set of properties which are constant and frequently accessed under the default active configuration of theories. Similarly, Weinreich's notion of default mass and count readings for lexical items can be seen as default cognitive perspectives. This is a subtly different concept from the notion of a default context which is justifiably criticised by Crain and Steedman 1985.

9.5.3. Theories, lexical concepts and the mass-count distinction

This section looks at the way mass and count readings can be derived from the noun-phrases analysed earlier in the chapter. Obviously, the way I represent theories is simplistic; the aim is to illustrate the structure of the system in terms of the relationship between theories and lexical concepts.

Consider *a cheese*. The currently active configuration of theories must instantiate an individuated entity with values on the parameters TEXTURE, COLOUR and FLAVOUR, say. In this case, we will assume that the current theories lexicalise this schematic phrase, and directly supply default values to actualise the phrase.

Now consider *a bread*. Let's assume that the same parameters must be satisfied but that on this occasion, no lexicalised concept is furnished. The hearer must work out the scale of the individual denoted. There are two ways of doing this that I want to emphasise. First, background knowledge may contain a disjunction like "individuals of bread are slices or rolls or loaves". It may also contain a probabilistic rule of inference like "at dinner an individual of bread is a roll" which can be combined with a premise from the local theory (for example, "this is dinner") to select the "roll" unit. The current theories may then supply default values for the parameters of "roll". The second possible mechanism involves pre-emption. For example, a roll may be being pointed to by the speaker. Because of the salience of the speaker (the hearer is tracking the speaker; that is the definition of listening) the object indicated will dominate the local theory.

Once a concept has been actualised it is likely to be cached in some fashion, ready for re-use. That is, although the concept is potentially transient, there is often some utility in retaining it temporarily in case, for example, the discourse subsequently refers back to it. However, the form of caching is a problem for theories of discourse (or text) processing and anaphora resolution and so I leave it for now. It is worth noting, however, that caching is probably the first stage of lexicalisation.

I have described how modification would be handled in chapter 5. Here, I simply illustrate it with a fresh example: *some blue cheese*. The adjective masks the COLOUR parameter of the head, constraining its instantiation. This example shows that the tailoring of *blue* is not a function of the default colour of *some cheese* but is instead subject to background knowledge about "cheese". In particular, cheese coloured the default "pale yellow-orange" is suffused with the colour; blue cheese is veined with the colour. In addition, there is no sense in which the hue or saturation of the blue is some function of the hue and saturation of "focal blue" and hue and saturation of "pale yellow-orange". It is largely independent of these colours, deriving instead from a theory about stilton. Often background knowledge will be less specific about the colour and the pattern of its application, but the general principle will remain the same.

So far, I have assumed a model in which the goal is to create a prototype. This is not always the case, of course. Consider (208) and (209) as alternative responses to (207)

(207) Did Eric say whether he was coming?

(208) He said he wanted to peel some tomatoes.

(209) He said he wanted to peel three tomatoes.

Perhaps the local theory has a goal to answer the question "should we keep Eric's seat for him?" It might use the responses simply to satisfy this goal. Of course, it's stretching things to imagine that the hearer really would use the distinction between a delimited and an undelimited activity to make his decision, but it is conceivable and, in any case, I only wish to demonstrate the mechanism in principle.

Drawing together the main points to emerge from this chapter:

The theoretical apparatus of Space Grammar has been helpful in general, and the attempt to translate the analysis into the notation of UCG has succeeded.

The type of a phrase and its main profile are the same thing.

While determiners have the syntactic category NP/N, quantifiers have the additional category NP. Quantifiers therefore present a local ambiguity. The quantifier is only profile determinant when it has the semi-subordinate role of NP/N.

Langacker's notion of a multi-dimensional semantic space has been useful in permitting us to differentiate types of delimitation.

The meaning of a nounphrase has four possible depths which can be distinguished grammatically. In addition, the "head" noun of a classifier construction can be placed on a semantic continuum running from quantifiers through measures to full substantives.

Participants in a proposition interact with the main verb in the actualisation of the trajectory of the process.

The dependency relation of semi-subordination is associated with semantic interaction at the level of profile.

Chapter 10.

The role of experimental psychology.

In the first chapter I argued that experimental psychology could play an important role in a cognitive science of lexical meaning. I claimed also that it could not be used simply to provide a transparent test of a linguistic theory. Rather, the two disciplines would each furnish their own data, each type of data carrying implications for different aspects of an integrated cognitive theory. In particular, different types of evidence are relevant at different stages in the development of a theory.

This chapter shows in detail how this approach can be applied to the topics of this thesis. I discuss some problems and issues that arise from the model developed so far. I then show how these can be resolved experimentally. Unfortunately, pressure to finish the thesis within four years means that the experimental program has not been pursued properly. Nevertheless, I describe two proposed experiments in detail, and present data obtained from one of them. Several other experimental possibilities are discussed in less detail.

I address three issues. First, is contextual influence on conceptual instantiation related to the combination of concepts, or is it unrelated to the syntactic structure of language. Second, can psychological techniques differentiate concepts which have been lexicalised from transient denotations which have been derived on-line. And, finally, does the cue validity of a referential function affect the representation of the alternative readings.

I begin by reporting an experiment addressing the first issue.⁶¹ Barsalou 1982 proposed that two kinds of property were associated with any concept: some activated on all occasions (CI) and others activated only when context encourages them forth (CD). For example, "has a smell" is said to be a CI property of skunks, while "where cooking can occur" is a

⁶¹ Many thanks to Joe Levy and Martin Shepherd for their help with programming the BBC. Thanks are due also to Terry for advice, encouragement and above all, his great patience in bearing with the gestation of this experiment. Statistical analysis was performed using GENSTAT and ISTAT software.

CD property of hospitals.

Barsalou correctly predicts in his own experiment, that subjects reading sentences containing some noun (the target noun) will verify that it possesses some property (the probe) more quickly if sentential context supports the property only if the property is CD. If the property is CI, then the presence of facilitatory context makes no difference to verification time because, he argues, the property is activated anyway. For example, consider the target-context-probe triples below.

- | | |
|--|--------------------------|
| 1 The hospital served meals all day. | -where cooking can occur |
| 2 The hospital had many famous surgeons. | -where cooking can occur |
| 3 The skunk stunk up the area. | -is smelly |
| 4 The skunk ran across the road. | -is smelly |

Verification of the probe is typically faster in one than two, but no quicker in three than four.

Barsalou's hypothesis is clearly of potential importance to our study, particularly as he offers an account of the acquisition of the two kinds of property and an indication of an interesting range of problems upon which the CI-CD distinction might have a bearing [much like the one constructed in chapter 1]. Further, his method could provide a tool with which to explore context effects (in particular, the effect of articles on the interpretation of nouns).

A striking aspect of the materials used by Barsalou is that, if it were the case that comprehension involved semantic decomposition, then the decomposition of his examples of "related" predicates would include as a member the property to be verified. For example:

(210) The hospital was quiet when dinner was served.

This leaves open the possibility that reaction time advantages accruing when related predicates are used are not due to modulation of the representation of the subject noun, but rather are due to activation of the property itself. This problem in fact covers two separate questions, both of which will be addressed in the experiment reported below.

First, is it necessary to include the property in the set of properties that *prima facie* would make up the semantic decomposition of the predicate in order to facilitate verification time? In practise, it will be necessary to generate predicates which can reinforce a property

without mentioning it directly. That is, predicates which would not inevitably introduce the property on every occasion of use. There is, however, a familiar phenomenon which can guide us to appropriate materials; many lexicographers discuss "syncategorematicity" in relation to adjectives like GOOD and GREAT. Such adjectives reinforce, or emphasise different properties according to which noun they are used to modify (Keil 1979, Miller 1977). For example, a *great painting* is not (typically) "great" for the same reason as a *great hammer*. Thus, if we can obtain syncategorematic predicates, and if these predicates can facilitate property verification, then it must be by modulation of the subject noun. In the design of experiment 1, then, a new factor is crossed with context type of predicate. This new factor has two levels: "original" predicates are similar to those used by Barsalou and mention the property fairly explicitly; but "syncategorematic" predicates might reinforce a different property if predicated of another noun. For example,

(211) The rose stained the carpet.

Here the verb phrase draws attention to the redness of the rose, but could equally well draw attention to the yellowness of custard or the darkness of coffee.

Secondly, is the grammatical relationship between subject and predicate required for facilitation to occur, or can it happen even if the property is simply mentioned in the verb phrase, without that VP standing in the relationship of predicate to the sentence subject? That is, is it necessary that the grammatical context explicitly connect property with subject, or is it sufficient that the property be mentioned in the general vicinity? If the latter holds, then the phenomenon is analogous to the transferred epithet figure of speech under whose auspices, for example, the phrase "the man drained the drunken glass" can attribute the property of being intoxicated to the man. If the process is, on the other hand, one of modulation, then we would expect the predicative relationship to be an essential component of the effect.

In order to test this, a third level is added to the factor of context-type of sentences: as well as using reinforcing and neutral predication, predicates will be used which mention the relevant attribute without predicating it of the subject noun. Thus the context can "juxtapose" a property and a noun, as in the following sentence (the property being "is red"):

(212) The rose made the girl blush.

A similar control appears in the context of a study of encoding specificity, in Anderson et

al. 1976.

Thus, to test the assumption that contextual interaction is due to the modulating effect of syntagmatically related material, we can test the generality of Barsalou's results with the experimental conditions described. The predicates he uses as reinforcing context are fairly explicit in their reference to the property under scrutiny. The question I ask is whether predicates reinforcing a property without mentioning it could provide reinforcing context as ably as ones as explicit as those used in Barsalou's original study. A second issue related to this is whether it is the mere mention of a property which matters as opposed to mention within the subject-predicate construction. If interaction depends on combination, we will find first that for both levels of subject-relatedness and at least Barsalou's levels of context-type, syncategorematic predicates affect reaction times in the same manner as original ones, and second that juxtaposition of targets and properties does not have an effect like reinforcing predication. If the data pattern within the constraints of these expectations, then the questions raised above will have received a satisfactory answer.

Finally, the relationship between target and property is said to underly the results, but in Barsalou's materials non-overlapping sets of properties were used for the two types of probe. Thus, the relationship between probe and target was confounded with probe. In Experiment 1, this confounding is undone by using each property in both conditions: will the results mentioned above still appear? They should unless Barsalou unwittingly used a procedure for assigning properties to conditions that was biased, such that the properties used were inherently CI or CD.

10.1. Method

10.1.1. Procedure

Subjects sat in front of a vdu and rested their forefingers on two response buttons about 8 inches apart. To start a trial, the subject pressed the space bar of a keyboard linked to the vdu. This lay behind the two keys. After a 500ms interval a sentence appeared two-thirds of the way up the screen. All sentences began with THE followed by a flashing subject noun (Gsubj) and a predication (Pred) of the subject noun. Subjects were instructed to understand the sentence and to read it aloud. After 6s a property label appeared below the sentence (about one-third of the way up the screen). If the flashing noun in the sentence

possessed the property, subjects pressed the "true" key (which was always beneath the forefinger of the dominant hand); otherwise they pressed "false". Reaction time was measured from first appearance of the property label to response. Subjects were instructed to respond as quickly as possible but to avoid error. They received 24 practice trials and 48 test trials. They could take a break at any time, but rarely did, typically completing the experiment in around twenty-five minutes.

10.1.2. Subjects

Subjects were 12 young adults (mean age = 21yrs 1mth) all of whom and all of whose parents were native speakers of English (self-report). They had between zero and four years experience of undergraduate psychology.

10.1.3. Materials

Materials consisted of 12 properties each linked to 6 sentences which represented the six treatment combinations obtained by crossing the factors of grammatical subject and kind of predicate. The Gsubj could be either highly or weakly related to the property, and the Pred could be related, unrelated or "juxtaposed" to the property. These relationships were verified by the ratings of an independent group of 4 judges. Each of these judges (graduate students of cognitive science) rated all Gsubj's and Preds. For "juxtaposed" Preds only the keyword was used. Each item was written on one side of an index card, on the back of which was written the property. Judges rated how much the item made them think of the property from 1 "it didn't come to mind at all" to 7 "it came to mind immediately". The cards were randomly ordered for each subject.

ANOVA's were performed on these ratings in order to check that they fell in the predicted patterns.

Looking first at predicates, the prediction is basically that the different levels of predicate should vary in relatedness, but that Preds of a given level should have similar relatedness whether they were to be used with a highly related grammatical subject or a weakly related one. It should be noted that frequently it was possible to use the same predicate with both grammatical subjects, which of course means that in such cases there was no difference in

the relatedness of the predicate to the property. The method used for syncategorematic predicates had to be slightly different since these were specifically designed not to be inherently related to the property ⁶². The technique used was a variant of the "free association" method. Subjects were presented with questions of the form:

(213) If the Gsubj Pred, then [what value does it have for] PARAMETER?

For example,

(214) If the bath fills quickly, then what does it contain?

Subjects had to supply a value or tick a "don't know" box. If the value was the one to be used in the experiment ("water" in the last example) then this was scored as 7. Any other value was scored as 1 unless it was a partial synonym of the value to be used, in which case it scored 4. "Don't knows" were scored as 2. For syncategorematic predicates, then, the prediction is slightly different: with a highly related Gsubj, scores should be higher than with a weakly related one. Means are shown in table 1.

Table 1 Ratings of predicate association.

	Pred	Original			Syncategorematic		
		REL	UNREL	JUXT	REL	UNREL	JUXT
Relatedness of Gsubj	HI	4.21	1.21	3.83	5.92	1.79	3.96
	LO	4.25	1.13	3.67	2.79	1.87	4.04

Separate analyses of variance were done for original and syncategorematic predicates. For originals, the main effect of relatedness was not significant ($F(1,3) < 1$), while the main effect of predicate was ($F(2,6) = 23.46, p < 0.01$). For syncategorematic predicates, both were significant: relatedness, $F(1,3) = 115.46, p < 0.01$; predicate, $F(2,6) = 89.4, p < 0.01$. These results support the view that related predicates were more related to the property than unrelated ones, and that this relationship was the same for both strongly and weakly related subjects.

⁶² Note that for all syncategorematic items, the same predicates instantiated a given condition for both highly and weakly related subjects.

Second, it was necessary to establish that highly related nouns had been rated more highly than weakly related ones, and that there was no difference between original and syncategorematic grammatical subjects in this respect. Note that because the same noun-pair was to be used in both CI and CD conditions, any effect of context which differs with condition cannot be attributed to a difference in the relatedness of noun-pairs used in the different conditions. The main effect of grammatical subject was significant as predicted ($F(1,3)= 51.39, p<0.01$), while the main effect of "Type" (syncategorematic versus original) was not ($F(1,3)<1$). There was no interaction of these factors either ($F(1,3)<1$). The means appear in table 2

Table 2 Mean ratings of grammatical subjects.

	Type	Original	Syncategorematic
Relatedness			
of Gsubj	HI	6.33	6.54
	LO	2.54	2.71

24 sentence-property pairs were also constructed, each sentence having a subject noun that clearly did not possess the property. (The same 12 properties being used as for the "true" items). They were, however, constructed in such a manner as to represent the six key treatment combinations (Gsubj x Pred) with respect to some implicit property. The purpose of this was to prevent subjects from being able to anticipate the correct response on the basis of the relationship between a grammatical subject and a predicate alone. For example, if a subject noticed on a given trial that the predicate was highly related to some property to which the grammatical subject was also highly related, this observation would not give him any information as to the appropriate response in advance of the appearance of the property; the correct response could only be decided by examination of the property label.

Finally, there were 24 sentences representing in roughly equal distribution the 12 treatment combinations of the experiment (true-false x Gsubj x Pred), and these were presented first to subjects and used as the practice items.

10.1.4. Design

For each property, there were six experimental sentences. These six were grouped into 6 pairs such that each subject could receive one pair from each of the twelve properties⁶³. With the aid of a computer a skeleton list of items was produced for each subject. These lists encoded which pair the subject would get for each property, and what order the 24 "true" items would appear in. This order of presentation was randomized separately for each subject, but constraints were placed on the assignment of pairs to subjects such that firstly all subjects received each treatment combination twice and secondly each property appeared in each treatment combination exactly four times across subjects.

A specially written lisp program (Appendix A) was then used to prepare materials files from these skeleton lists, adding a random number of false items between successive true ones. The relative order of the 24 false items was the same for all 12 subjects. Each subject received his list of 48 true and false items after first receiving the 24 practice items. All subjects received the practice items in the same pseudo-random order.

10.2. Results

Analyses of variance were performed on the correct "true" item's latencies by subjects (F1) and by items (F2). For subjects, the latency used was the average of two latencies within each treatment combination, and for items it was the average of four.

Planned comparisons were evaluated using the Bonferroni t test, with the adjustment for repeated measures recommended in Myers (1979).

F1 and F2 were first calculated without juxtaposed predicates. The overall main effect of grammatical subject was significant for RT's ($F(1,11) = 17.475$, $p < 0.01$; $F(1,10) = 36.37$, $p < 0.01$), and also for errors ($F(1,11) = 15.476$, $p < 0.01$; $F(1,10) = 62.83$, $p < 0.01$). However, the main effect of Pred was not significant for RT's ($F(1,11) < 1$; $F(1,10) < 1$), and,

⁶³ These pairs were selected in such a way as to avoid repetition of lexical material, and were as follows (in terms of subject relatedness and predicate relatedness respectively): a) HI,REL & LO,UNREL; b) HI,UNREL & LO,JUXT; c) LO,UNREL & HI,JUXT; d) HI,JUXT & LO,REL; e) LO,REL & HI,UNREL; and f) LO,JUXT & HI,REL.

although it was significant for errors in the items analysis ($F(2,1,10)= 5.517, p<0.05$), it wasn't by subjects ($F(1,1,11)= 1.658, p> 0.20$). Contrary to the expected result, the Gsubj by Pred interaction was not significant for RT's either ($F(1,1,11)< 1; F(2,1,10)< 1$), but it was significant in terms of errors ⁶⁴ ($F(1,1,11)= 4.459, p< 0.10; F(2,1,10)= 5.976, p<0.05$). Indeed, as can be seen from table 3, the error rate was rather high.

Table 3 Mean response latencies and error rates.

	Pred	RELATED		UNRELATED		(JUXTAPOSED)	
		Lms	E%	Lms	E%	Lms	E%
Relatedness							
of Gsubj	HI	1191	14.6	1227	10.4	1370	8.4
	LO	1752	31.2	1756	58.4	1613	47.9

Turning now to the specific comparisons indicated by the experimental hypotheses regarding the levels of Pred, let us examine first the difference between related and unrelated predicates. With highly associated grammatical subjects, related predicates did not permit faster verification than unrelated ones ($t_1= 0.243, df=11; t_2= 0.182, df= 10$) or fewer errors ($t_1= 0.407, df=11; t_2= 0.247, df=10$). For weakly associated grammatical subjects, however, there was unexpectedly no difference between verification times ($t_1= 0.117, df=11; t_2= 0.217, df=10$), although there were significant differences in error rates ($t_1= 2.59, p< 0.10; t_2= 3.704, p< 0.05$) for the two levels of predicate.

The main effect of Type was not significant ($F(1,1,11)< 1; F(2,1,10)= 1.631, p> 0.20$) and neither was its interaction with Pred ($F(1,1,10)< 1; F(2,1,10)= 1.313, p> 0.20$). Furthermore, although the interaction with Gsubj was significant by items ($F(2,1,10)= 5.516, p< 0.05$), and the second order interaction with Gsubj and Pred approached significance by subjects ($F(1,1,10)= 3.893, p< 0.10$), neither interaction achieved significance on the other analysis ($F(1,1,10)< 1$; and $F(2,1,10)< 1$, respectively). I shall return to the three-way interaction shortly, but the conclusion from these results that syncategorematic type predicates affect

⁶⁴ "... CD properties in irrelevant contexts are normally inactive and errors for these sentences occur when subjects decide to respond prior to this information's becoming active." (Barsalou 1982:86)

RT in the same manner as original ones is borne out by the analyses of errors, which found the main effect and all interactions of Type to have F's uniformly <1. The mean RT and error rate for each condition in the subject's analysis is shown in table 4

Table 4 Mean response latencies and error rates.

Type of	Pred	RELATED		UNRELATED	
		Lms	E%	Lms	E%
Pred	SYNCAT	1485	25	1422	35.4
	ORIGINAL	1456	20.8	1561	33.3

As was noted above, the Type-Gsubj-Pred interaction approaches significance. An unexpected feature of table 4 is that for syncategorematic types, weak predicates were faster than related ones. Although no specific prediction was made about this pair of means, it's clear that we would not have expected the weak predicates to be quicker. Table 5 gives a three-way breakdown of RT's and error rates.

Table 5 Mean response latencies and error rates.

	Pred	RELATED		UNRELATED	
		Lms	E%	Lms	E%
	Gsubj				
SYNCAT	High	1196	16.5	1277	12.5
	Weak	1771	33	1563	58.4
ORIGINAL	High	1211	12.5	1210	8.4
	Weak	1683	29.2	1933	58.4

The correlation between mean verification time and error rate for these eight cells is high and positive ($r= 0.813$, $p< 0.02$), indicating that no speed-accuracy trade-off is taking place.

Focussing now on the reaction times to sentences whose grammatical subject was weakly associated with the probe, there is a nearly significant difference between related and unrelated predicates for original-style sentences ($t_1= 1.828$, $df= 11$, $p< 0.10$; $t_2= 0.615$, $df=10$, not sig.), but not for syncategorematic ones ($t_1= 0.007$, $df=11$; $t_2= 0.172$, $df=10$). The most obvious conclusion is that the relative explicitness of mention is vital for context (ie. the verb phrase) to facilitate verification, but this will be re-examined in the Discussion. Relevant will be the observation that for syncategorematic types, related predicates are on average 200ms slower than unrelated ones when the subject is weakly associated with the target property.

In contrasting the two types of sentence, it has been important to be abnormally lenient with the alpha level, because the null hypothesis is that there will be a difference. That is, a restrictive alpha level would be increasing the chance of supporting the experimental hypothesis of no difference. Note further that, although $t_2(\text{ORIGINAL, WEAK, REL v UNREL})$ is not statistically significant (as noted in the previous paragraph), the magnitude of the difference was 296ms, implying a lack of power.

In order to compare juxtaposed and related predicates, analyses by subjects and by items were also carried out with these as the only levels of Pred. Again, the main effect of Gsubj was significant, both in terms of RT ($F_1(1,11)= 6.697$, $p< 0.05$; $F_2(1,10)=15.868$, $p< 0.01$) and error rate ($F_1(1,11)= 10.735$, $p< 0.01$; $F_2(1,10)= 17.521$, $p< 0.01$). The main effect of predicate is not significant in terms of RT ($F_1(1,11)< 1$; $F_2(1,10) <1$), and although

there is a difference in error rate by items ($F2(1,10)= 4.416, p=0.062$), it does not hold in the subjects analysis ($F1(1,11)< 1$). The interaction between grammatical subject and predicate was not significant for RT ($F2(1,10)< 1$) or errors ($F2(1,10)= 1.421, p> 0.20$) by items. By subjects, however, both RT ($F1(1,11)= 8.461, p< 0.025$) and error rate ($F1(1,11)= 4.452, p< 0.10$) achieve significance. See table 6.

Table 6 Mean response latencies and error rates.

	Pred	RELATED		JUXTAPOSED	
		Lms	E%	Lms	E%
Gsubj	High	1191	14.6	1370	8.4
	Weak	1752	31.2	1613	47.9

.mn Specific comparisons between means at these two levels of predicate were non-significant for highly related grammatical subjects on RT's ($t1= 1.772, df= 11; t2= 0.685, df= 10$) and error rates ($t1= 0.814, df= 11; t2= 0.187, df= 10$). For weakly related grammatical subjects, the subjects comparisons just miss the 0.10 level ($t1= 2.315, df=11$ and $t1= 2.168, df=11$ for RT's and errors respectively), but the items analyses fall short ($t2= 0.685, df=10; t2= 1.873, df= 10$, respectively).

Finally, a full analysis of variance was also run by both subjects and items, and it revealed the anomaly that, with juxtaposed predicates included, the overall main effect of Type achieved a degree of significance ($F1(1,11)= 12.148, p< ; F2(1,10)= 2.896, p=0.12$). This is odd since it was not significant with juxtaposed predicates, and since these two cells of juxtaposed predicates are meant to be functionally identical. That is, no syncategorematic-original distinction can be made for juxtaposed predicates since they must involve explicit mention. There is clearly, however, a large difference in mean reaction time to the two (289ms in the items analysis; 372ms by subjects). Since these conditions are equivalent in terms of the experimental variables (by definition, as I have indicated), the difference between them must be due to some difference between the particular items used (the factor type is confounded with items in the design, see above).

10.3. Discussion

The results are superficially unsatisfactory. There is a lack of differentiation of related and juxtaposed predicates, and the difference between related and unrelated predicates with a weak grammatical subject is found for original style sentences (thus replicating Barsalou (1982)), but not for syncategorematic ones. It would appear to be the case that the key to making context-dependent properties more available is no more and no less than simple mention in physical proximity to the test stimulus. However, two principle features of the data force a qualification of this interpretation. Firstly, differences that were not significant in terms of reaction time were so in respect of error rate. Thus, more errors were made when a weakly related grammatical subject was collocated with an unrelated predicate for both sentence types. Secondly, there appeared to be a lack of power, with relatively large differences (around 200ms) being declared insignificant.

Clearly, the experiment needs to be repeated, and a key problem that will have to be tackled first will be the need to reduce subject error. On reflection, it was felt that the instructions to subjects could have indicated more clearly that they were to judge whether the grammatical subject *could* possess the property in question, rather than whether this particular exemplar *did* have it. Furthermore, feedback could be given to subjects as to their accuracy during the practice items. A reduction in error rate might be made a criterion for beginning the experimental items. It would seem advisable also to increase the numbers both of subjects and of items, in order to improve the power of the design. Having now obtained estimates of error variance from this experiment, we are in a position to work out with some precision how many to add.

Two anomalies emerged in the results. First, the expected relationship between related and unrelated predicates (the latter should not be quicker under any circumstances) was reversed for syncategorematic sentences with a weakly related grammatical subject. A careful review of experimental technique has identified the difficulty in the method of rating the relatedness of syncategorematic predicates to properties. The reader may recall that it was realized that precisely because of the nature of such predicates, they could not simply be rated, and so a variant of the method of free association was used, with weakly related ones being defined as those which did not cause the judges to respond with the target property. What I now realize is that this definition entails that judges must have been inclined to respond with *some other* property. This theorem enables us to explain the anomaly; subjects were

slower responding because a preferred association was interfering with their response to the probe; unrelated predicates carried no such baggage.

The second anomaly we have drawn attention to is the apparent failure to successfully match juxtaposed predicates in the two groups distinguished by the factor Type. The nature of this failure is perhaps revealed by a post hoc qualitative analysis. Juxtaposed predicates were subjected to a grammatical analysis, and the results are tabulated in table 7.

Table 7 Frequency table of juxtaposed predicate types.

	Syncategorematic		Original	
	High	Weak	High	Weak
Direct_obj	1	1	3	3
Locative	3	2	2	2
Adjective	2	3	1	1

Obviously, it would be better if there was no discrepancy in the formal properties of "matched" predicates.

10.4. Proposals for experiments

A possibility worth systematic investigation is that the relationship between contextual combination and interaction is subject to the processing mode of the hearer. For example, it might be the case that subjects attending more carefully enjoy context effects more closely tied to the combinatorial structure of the text. This possibility is suggested by the discussion of undifferentiated reference in chapter 2. It could be tested by blocking subjects by a locus of control variable: field independent subjects would be predicted to process text more analytically than field dependent subjects. Another possibility is that the foreground—background structure of a text affects contextual interaction. In particular, the discussion in chapter 2 and chapter 9 suggested that backgrounded entities, such as landmarks, may not be projected so strongly. One can ask, then, for example, whether the contextual reinforcement of properties (Barsalou, 1982) or contextual adaption between head and modifier occurs even in non-focal positions.

Unfortunately, there is not time to pursue this experimental program further. I turn now to the second and third issues raised at the beginning of the chapter. How can we differentiate lexicalised readings from derived ones? And does the cue validity of the referential function relating two readings affect the likelihood of their separate lexicalisation?

To tackle these issues, I propose using a technique investigated by Barsalou and Ross 1986. Their experiments involved presenting subjects with a list of nouns. Subjects then had to estimate how frequently given superordinate categories (or context-independent properties) had been instantiated by exemplars in the list. For example, the list might contain *apple*, *orange* and *pear* among the nouns. After reading the list, the subject would be asked to estimate how many *fruits* were in the list. Barsalou and Ross (1986) were particularly concerned to discover whether frequency sensitivity was tied to strategic processing or was a consequence of automatic processing. The result I rely on, however, is that subjects were sensitive to the frequency of occurrence of superordinate categories in a condition in which they were trying to learn the list for free recall. This is the condition I propose to employ.

Consider two readings of the noun *lamb*. On one reading, it is an "animal"; on the other, it is "meat". If derivation of the reading "meat", for example, needs to pass through an intermediate designation of an "animal", then one would expect the frequency counting mechanism to increment its "animal" tally by one.⁶⁵ If, on the other hand, neither sense is basic, only the frequency estimate for "meat" should be driven up by this item. This would also be the case if both readings were lexicalised separately.

The set of exemplars for each category should be either newly derived in one reading or lexicalised in both readings. This would be a between subjects variable at the category level. Each subject would receive both types of exemplar set, but for different categories. One could then compare frequency estimates by subjects receiving a set of dually lexicalised terms with estimates from subjects who have to derive one of the readings. The prediction is that the frequency estimate for "animal" should not be affected in the dually lexicalised case, and will only be affected in the other case if subjects must derive the reading via an intermediate designation.

⁶⁵ I follow Barsalou and Ross (1986), of course, in not wishing to suggest frequencies are stored in a literal way. See footnote 1, page 117 (*ibid.*).

One difference in design between such a study and Barsalou and Ross (1986) would have to be the incorporation of each noun in a context sentence that made the reading clear. Barsalou and Ross merely presented a list of nouns and relied on a default background context to ensure subjects generated the intended readings.

The other question was whether the nature of the referential function would make a difference. Specifically, would a function with lower cue validity such as "fur-of" (cf. chapter 7) force derivation via an intermediate designation of the "animal"? This would be accomplished by orthogonally varying the cue validity of the referential function. For example, one subject's list might contain the items *mink*, *ocelot* and *seal*, contained in sentences like:

(215) The lady's coat was made of mink.

For the category animal, this subject's data would be assigned to the condition "low cue validity, both readings lexicalised". The interesting condition would be the one involving similar sentences containing nouns like *cat* or *horse*, for which the "fur-of" reading is novel. Would this condition differ from the more straightforward one involving novel "meat-of" readings? If only the former inflates the frequency estimate for the category "animal", then the hypothesis that cue validity affects the derivation of readings would be supported.

An alternative to the frequency estimation approach would be the phrase-by-phrase reading time task described by Janus and Bever 1985, and discussed in relation to metaphor in chapter 5. One would predict relatively slower reading times for novel readings. This would be particularly useful in cases of combination, such as the mass—count distinction. Again, one could manipulate a "complexity of referential function variable". The frequency estimation task, too, could be used to look at the lexicalisation of complex concepts.

In connection with referential functions, an interesting question from a psychological point of view would be to ask whether the function is computed using a "tool concept" like a Piagetian operation (cf Cometa and Eson 1978). or, instead, using background knowledge in the form of theories. This is a key issue in developmental psychology. It could be tested, for example, by looking at the generalisability of referential functions to new knowledge domains, for children and adults of various ages.

This chapter has shown how experimental psychology can be used to test issues that can be difficult to resolve using linguistic techniques alone (the second question); to show how linguistic analysis can show us what data to look for (the third question); and to test assumptions of the linguistic analysis (the first question). Unfortunately, it has more of the character of a prospectus than I would have liked.

Chapter 11.

Conclusions.

My goal has been to understand better a range of phenomena of flexibility of meaning, while at the same time developing an interdisciplinary approach within which the issues could be studied further. I have argued that the lexicon plays a role in making information available, and I have attempted to formulate an approach that lets it do this in a way that takes account of the flexibility of lexical meaning. At the same time, my approach allows for phenomena of precization. In this approach, lexical concepts are actualised by a configuration of, more or less naive, theories. Actualisation may or may not yield a concept, which may or may not have a prototype structure. The form actualisation takes will depend on the current goals of the hearer and the state of his background knowledge. Neither concepts nor the edifice of background knowledge can be regarded as static. Actualised lexical concepts may be cached or even lexicalised.

Precization is the complement of flexibility. I have argued that both are a function of lexical combination and integration: they are two aspects of this single phenomenon. Precization is the consequence of a particular combination; flexibility the result of (or, better, the prerequisite of) the generality of combination. I presented a fairly detailed description of a way lexical concepts might combine with other lexical concepts, and of a way they might combine with theories. Particular attention was paid to the role of dependency relations in determining the manner of combination, especially the effect of semi-subordinate items on the GESTALT parameter. I argued that the instantiation of this parameter, the main profile of a construction, is its type.

This conception of the lexicon is radically minimalist (cf Cruse 1988). I won't repeat the earlier discussion of these points here, but I do wish to elaborate on a small set of issues that arise from them.

From the beginning, I have emphasised a cognitive approach to meaning. The meaning of a word is the contribution it makes to the understanding the listener arrives at. It can contribute to this understanding in different ways, but it should always be understood in terms of this role.

The cognitive approach to semantics has allowed us to avoid certain problems that do occur for a more realist approach. For example, problems with the minimal parts hypothesis were avoided in this way. In short, not all problems in philosophy need to be tackled by Cognitive Science because a cognitive approach to meaning, and the willingness to allow a single entity to be seen in different ways in particular, means that the issue does not really arise. We will see another example of this shortly.

It is a mistake, standard in cognitive psychology, to regard the semantics of the lexicon as a list of concepts which can be known by the categorisation judgements of the subject. The concept of lexical semantics that pervades cognitive science is based on this idea of semantic material, taken (or copied?) from the lexical entry and placed somewhere else. I have argued, instead, that the lexicon supplies tools which can be used to construct all manner of semantic objects. My approach in this regard is similar to one discussed by Brian Cantwell-Smith (Cantwell-Smith 1986). He argues that one conception of the way language influences the mind would be in terms of a sort of electro-magnetic causal force operating at a distance.

I want to draw out a distinction between two modes of signification which lexical items may use. I don't believe this distinction can be drawn sharply, but nonetheless the poles can be clearly differentiated.

First, they can be used in the cognitive manner I have concentrated on. They interact with other parts of the cognitive system to construct a new semantic object - something I have termed a transient denotation. And on the other hand, they can be used in the manner of a gesture to nudge an evolving representation (local theory) along the appropriate path of elaboration. This mode becomes increasingly likely when the shared context is rich. It is therefore more common in spoken language and especially common towards the end of a spoken utterance. The sort of example one finds is:

(216) She is a fiery, tempestuous, vicious woman.

Predicates are added like splashes of colour on an impressionist canvas, until the speaker is satisfied that the portrayal is correct.⁶⁶

A special case of the second mode of signification is allusion to a culturally conventional lexicalised concept. The stereotype of this is the idiom (see Becker 1975). The speaker uses a conventional formula to remind the hearer of an idea they share (with the presupposition that the speaker typically believes it is relevant to the situation and that the hearer has not realised its relevance).

Lexicalised concepts, in general, are not lexical concepts. Rather, they exist only in configurations of theories. On this view, ambiguity only arises if both senses are lexicalised in the current active configuration of theories. A lexicalisation in another active configuration can only be accessed if the word-form is marked in some way. Similarly, if a lexicalisation is available it is used; a novel reading is only derived if there is no lexicalisation accessible or if the word-form was marked (for example, stressed). This view of ambiguity allows me to account for the register-specific nature of "blocking". In different registers, the active configuration of theories differs, and so the newly derived form does not have to compete with the same set of lexicalised concepts. It may therefore find a lexical gap in one register that does not exist in another.

One issue worthy of more investigation is the question of what motivates the lexicalisation of a concept. Typically one finds the suggestion that it is a matter of rote learning (for example, Bybee 1985), but this is not very plausible psychologically. It is likely that two types of factor are important - personal factors and social factors. On the one hand, people are likely to lexicalise concepts that are vivid or useful (Barsalou 1985) to them personally. On the other, they are likely to be influenced by the way words are used by significant others and public commentators. In this respect, it would be interesting to look at the use of the word *strike* among British television and radio football commentators in the period during and after the 1986 World Cup, for example, and its spread among the printed media and everyday language. Socially, one might expect some distinction between the effect on Scottish English as opposed to the sassenach dialects. Another interesting factor would be the relationship between the timing of repetition and lexicalisation - how soon must the first

⁶⁶ How the speaker gauges this is an interesting question.

re-iteration follow; does there need to be a gap between an initial group of uses and a "confirmation group"? With Romy, I noticed that when she was learning her first words she used a technique one might call "parental cross-bearing" to identify interpretations worth learning. A concept was attractive, then, if mummy and daddy both used it in the same way independently of one another. That is, she was sensitive to the utility of social consensus. An alternative starting point would be to look at the motives likely to produce the caching of lexical concepts, since this is probably the first step towards eventual lexicalisation (rather as canonisation is the first step in beatification) and since one might be able to identify some linguistic variables affecting the probability of caching.

I have emphasised the cognitive approach to meaning. Within this approach I have motivated the feature DELIMITATION, in particular, as an ontological primitive. Support was found in the way the perfective—imperfective distinction appears to use the same distinction, and from the way the two distinctions interact. This feature was built into an ontology, that is, a type theory, by arguing that the main profile of a construction was its sort. Using Langacker's notion of a multi-dimensional semantic space, this theory of types could be extended to a theory of categories. In particular, a category can be defined as a pair of a type (as just described) and a set of parameters (ie. a semantic dimensionality). As Ryle 1954 says:

"there are not just two or just ten different logical *metiers* open to the terms or concepts we employ in ordinary or technical discourse, there are indefinitely many such different *metiers* and indefinitely many dimensions of these differences."
(p10, quoted in Drange 1966, p75).

This definition gives some useful results. First as I suggested in chapter 5, it can help us to account for anomaly. In chapter 5, I argued that anomaly was the result of trying to apply a value to a parameter which a lexical concept did not possess. If the lexical concept did possess an appropriate parameter, the bizarreness reaction did not result even if the concept realised contradicted background knowledge (eg. *carnivorous sheep*). This approach is similar to that advocated by Drange 1966, Strawson 1952. Strawson is quoted by Drange:

"The reason why 'loud smell' makes no sense is [that] ... [r]eferring to something as a smell proscribes as senseless the question whether it is loud or not loud: it does not commit the speaker to one of two incompatible answers to the question."
(p27)

Drange (1966) argues at length that "type crossings", as he calls them, are not attributable to definitional falsity if an ordinary knowledge restriction is placed on the notion of

definition to exclude clauses like "what pertains to the olfactory sense cannot have an auditory property" (p30).

The other useful result, which comes from the nature of parameters, is that the actualisation of a category (or lexical concept) is constrained. A word cannot be used to denote any concept, it can only denote a concept in the semantic space defined by its dimensionality. It cannot occupy other dimensions and it cannot be realised in fewer dimensions. This is a refinement of Langacker's notion of subtypes (Langacker, 1983) as a delimited zone within the space defined by the dimensions of the superordinate: I have added the constraint that the exemplar must exist in the same set of semantic dimensions, no fewer and no more. At the same time as constraining the actualisation of a lexical concept, this makes certain predictions regarding diachronic change. It allows, for example, for diachronic change from one meaning to its opposite.

The constraints expressed in the last paragraph, however, express tendencies rather than firm rules. Drange (1966) argues that there is a sharp distinction between phrases which are type crossings and phrases which are not. For Drange, a type crossing is an objective fact.

"Thus, just as pegs come in all shapes and sizes, so also things are of many different *types*. And just as pegs are not the right shape or size to fit in a give hole, so also things are sometimes not of the right type to have the properties ascribed to them in propositions." (p148)

I would prefer to talk of a graded distinction. From the point of view of cognitive semantics, it is not, in general, useful to focus on objectivity. The shape of both hole and peg is a function of the cognitive system which is construing them. Within the theoretical framework I have developed, one must allow for the flexibility of construal. For example, I would allow categorial promotion (cf the discussion of the promotion of measures to living beings in chapter 9), which Drange explicitly forbids. Drange says that a stone moving under its own steam is no longer a stone. From the point of view I have described, whether a phrase expresses a type crossing will be a function of the conceptual creativity of the hearer and the current active configuration of theories. The current theories can ameliorate what is superficially a type crossing by sanctioning the construction directly (eg. by lexicalising a concept for the construction) or by allowing the cognitive system to move with more or less difficulty to a configuration that does sanction the construction.

I have been at pains to emphasise the distinction between a lexicographic attitude to word meaning, and the approach of the cognitive psychologist. These differences result in very different characterisations of lexical meaning. The lexicographic technique uses concordances - samples of actual usage together with the (linguistic) context of use - to provide data. The lexicographer categorises the uses and assembles constellations of readings. This process reflects emergent meaning back into the lexical entry. The lexicographer has the task of apportioning meaning to the lexical items in the example string, and tends to do this in such a way that none of the meaning, none of the semantic material, is left over. In other words, lexicography ascribes the precise interpretation of strings to the precise senses of words. My approach, in contrast, ascribes precision to the process of combination and integration inherent in lexical actualisation - lexical concepts are not so much vague as meaningless. A corollary of this approach is the lexical version of Heisenberg's uncertainty principle. The issue is whether the structure and representation I have described in this cognitive account are, in principle, accessible to processes of introspection.

Consider representations which we feel, subjectively, we can be conscious of: memories of specific events, or the meanings of words. Our everyday experience is that we can bring these things to mind. Nevertheless, there is evidence that the process of access itself contributes to and alters meaning. Take, for example, the literature on eyewitness testimony. Loftus has demonstrated that the words used in asking about an event can affect the way the event is remembered. In particular, presupposing something in a question (for example, *Did the man with the scarf hit her?*, which presupposes that the man wore a scarf) tends to convince people that that thing happened or was present. In other words, we do not have privileged access to even our own cognitive lexicons.

Lexicography, as a psychological activity, can be defined with some precision in terms of the theory I have presented. Specifically, lexicography is the search for configurations of theories which yield lexicalised actualisations of lexical concepts.

My starting point was a loose collection of phenomena, linked by the common theme of lexical flexibility. Although the discipline of thesis-writing forbids too deep an exploration of a wide range of phenomena, I hope that I have clarified the relationships among these phenomena to a small degree (see chapter 5). I want to focus here on a distinction between phenomena of depth of meaning and phenomena of the breadth of meaning.

Phenomena of depth of meaning include differences in degrees of elaboration or availability due to the manipulation of focus, and the semantic illusions studied by Erickson and Mattson (see chapter 2). Evidence for at least a four-way distinction (figure, ground, landmark, assimilation) in depth of availability was presented in chapter 9 and related to Langacker's Space Grammar. Langacker's notion of layers of profiling structure is especially useful and, I believe, important. The neutralisation of the English plural in examples like *drains* and *woods* (it no longer denotes individuation) is related to this, but is a phenomenon more of diachronic lenition. Bybee (1985) discusses this in relation to morphology, and the argument can be extended to closed class words (see chapter 9).

Depth of meaning is also related to mode of information processing, I have claimed. In a holistic mode of processing, meaning is elaborated less deeply than in an analytic mode. The analytic mode is less likely in what I take to be the most usual mode of language processing - goal driven augmentation of the local theory (see chapter 9).

Phenomena of breadth of meaning are connected to combination and integration. By breadth of meaning, I mean the range of concepts a lexical concept has the potential to be realised as. Adjectival modification, type crossing, theory change, pre-emption, profile fixing and emergent meaning exemplify this.

The distinction that I have drawn here between depth and breadth of meaning cross-cuts some traditional distinctions, such as the distinction between diachronic phenomena and phenomena of figurative language (cf Ullman 1962). For example, metaphor and metonymy are breadth phenomena, while transferred epithets are, probably, a depth phenomenon.

There is a major class of phenomena upon which I have barely touched. These are phenomena concerned with non-monotonicity. I have tried to present an approach to the breadth and depth of lexical meaning that is monotonic in spirit, and therefore do not explore such issues here. However, in Dunbar (in preparation, a) I discuss them more fully.

The key to my approach on breadth of meaning is the complementarity of vagueness and precization (see also Bosch 1983). The contextually driven notion of actualisation means, of course, that a given lexical concept has the potential to be realised as any of an infinite number of concepts. The proof is straightforward. Because realisation depends on context

and because there is an infinite number of possible contexts (where contexts are indexed by times, worlds, cognitive states, situations and so on), there are an infinite number of possible actualisations. I have already argued that parameterization of lexical concepts constrains actualisation - it is not the case that just anything is possible. This, however, is not sufficient to explain the large degree of agreement between subjects about the correct interpretation of a particular use of a word. This must be explained in terms of the intersubjective similarity of knowledge structures. People's theories tend to be similar because people are constrained by the need to survive in the same world and because they are subject to similar socialising forces. For this reason, I have sometimes referred to background knowledge as "folk theories" (Dunbar 1989). Folk theories facilitate allusion to culturally conventional concepts. Agreement on meaning is also controlled by the judgement of the speaker, who can add predicates until he is satisfied that he has communicated the correct message (see above). This is something people do skilfully. Well, I hope so anyway.

References.

- Allan, K. (1986) *Linguistic Meaning*, Volume I. London: Routledge and Kegan Paul.
- Allan, K. (1980) Nouns and countability. *Language*, **56**,281-567.
- Allerton, D. J. (1982) *Valency and the English Verb*. London: Academic Press.
- Anderson, R. C. and McGraw, B. (1973) On the representation of meanings of general terms. *Journal of Experimental Psychology*, **101**, 301-306.
- Anderson, J. R. and Bower, G. H. (1973) *Human Associative Memory*. Washington D.C.: Hemisphere.
- Anderson, J. R. (1976) *Language, Memory and Thought*. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- 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 Behavior*, **15**, 667-679.
- Anglin, J. M. (1983) Extensional aspects of the pre-school child's word concepts. In Seiler, T. B. and Wannenmacher, W. (eds.) *Concept Development and the Development of Word*, pp245-268. Berlin: Springer-Verlag.
- Armstrong, S. L., Gleitman, L. R. and Gleitman, H. (1983) What Some Concepts Might Not Be. *Cognition*, **13**, 263-308.

- Aronoff, M. (1976) *Word Formation in Generative Grammar*. Cambridge, Mass.: MIT Press.
- Barclay, J. R., Bransford, J. D., Franks, J. J., McCarrell, N. S. and Nitsch, K. (1974) Comprehension and semantic flexibility. *Journal of Verbal Learning and Verbal Behavior*, **13**, 471-481.
- Barsalou, L. W. (1981) PhD Thesis, Psychology, Stanford University.
- Barsalou, L. W. (1982) Context-independent and context-dependent information in concepts. *Memory and Cognition*, **10**, 82-93.
- Barsalou, L. W. (1985) Ideals, central tendency and frequency of instantiation as determinants of graded structure in categories. *Journal of Experimental Psychology (Learning, Memory and Cognition)*, **11**, 629-654.
- Barsalou, L. W. and Ross, B. H. (1986) The roles of automatic and strategic processing in sensitivity to superordinate and property frequency. *Journal of Experimental Psychology (Learning, Memory and Cognition)*, **12**, 116-134.
- Barsalou, L. W. (1986) Are there static category representations in long-term memory?. *Behavioural and Brain Sciences*, **9**, 651-652. Commentary on Schank et al (1986).
- Bates, E., Bretherton, I., Shore, C. and McNew, S. (1983) Names, gestures and objects: symbolization in infancy and aphasia. In Nelson, K. (ed.) *Children's Language*, Volume Vol. 4.
- Becker, J. D. (1975) The Phrasal Lexicon. In Schank, R. and Nash-Webber, B. L. (eds.) *Theoretical Issues in Natural Language Processing*, Cambridge, Mass, June 10-13, 1975, pp60-63.
- Benveniste, E. Problemes semantiques de la reconstruction. *Word*, **10**, 251-264.

- Berry (1975) *Introduction to Systemic Linguistics I : Structures and Systems*. Batsford.
- Bever, T. G., Garrett, M. F. and Hurtig, R. (1973) The interaction of perceptual processes and ambiguous sentences. *Memory and Cognition*, **1**, 277-286.
- Bierwisch, M. (1970) Semantics. In Lyons, J. (ed.) *New Horizons in Linguistics*, pp166-184. Harmondsworth, Middlesex: Penguin.
- Bloom, L. (1970) *Language Development: form and function in emerging grammar*. MIT.
- Boas, H. U. (1988) The internal structure of lexical entries: structural and/or 'definitional' semantics. In Hullén, W. and Schulze, R. (eds.) *Understanding the Lexicon: Meaning, Sense and World Knowledge in Lexical Semantics*, pp50-61. Tübingen: Niemeyer.
- Bobrow, D. and Winograd, T. (1977) An overview of KRL, a knowledge representation language. *Cognitive Science*, **1**.
- Boden, M. *Artificial Intelligence and Natural Man*. Sussex: Harvester Press.
- Bolinger, D. (1965) The atomization of meaning. *Language*, **41**, 555-573.
- Bosch, P. (1983) Vagueness is context-dependence. A solution to the Sorites paradox. In Ballmer, T. and Pinkal, M. (eds.) *Approaching Vagueness*, pp189-211. Amsterdam: North Holland.
- Brachman, R. J. and Schmolze, J. G. (1985) An Overview of the KL-ONE Knowledge Representation System. *Cognitive Science*, **9**, 171-216.
- Bradley, D. C. (1983) *Computational Distinctions of Vocabulary Type*. Bloomington, Indiana: Indiana University Linguistics Club.

- Bradshaw, J. L. (1974) Peripherally presented & unreported words may bias the perceived meaning of a centrally fixated homograph. *Journal of Experimental Psychology*, **103**, 1200-1202.
- Bransford, J. D. and McCarrell, N. S. (1977) A sketch of a cognitive approach to comprehension: some thoughts about understanding what it means to comprehend. In Johnson-Laird, P. N. and Wason, P. C. (eds.) *Thinking: Readings in cognitive science*, pp377-399. Cambridge: Cambridge University Press.
- Broadbent, D. E. (1981) Perceptual experiments and language theories. *Philosophical Transactions (Series B) of the Royal Society*, **295**, 375-385.
- Bunt, H. (1981) On the why, the how and the whether of a count/mass distinction among adjectives. In Groenendijk, J., Janssen, T. and Stokhof, M. (eds.) *Formal Methods in the Study of Language*, pp51-77. Amsterdam: Mathematical Centre Tracts.
- Bunt, H. (1985) *Mass-terms and Model-theoretic Semantics*. Cambridge: Cambridge University Press.
- Bybee, J. L. (1985) *Morphology: A study of the relation between meaning and form*. Amsterdam: John Benjamins Publishing Co.
- Cantwell-Smith, B. (1986) The Computational Metaphor. .
- Caplan, D. (1972) Clause Boundaries and Recognition Latencies for Words in Sentences. *Perception and Psychophysics*, **12**, 73-76.
- Carey, S. (1983) Constraints on word meaning - natural kinds. In Seiler, T. B. and Wannenmacher, W. (eds.) *Concept Development and the Development of Word Meaning*, No. 12, pp126-146. Berlin: Springer-Verlag.
- Carey, S. (1984) Cognitive development: the descriptive problem. In Gazzaniga, M. (ed.) *Handbook of Cognitive Neuroscience*, pp37-79. New York: Plenum Press.

- Carrithers, C. and Bever, T. G. (1984) Eye-fixation patterns during reading confirm theories of language comprehension. *Cognitive Science*, **8**, 157-172.
- Cartwright, H. M. (1975) Some remarks on mass nouns and plurality. *Synthese*, **31**, 395-410.
- Chomsky, N. (1965) *Aspects of the Theory of Syntax*. Cambridge, Mass.: MIT Press.
- Clark, E. (1973) What's in a word? On the child's acquisition of semantics in his first language. In *Cognitive Development and the Acquisition of Language*. New York: Academic Press.
- Clark, H. H. and Haviland, S. E. (1977) Comprehension and the given-new contract. In Freedle, R. O. (ed.) *Discourse Production and Comprehension*, Volume 1, pp1-40. Norwood, N.J.: Ablex.
- Clark, H. H. (1983) Making Sense of Nonce Sense. Chapter 9 in d'Arcais, G. B. F. and Jarvella, R. J. (eds.) *The Process of Language Understanding*, pp297-331. Chichester: John Wiley and Sons. contextual expressions.
- Clark, H. H. and Gerrig, R. J. (1983) Understanding old words with new meanings. *Journal of Verbal Learning and Verbal Behavior*, **22**, 591-608.
- Close, R. A. (1975) *A Reference Grammar for Students of English*. London: Longman.
- Cohen, L. J. (1979) The semantics of metaphor. In Ortony, A. (ed.) *Metaphor and Thought*, pp64-77. Cambridge: Cambridge University Press.
- Cohen, G. (1983) *The Psychology of Cognition*, 2nd. London: Academic Press.
- Cohen, B. and Murphy, G. L. (1984) Models of Concepts. *Cognitive Science*, **8**, 27-58.

- Cole, R. A. and Jakimik, J. (1980) A Model of Speech Perception. Chapter 6 in Cole, R. (ed.) *Perception and production of fluent speech*, pp133-163. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Collins, A. M. and Quillian, M. R. (1969) Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, **8**, 240-247.
- Collins, A. M. and Loftus, E. F. (1975) A Spreading Activation Theory of Semantic Processing. *Psychological Review*, **82**, 407-428.
- Cometa and Eson (1978) Logical operators and metaphor interpretation. *Child Development*, **49**, 649-659.
- Conrad, F. G. and Rips, L. J. (1986) Conceptual combination and the given/new distinction. *Journal of Memory and Language*, **25**, 255-278.
- Coseriu, E. and Geckeler, H. (1981) *Trends in Structural Semantics*, No. 158. Tübingen: Gunter Narr Verlag.
- Coulmas, F. (1983) Underdeterminacy and plausibility in word-formation. In Bauerle, R., Schwarze, C. and von Stechow, A. (eds.) *Meaning, Use and Interpretation of Language*, pp46-61. Berlin: de Gruyter.
- Crain, S. and Steedman, M. J. (1985) On not being led up the garden path: the use of context by the psychological parser. In Dowty, D., Karttunen, L. and Zwicky, A. (eds.) *Natural Language Parsing: Psychological, Computational, and Theoretical perspectives*.
- Crangle, C. E. (1984) A Computational Approach to Lexical Meaning. PhD Thesis, Stanford University. UMI.
- Cruse, D. A. (1988) Word meaning and encyclopaedic knowledge. In Hüllen, W. and Schulze, R. (eds.) *Understanding the Lexicon: Meaning, Sense and World Knowledge in Lexical Semantics*, pp73-84. Tübingen: Niemeyer.

- DeJong, G. F. and Waltz, D. L. (1983) Understanding novel language. In Cerconne, N. (ed.) *Computational Linguistics*, pp131-147. London: Pergamon Press.
- Dekeyser, X., Devriendt, B., Tops, G. A. J. and Geukens, S. (1979) *Foundations of English Grammar*. Antwerpen: Uitgeverij De Nederlandsche Boekhandel.
- Donaldson, M. (1978) *Children's Minds*. London: Fontana.
- Donaldson, M., Grieve, R. and Pratt, C. (1983) General introduction. In Donaldson, M., Grieve, R. and Pratt, C. (eds.) *Early Childhood Development and Education*, pp1-9. Oxford: Basil Blackwell.
- Donnellan, K. S. (1966) Reference and Definite Descriptions. *Philosophical Review*, 75, 281-304.
- Drange, T. (1966) *Type Crossings: Sentential meaninglessness in the border area of linguistics and philosophy*. The Hague: Mouton.
- Dunbar, G. L. (1984) The Perception of the Meaning of Words. MA dissertation, Department of Psychology, University of Edinburgh.
- Dunbar, G. L. (1988) Nouns and verbs. Talk given to the Mental Lexicon Workshop, Centre for Cognitive Science, University of Edinburgh.
- Dunbar, G. L. and Myers, T. F. (1988) Concept combination and the characterisation of lexical concepts. In Hullen, W. and Schulze, R. (eds.) *Understanding the Lexicon: Meaning, Sense and World Knowledge in Lexical Semantics*, pp292-302. Tübingen: Niemeyer.
- Dunbar, G. L. (1989) Folk Theories and Componential Analysis. Forthcoming. Paper presented at XXII Linguistische Kolloquium, Berlin, August 1988.

Dunbar, G. L. (in preparation, a) The meaning of adjectival modification.

Dunbar, G. L. (in preparation, b) The structure of folk theories.

Engelkamp, J. and Zimmer, H. D. (1983) Foci of attention in comprehension and production of sentences. In Rickheit, G. and Bock, M. (eds.) *Psycholinguistic Studies in Language Processing*, pp119-133. Berlin: de Gruyter.

Erickson, T. B. and Mattson, M. E. (1981) From words to meaning: a semantic illusion. *Journal of Verbal Learning and Verbal Behavior*, **20**, 540-551.

Fillenbaum, S. and Rapaport, A. (1971) *Structures in the Subjective Lexicon*. New York: Academic Press. lexical fields..

Fillmore, C. (1988) The Pisa Lectures. Unpublished. European Science Foundation Summer School Lectures.

Fodor, J. A., Bever, T. G. and Garrett, M. F. (1974) *The Psychology of Language*. New York: McGraw Hill.

Fodor, J. A. (1975) *The Language of Thought*. New York: Thomas Crowell.

Fodor, J. A., Garrett, M. F., Walker, E. C. T. and Parkes, C. H. (1980) Against definitions. *Cognition*, **8**, 263-367.

Foss, D. J. (1969) Decision processes during sentence comprehension: Effects of lexical item difficulty and position upon decision times. *Journal of Verbal Learning and Verbal Behavior*, **8**, 457-462.

Foss, D. J. (1970) Some effects of ambiguity upon sentence completion. *Journal of Verbal Learning and Verbal Behavior*, **9**, 699-706.

- Foss, D. J. (1982) A discourse on semantic priming. *Cognitive Psychology*, **14**, 590-607.
- Franks, B. The Symptoms of Reality. PhD Thesis, Centre for Cognitive Science, University of Edinburgh. In preparation.
- Fraser, B. (1979) The interpretation of novel metaphors. In Ortony, A. (ed.) *Metaphor and Thought*, pp172-185. Cambridge: Cambridge University Press.
- Ganong, W. F. (1980) Phonetic categorization in auditory word perception. *Journal of Experimental Psychology (Human Perception and Performance)*, **6**, 110-125.
- Gardener, H. (1974) Metaphors and modalities: how children project polar adjectives onto diverse domains. *Child Development*, **45**, 84-91.
- Garnham, A. (1981) Mental models as representations of text. *Memory and Cognition*, **9**, pp560ff. part of his PhD thesis.
- Garrod, S. and Sanford, A. J. (1977) Interpreting anaphoric relations: the integration of semantic information while reading. *Journal of Verbal Learning and Verbal Behavior*, **16**, 77-90.
- Gazdar, G., Klein, E., Pullum, G. and Sag, I. (1985) *Generalized Phrase Structure Grammar*. London: Basil Blackwell.
- Gibson, J. J. (1950) *The Perception of the Visual World*. Boston, Mass.: Houghton Mifflin.
- Gibson, J. J. (1979) *The Ecological Approach to Visual Perception*. Boston, Mass.: Houghton Mifflin.
- Gombrich, E. H. (1960) *Art and Illusion*. London.

- Gordon, B. and Caramazza, A. (1985) Lexical access and frequency sensitivity: frequency saturation and open/closed class equivalence. *Cognition*, **21**, 95-115.
- Greenfield, P. and Smith, J. (1976) *The Structure of Communication in Early Language Development*. New York: Academic Press.
- Guillaume, G. (1984) *Foundations for a Science of Language*. Amsterdam: John Benjamins Publishing Co.
- Haberlandt, K. F. and Graesser, A. C. (1985) Component processes in text comprehension and some of their interactions. *Journal of Experimental Psychology*, **114**, 357-374.
- Half, H. M., Ortony, A. and Anderson, R. C. (1976) A context-sensitive representation of word meanings. *Memory and Cognition*, **4**, 378-383.
- Hampton, J. A. (1987) Overextension of conjunctive concepts; evidence for a unitary model of concept typicality and class inclusion. *Journal of Experimental Psychology (Learning, Memory and Cognition)*.
- Hawkins, J. A. (1978) *Definiteness and Indefiniteness*. London: Croom Helm.
- Herriot, A. (1969) The comprehension of active and passive sentences as a function of pragmatic expectations. *Journal of Verbal Learning and Verbal Behavior*, **8**, 166-169.
- Hewson, J. (1972) *Article and Noun in English*. The Hague: Mouton.
- Hill, J. C. (1983) A Computational Model of Language Acquisition in the Two-year-old. *Cognition and Brain Theory*, **6**, 287-317.
- Hill, J. C. (1986) A computational model which addresses errors of over-generalization and their subsequent disappearance in early child language acquisition. In *Eighth Annual Conference of the Cognitive Science Society*, University of Massachusetts, Amherst, 1986, pp407-419.

- Hill, J. (1987) Modelling Concept Acquisition. Unpublished. Mental Lexicon Workshop, Centre for Cognitive Science, University of Edinburgh..
- Hollan, J. D. (1975) Features and semantic memory: set theoretic or network model? *Psychological Review*, **82**, 154-155.
- Hopper, P. J. (ed.) (1982) *Tense-Aspect: between semantics & pragmatics*, Volume 1. Amsterdam: John Benjamins Publishing Co. From a 1979 conference at UCLA.
- Jackendoff, R. (1978) Grammar as evidence for conceptual structure. In Halle, M., Bresnan, J. and Miller, G. (eds.) *Linguistic Theory and Psychological Reality*. Cambridge, Mass.: MIT Press.
- Jackendoff, R. J. (1983) *Semantics and Cognition*. MIT.
- Janus, R. A. and Bever, T. G. (1985) Processing of metaphoric language: An investigation of the three-stage model of metaphor comprehension. *Journal of Psycholinguistic Research*, **14**, 473-487.
- Jespersen, O. (1949) *A Modern English Grammar*, Volume VII. London: George Allen and Unwin.
- Jespersen, O. (1914) *A Modern English Grammar*, Volume II. Heidelberg: Streitberg. Syntax. First volume.
- Johnson-Laird, P. N., Gibbs, G. and de Mowbray, J. (1978) Meaning, amount of processing, and memory for words. *Memory and Cognition*, **6**, 372-375.
- 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.

- Johnson-laird, P. and Garnham, A. (1979) Description and discourse models. *Linguistics and Philosophy*, **3**, 371-393.
- Jones, G. V. (1982) Stacks not fuzzy sets: an ordinal basis for a prototype theory of concepts. *Cognition*, **12**, 281-290.
- Just, M. and Carpenter, P. A. (1980) A theory of reading: From eye fixations to comprehension. *Psychological Review*, **87**, 329-354.
- Karmiloff-Smith, A. (1979) *A Functional Approach to Child Language: A study of determiners and reference*. Cambridge: Cambridge University Press.
- Katz, J. J. and Fodor, J. A. (1963) Structure of a semantic theory. *Language*, **39**, 170-210.
- Keenan, E. L. (1978) The syntax of subject-final languages. In Lehman, W. P. (ed.) *Syntactic Typology: studies in the phenomenology of language*, pp268-327. Austin, Texas: University of Texas Press.
- Keil, F. C. (1979) *Semantic and Conceptual Development: An Ontological Perspective*. Cambridge, Mass.: Harvard University Press. see also Gerard & Mandler (1983).
- Keil, F. C. (1983) Inferences and the acquisition of word meaning. In Seiler, T. B. and Wannemacher, W. (eds.) *Concept Development and the Development of Word Meaning*, pp103-125. Berlin: Springer-Verlag.
- Keil, F. C. (1984) Mechanisms in cognitive development and the structure of knowledge. Chapter 4 in Sternberg, R. J. (ed.) *Mechanisms of Cognitive Development*, pp81-101. New York: W H Freeman and Company.
- Kempson, R. M. (1977) *Semantic Theory*. Cambridge: Cambridge University Press.

- Kilby, D. (1984) *Descriptive Syntax and the English Verb*. London: Croom Helm.
- Kintsch, W. (1974) *The Representation of Meaning in Memory*. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Kittay, E. and Lehrer, A. (1981) Semantic fields and the structure of metaphor. *Studies in Language*, **5**, 31-63.
- Kunzendorf, R. G. (1976) Selected features of word meaning. *Journal of Verbal Learning and Verbal Behavior*, **15**, 633-640.
- Labov, W. (1972) The boundaries of words and their meanings. In Bailey, C. N. and Shuy, R. W. (eds.) *New Ways of Analyzing Variation in English*, Volume 1. Washington: Georgetown University Press. (Papers from conference, 1972).
- Lackner, J. and Garrett, M. (1973) Resolving ambiguity: effects of biasing context in the unattended ear. *Cognition*, **1**, 359-372.
- Ladefoged, P. (1971) *Preliminaries to linguistic phonetics*. Chicago: University of Chicago Press.
- Lakoff, G. (1982) Categories and Cognitive Models. Cognitive Science Report No. 2, University of California at Berkeley, November, 1982.
- Lakoff, G. (1984) Classifiers As a Reflection of Mind: A Cognitive Model Approach to Prototype Theory. Cognitive Science Report No. 19, University of California at Berkeley, January, 1984.
- Langacker, R. W. (1982) Remarks on English aspect. In Hopper, P. J. (ed.) *Tense and Aspect*, pp265-303. Amsterdam: John Benjamins Publishing Co.
- Langacker, R. W. (1982) Space Grammar, Analysability, and the English Passive. *Language*, **58**, 22-80.

- Langacker, R. W. (1983) *Foundations of Cognitive Grammar 1*:. Paper no. 100.
- Lass, R. (1976) *English Phonology and Phonological Theory: Synchronic and diachronic studies*. Cambridge: Cambridge University Press.
- Leech, G. (1969) *Towards a Semantic Description of English*. London: Longman.
- Leech, G. (1981) *Semantics*, 2nd. Harmondsworth: Penguin.
- Lehrer, A. (1985) *Wine and Conversation*. Bloomington, Indiana: Indiana University Linguistics Club.
- Lehrer, A. (1986) English classifier constructions. *Lingua*, **68**, 109-148.
- Lipka (1972) *Semantic Structure and Word-formation: Verb—particle constructions in contemporary English*. Munchen: Wilhelm Fink Verlag.
- Lipka, L. (1987) Metaphor and Metonymy as Productive Processes on the Level of the Lexicon.. Paper to XIV Internationaler Linguistenkongress, Berlin.
- Lyons, J. (1963) *Structural Semantics*. Oxford: Philological Society.
- Lyons, J. (1968) *Introduction to Theoretical Linguistics*. Cambridge: Cambridge University Press.
- Lyons, J. (1977) *Semantics*. Cambridge: Cambridge University Press.
- Lyons, J. (1981) *Language, Meaning and Context*. London: Fontana.
- MacKay, D. G. and Bever, T. G. (1967) In Search of Ambiguity. *Perception and Psychophysics*, **2**, 193-200.

- Marslen-Wilson, W. D. (1973) Linguistic Structure and Speech Shadowing at Very Short Latencies. *Nature*, **244**, 522-523.
- Marslen-Wilson, W. and Welsh, A. (1978) Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, **10**, 29 - 63.
- Marslen-Wilson, W., Tyler, L. and Seidenberg, M. (1978) Sentence processing and the clause boundary. In Levelt, W. J. and d'Arcais, G. F. (eds.) *Studies in the Perception of Language*. Chichester: John Wiley and Sons.
- Marslen-Wilson, W. and Tyler, L. K. (1980) The Temporal Structure of Spoken Language Understanding. *Cognition*, **8**, 1-74.
- Marslen-Wilson, W. D. (1986) Parallel Processing in Spoken Word-Recognition. *Cognition*.
- Marx, W. (1983) The meaning-confining function of the adjective. In Rickheit, G. and Bock, M. (eds.) *Psycholinguistic Studies in Language Processing*, pp70-81. Berlin: de Gruyter.
- Matthews, P. H. (1981) *Syntax*. Cambridge: Cambridge University Press.
- Medin, D. L. and Smith, E. E. (1984) Concepts and Concept Formation. In *Annual Review Psychology*, Volume 35, pp113-138.
- Miller, G. A. and Isard, S. (1963) Some perceptual consequences of linguistic rules. *Journal of Verbal Learning and Verbal Behavior*, **2**, 217-228.
- Miller, G. A. and Johnson-Laird, P. N. (1976) *Language and Perception*. Cambridge: Cambridge University Press.
- Miller, G. A. (1977) Practical and Lexical Knowledge. Chapter 24 in Johnson-Laird, P. N. and Wason, P. C. (eds.) *Thinking: Readings in Cognitive Science*, pp400-410. Cambridge: Cambridge University Press.

- Miller, G. A. (1978) Semantic relations among words. In Halle, M., Bresnan, J. and Miller, G. A. (eds.) *Linguistic Theory and Psychological Reality*, pp60-118. MIT.
- Minsky, M. (1975) Frame-system theory. In Schank, R. and Nash-Webber, B. L. (eds.) *Theoretical Issues in Natural Language Processing*, Cambridge, Mass, June 10-13, 1975.
- Moens, M. and Steedman, M. (1986) Temporal Information and Natural Language Processing. ACORD Deliverable 2.5, Centre for Cognitive Science, Edinburgh.
- Moens, M. (1987) Tense, Aspect and Temporal Reference. PhD Thesis, Centre for Cognitive Science, University of Edinburgh.
- Morton, J. and Long, J. (1976) Effect of word transitional probability on phoneme identification. *Journal of Verbal Learning and Verbal Behavior*, **15**, 43-51.
- Mourelatos, A. P. D. (1978) Events, Processes and States. *Linguistics and Philosophy*, **2**, 415-434.
- Murphy, G. L. and Wright, J. C. (1984) Changes in conceptual structure with expertise: differences between real-world experts and novices. *Journal of Experimental Psychology (Learning, Memory and Cognition)*, **10**, 144-155. lexicon concept.
- Murphy, G. L. and Medin, D. L. (1985) The role of theories in conceptual coherence. *Psychological Review*, **92**, 289-316.
- Neisser, U. (1967) *Cognitive Psychology*. New York: Appleton-Century-Croft.
- Neisser, N. (1976) *Cognition and Reality: Principles and Implications of Cognitive Psychology*. San Francisco: W. H. Freeman.
- Nelson, K. (1983) Concept development. In Seiler, T. B. and Wannenmacher, W. (eds.) *Concept Development and the Development of Word Meaning*, pp160-188. Berlin: Springer-Verlag.

- Nida, E. (1975) *Componential Analysis of Meaning*. The Hague: Mouton.
- Nunberg, G. (1978) *The Pragmatics of Reference*. Bloomington, Indiana: Indiana University Linguistics Club.
- Nunberg, G. (1979) The Non-uniqueness of Semantic Solutions: Polysemy. *Linguistics and Philosophy*, **3**, 143-184.
- Oden, G. C. (1977) Integration of fuzzy logical information. *Journal of Experimental Psychology (Human Perception and Performance)*, **3**, 565-575.
- Ortony, A., Schallert, D. L., Reynolds, R. E. and Antos, S. J. (1978) Interpreting metaphors and idioms: some effects of context on comprehension. *Journal of Verbal Learning and Verbal Behavior*, **17**, 465-477.
- 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**, 299-318.
- Osherson, D. N. (1987) A similarity metric. *Journal of Mathematical Psychology*.
- Palmer, F. R. (1976) *Semantics*. Cambridge: Cambridge University Press.
- Parsons, T. (1970) Mass terms and amount terms. *Foundations of Language*, **6**, 353-388.
- Pelletier, F. J. (ed.) (1979) *Mass Terms: Some Philosophical Problems*. Dordrecht: D. Reidel.
- Piaget, J. (1926) *The Language and Thought of the Child*. London: Routledge and Kegan Paul.

- Piaget, J. and Inhelder, B. (1968) *Memory and Intelligence*.
- Pichert, J. W. and Anderson, R. C. (1977) Taking different perspectives on a story. *Journal of Educational Psychology*, **69**, 309-315.
- Poutsma, H. (1926) *A Grammar of Late Modern English*, Volume II. Groningen: P. Noordhoff.
- Quine, W. V. (1960) *Word and Object*. Cambridge, Mass.: MIT Press.
- Radford, A. (1981) *Transformational Syntax: A student's guide to Chomsky's Extended Standard Theory*. Cambridge: Cambridge University Press.
- Reichgelt, H. A review of Hans Kamp's "A theory of truth and semantic representation". Unpublished paper, Centre for Cognitive Science, University of Edinburgh.
- Rips, L. J., Shoben, E. J. and Smith, E. E. (1975) Set theoretic and network models reconsidered. *Psychological Review*, **82**, 156-157.
- Robinson, R. (1954) *Definition*. Oxford: Oxford University Press.
- Romaines, S. (1983) On the productivity of word-formation rules and the limits of variability in the lexicon. *Australian Journal of Linguistics*, **3**, 177-198.
- Rosch, E. H. (1973) On the internal structure of perceptual and semantic categories. In *Cognitive Development and the Acquisition of Language*. New York: Academic Press. First extension beyond colour terms of prototype theory..
- Rosch, E. and Mervis, C. B. (1975) Family Resemblances: Studies in the Internal Structure of Categories. *Cognitive Psychology*, **7**, 573-605.
- Roth, E. M. and Shoben, E. J. (1983) The effect of context on the structure of categories. *Cognitive Psychology*, **15**, 346-378.

- Roth, E. M. and Mervis, C. B. (1983) Fuzzy set theory and class inclusion relations in semantic categories. *Journal of Verbal Learning and Verbal Behavior*, 22, 509-525.
- Rumelhart, D. E. (1979) Some problems with the notion of literal meanings. In Ortony, A. (ed.) *Metaphor and Thought*, pp78-90. Cambridge: Cambridge University Press.
- Rumelhart, D. and Norman, D. The Active Structural Network. In Norman, D. and Rumelhart, D. (eds.) *Explorations in Cognition*. San Francisco: W. H. Freeman.
- Ryle, G. (1954) *Dilemmas*. Cambridge: Cambridge University Press.
- Sadock, J. M. (1979) Figurative speech and linguistics. In Ortony, A. (ed.) *Metaphor and Thought*, pp46-63. Cambridge: Cambridge University Press.
- Sakamoto, Y., Ishikawa, T. and Satoh, M. (1986) Concept and structure of semantic markers for machine translation in Mu-Project. In *Proceedings of the 11th International Conference on Computational Linguistics and the 24th Annual Meeting of the Association for Computational Linguistics*, Institut fuer Kommunikationsforschung und Phonetik, Bonn University, Bonn, 25-29 August, 1986, pp13-20.
- Sanford, A. J. and Garrod, S. C. (1981) *Understanding Written Language*. Chichester: John Wiley and Sons.
- Sapir, E. (1951) The function of an international auxiliary language. In Mandelbaum, D. G. (ed.) *Selected Writings of Edward Sapir in Language, Culture and Personality*, pp110-121. Berkeley and Los Angeles: University of California Press. First published in 1931.
- Saussure, F. (1966) *Course in General Linguistics*. New York: McGraw Hill. Translated by Wade Baskin, first published 1916.

- Schank, R. C. and Abelson, R. P. (1977) Scripts, Plans, and Knowledge. In Johnson-Laird, P. N. and Wason, P. C. (eds.) *Thinking*. Cambridge: Cambridge University Press.
- Schank, R. C., Collins, G. C. and Hunter, L. E. (1986) Transcending inductive category formation in learning. *Behavioural and Brain Sciences*, **9**, 639-686.
- Searle, J. R. (1979) Metaphor. In Ortony, A. (ed.) *Metaphor and Thought*, pp92-123. Cambridge: Cambridge University Press.
- Sharkey, N. E. and Sharkey, A. J. C. (1987) What is the point of integration? The loci of knowledge-based facilitation in sentence processing. *Journal of Memory and Language*, **26**, 255-276.
- Shillcock, R. C., Altmann, G. T. M. and Bard, E. G. Subsequent context affects word recognition in spontaneous speech.
- Sinclair, J. M. (ed.) (1987) *Looking Up: An account of the COBUILD Project in lexical computing*. London: Collins ELT.
- Slobin, D. I. (1982) Universal and particular in the acquisition of language. In Wanner, E. and Gleitman, L. R. (eds.) *Language Acquisition: the State of the Art*, pp128-170. Cambridge: Cambridge University Press.
- Smith, E. E., Shoben, E. J. and Rips, L. J. (1974) a featural model for semantic decisions. *Psychological Review*, **81**, 214-241.
- Smith, N. V. (1980) Review of Lehman (1978). *Journal of Linguistics*, **16**, 150-154.
- Strawson, P. F. (1952) *Introduction to Logical Theory*. London: Methuen.
- Strawson, P. F. (1967) On referring. In *Contemporary Readings in Logical Theory*, pp105-126. London: Macmillan.

- Strohner, H. and Nelson, K. E. (1974) The young child's development of sentence comprehension. *Child Development*, **45**, 567-576.
- Tourganeau, R. and Sternberg, R. J. (1982) Understanding and appreciating metaphors. *Cognition*, **11**, 203-244.
- Townsend, D. J. and Bever, T. G. (1982) Natural units of representation interact during sentence comprehension. *Journal of Verbal Learning and Verbal Behavior*, **21**, 688-703.
- Trier, J. (1931) *Der Deutsche Wortschatz im Sinnbezirk des Verstandes: die Geschichte eines Sprachlichen Feldes*. Heidelberg.
- Ullmann, S. (1962) *Semantics: An Introduction to the Science of Meaning*. New York: Barnes and Noble.
- Vendler, Z. (1957) Verbs and Times. *Philosophical Review*, **66**, 143-160.
- Verschueren, J. (1981) Problems of lexical semantics. *Lingua*, **53**, 317-351.
- Vygotsky, L. S. (1962) *Thought and Language*. MIT. First published (posthumously) in 1934.
- Ware, R. X. (1975) Some bits and pieces. *Synthese*, **31**, 379-393.
- Warren, R. M. and Sherman, G. (1974) Phonemic restorations based on subsequent context. *Perception and Psychophysics*, **16**, 150-156.
- Wason, P. C. (1977) Self-contradictions. In Johnson-Laird, P. N. (ed.) *Thinking: Readings in Cognitive Science*, pp114-128. Cambridge: Cambridge University Press.

- Weinreich, U. (1963) On the Semantic Structure of Language. In Greenberg, J. H. (ed.) *Universals in Language*, pp142-163. Cambridge, Mass.: MIT Press.
- Weinreich, U. (1963) On the semantic structure of language. Chapter 7 in Greenberg, J. H. (ed.) *Universals of Language*, 2nd edition, pp142-217. Cambridge, Mass.: MIT Press. Only read pages 154-158 where he discusses deixis.
- Weinreich, U. (1966) Explorations in Semantic Theory. Chapter 7 in Sebeok, T. A. (ed.) *Theoretical Foundations*, Volume III, pp395-477. The Hague: Mouton.
- Wittgenstein, L. (1953) *Philosophical Investigations*. Oxford: Basil Blackwell. Translated by G E M Anscombe.
- Wundt, W. (1894) *Logik*, Volume II: *Eine Untersuchung der Prinzipien der Erkenntnis*. Stuttgart: Ferdinand Enke. A study of the principles of knowledge."
- Zadeh, L. (1965) Fuzzy sets. *Information and Control*, **8**, 338-353.
- Zadeh, L. A. (1982) A note on prototype theory and fuzzy sets. *Cognition*, **12**, 291-297.
- Zaslow, R. W. (1961) A study of concept formation in normals, mental defectives and brain damaged adults. *Genetic Psychology Monographs*, **63**, 279-338.
- Zeevat, H., Klein, E. and Calder, J. (1987) An Introduction to Unification Categorical Grammar. In Haddock, N. J., Klein, E. and Morrill, G. (eds.) *Edinburgh Working Papers in Cognitive Science*, Volume 1: *Categorical Grammar, Unification Grammar, and Parsing*.
- Zemach, E. (1979) Four ontologies. In Pelletier, F. J. (ed.) *Mass Terms: Some Philosophical Problems*, pp63-80.
- Ziff, P. (1964) About grammaticalness. *Mind*, **LXXIII**.

Zwicky, A. M. and Sadock, J. M. (1975) Ambiguity tests and how to fail them. In Kimball, J. P. (ed.) *Syntax and Semantics*, Volume Vol. 4, pp1-36. New York: Academic Press.

Appendix A.

Materials randomisation program listing

These functions are intended to do some randomizing, and then to assemble data files from the code numbers generated.

```
(defun Fire (subjlist anslist) ;This arg is a list of files to fill.
  (ports) ;Opens ports once & for all to Agendas
  (Filer subjlist anslist) ;and corresponding ESIT's.

(defun Filer (subjlist anslist)
  (cond((null subjlist) ;Once there are no files left
        (clam) ;close the ports & go home
        (t(EFwrite (read esitport) ;else, get to work on filling
                   (car subjlist)(car anslist)))))) ;the first file.

(defun EFwrite (ESIT subjfile ansfile)
  (setq output (fileopen subjfile 'w)) ;open up the newfile for writing
  (setq ansport (fileopen ansfile 'w)) ;and one to record approp responses
  (Falseopen) ;open up port to the false items
  (Copy (read agendaport) ESIT)) ;set up a recursion over Agenda
```

;ESIT can be plugged in by hand, subj by subj. It's a list of nos, the same length as the number of true items to be finally included. Adjacent nos specify which instances of an item are to be used. Successive pairs of numbers are associated with successive items in MAINLIST. If only one instance of each item is to be used, then an amended version would have to be used.

;Agenda required more complex randomizing, so read it from Aitch.

```
(defun Copy (Agenda ESIT)
  (if (null Agenda)
      then (clam)
      else (let((item (add1(quotient(car Agenda)2)))
                (specifier (nth (car Agenda) ESIT)))
            (print (cdr(nth (sub1 specifier)
                            (nth (sub1 item) MAINLIST)))
                  output)
            (terpr output)
            (print (cdr(nth 6
```

```

                                (nth (sub1 item) MAINLIST)))
                                output)
                                (terpr output)
                                (print 1 ansport)
                                (terpr ansport)
                                (readf))) ;collect a variable number of false items
                                        ;before returning to true items.

(defun readf ()
  (if (eq (Toss) 1)
      then(let((lie (read falseport)))
            (if lie
                then(print lie output) ;send the first false item
                    (terpr output) ;newline
                    (print (read falseport) output) ;get the property
                    (terpr output) ;newline
                    (print 0 ansport)
                    (terpr ansport)
                    (readf)
                    else(Copy (cdr Agenda) ESIT)))
            else(Copy (cdr Agenda) ESIT))) ;termination - get a true

(defun rollo ()
  (let ((tale (read falseport)))
      (if tale
          then (print tale output)
              (terpr output)
              (print (read falseport) output)
              (terpr output)
              (print 0 ansport)
              (terpr ansport)
              (rollo)
              else (close falseport))))

(defun Toss ()
  (Trandom 2))

(defun Falseopen () ;This opens a channel to file
  (setq falseport (fileopen 'Falsies 'r)) ;containing the false items.

(defun clap ()
  (rollo) ;tip residual falsies into the file,
  (close output) ;close this particular subjfile, then
  (close ansport)
  (Filer (cdr sublist)(cdr anslist)) ;recycle with rest of the sublist

```



```

(defun clam ()
  (close agendaport)
  (close esitport)) ;tidies up at the end.

(defun ports ()
  (setq agendaport (fileopen 'Aitch 'r))
  (setq esitport (fileopen 'East 'r)))

(load 'sents.l) ;This binds a table containing the true items and their
                ;associated properties to the global variable MAINLIST.

(def Trandom
  (lambda (n)
    (if n
        then (abs(mod(rot(random)16)n))
        else (rot(random)16))))

(defun Rord (n lits) ;builds a list ("lits") by randomly
                    ;ordering the nos between 0 & n.
  (if (eq (length lits) n)
      then (print lits fvar)
          (terpr fvar)
      else (let ((N (Trandom n))) ;Set "N" to one of the nos, but
              (if (memq N lits) ;if it's already in "lits"
                  then (Rord n lits) ;then try again with another number
                  else (Rord n (append (ncons N) lits)))))) ;else stic it on lits
          ;you have to open
          ;an out file for
          ;yourself first.

```

Plain old "close" does the job when you've finished. There's an upper bound on the size of the list you can randomly order with this function: between 40 and 50 on this version, beyond which a namestack overflow occurs.

Understanding the Lexicon

Meaning, Sense and World Knowledge
in Lexical Semantics

Edited by
Werner Hüllen and Rainer Schulze

Sonderdruck

Max Niemeyer Verlag
Tübingen 1988



GEORGE L. DUNBAR

TERRY F. MYERS

CONCEPT COMBINATION AND THE CHARACTERIZATION OF LEXICAL CONCEPTS¹

0. This paper examines the view that the meaning of a word can be characterized as a conceptual prototype.² This view has gained currency because of its advantages over the so-called classical view of concepts in relation to certain well-known psychological results. Here, it is argued that although goodness-of-exemplar ratings indirectly reflect the meaning of a word, there are some good reasons for doubting whether the semantic part of a lexical entry, the lexical concept, should be characterized in such terms. Four arguments will be provided in support of this contention: (1) prototypes cannot be combined; (2) some properties inherent in prototypes are not inherent in lexical items; (3) prototypes are not coherent in isolation from cognitive models; and (4) prototypes are not constant (i.e. their inferential relations vary). In sum, prototypes do not display those properties that would enable lexical concepts to play a computable role in a natural language understanding system. Let us examine each of these deficiencies in turn.

1 This paper is based on part of a chapter in Dunbar (forthcoming).

2 'Prototype' is defined here as in Dunbar (1987: 2): "a (possibly structured) set of default values along certain dimensions to which exemplars can be compared and rated for typicality in terms of some similarity metric (examples departing least from the defaults being the most prototypical)".

1. Syntactic non-compositionality of prototypes

One way of formalising the concept of prototype that has been adopted in the literature is Zadeh's 'fuzzy logic' (1965). The key attraction of fuzzy logic is that it replaces the characteristic functions of set theory, which are binary (an entity is either in the set or out of it; the predicate is either true or false of any particular entity; the 'middle' is excluded), with characteristic functions which assign values anywhere on a continuum between 0 and 1. Such functions offer an appealing analogy with distributions of 'goodness of exemplar' (GOE) ratings (Roth/Mervis 1983). Thus, it has been argued, degree of membership can be formalized in terms of the relative truth of the predication.

Translation rules function in a logic to indicate how, given the interpretations of its constituents, the interpretation of some construction may be determined. For example, in a standard predicate logic the set of objects which are *red squares* is defined by a translation rule as being the intersection of the set of objects which are red with the set of objects which are square.

One contention of critics is that the corresponding function in fuzzy logic telling you how relatively true it is that some object is a *red square* from information about how red and how square it is, cannot be defined. Osherson/Smith (1981) examine fuzzy intersection in relation to the concept *striped apple*, which they take to be intersective. The obvious translation rule in fuzzy logic (the 'min' rule) entails that the 'goodness' of a complex cannot exceed its goodness as an example of either constituent concept. But, argue Osherson/Smith (1981), a *striped apple* will be a better exemplar of the category *striped apple* than of *apple*, and so this translation rule is not adequate.

Oden (1977) proposed as an alternative a multiplicative rule, which he attributes to Goguen (1969), whereby the GOE rating of a complex concept will be the product of the GOE ratings of its constituents. To test this he asked subjects to judge the truth of pairs of sentences, for example *robins are birds* and *beds are furniture*. He systematically varied the degree of rated truth of the individual statements in such pairs and attempted to predict the truth of their conjunction. What he found was an interaction between the ratings of the constituent statements. This could be better modelled by the multiplicative rule than by the minimum rule. Although the multiplicative rule is a better match to subjects' intuitions about complex statements, it too has serious shortcomings. For example, an object cannot have a higher rating with respect to

how good a member of the conjunction of A and B it is than its rating for exemplariness in either of the constituents. Indeed its GOE rating must be lower, unless both of its constituent ratings are 1. Hence this rule, like the minimum rule, cannot allow any striped apple to be a better striped apple than it is an *apple*.

Osherson/Smith (1982) sketch a proof that, in fact, no simple function of the constituents can adequately characterize their conjunction. This is done by imagining a cube and a sphere which metamorphose in such a way that each gradually approaches the other's shape. At the half-way point, the object should be equally as good a *ball* as a *cube*. If the goodness of something's membership in a conjunction is a simple function of its goodnesses in the constituent categories, then this entity should be a *round ball* to the same degree that it is a *round cube* - a conclusion that violates one's intuitions. Thus, no simple function of constituent goodnesses can predict correctly the goodness of an exemplar in a complex category.

By way of response to these arguments, Zadeh (1982) discusses the 'normalization' of complex concepts (raising the maximum value of the set to 1 by dividing all values by the pre-normalization maximum). This is said to be brought about through focussing on an intersection "by giving it a label" (1982: 291). This makes it technically feasible for a member of this set to have a higher degree of membership of it than of either of the constituent sets. Like ranking (Jones 1982) this result is achieved by considering the 'absolute' goodness of members of a complex category to be a function of their relative goodness with respect to one another independently of their goodness in the constituent categories, while deriving that relative goodness from goodness in constituent categories. Both work (to oversimplify) by assuming that in any category there will be a 'best' member, and stipulating that that member's rating for goodness should be 1. The crucial difficulty with normalization, pointed out by Osherson/Smith (1982), is that, if it is based on the 'min' rule, it only takes account of one half of the conjunct. For example, two red squares will be equally good as *red squares* so long as their values for the poorer of these properties are equal. Thus, if they are equally poor shades of red, even if one is a worse square than the other, they will be equally good *red squares* (so long as both are better 'squares' than 'reds'). This is counter-intuitive. Osherson/Smith (1982) also demonstrate that if a multiplicative rule underpinned normalization, then if the best *red flower* were a perfect flower it could not be a perfect red thing, because its goodness as a *red flower* would have to be less than 1 in

order to prevent the product of the constituent goodnesses of any *red flower* equalling their normalized product, and hence the goodness of the complex concept. But, these authors ask, what contingency is there among properties that impels no entity to be perfect in more than a single respect?

Jones' (1982) attempt to tackle these issues founders on technical problems associated with the need to rank entities before determining their typicality (Osherson/Smith 1982). However, even if ranks could be assigned, there are instances of concept conjunction that Osherson/Smith (1982) assert would defeat the scheme. Jones' rule could not permit an entity which is *prima* in two categories to be a better exemplar of their conjunction. But, although Paul Newman [was] both the best actor and the most handsome entity, he [was] more highly rated as a *handsome actor*. Paragons, it would seem, rise above the canons of logic.

Cohen/Murphy (1984: 51) describe concept combination as:

a process of combining representations according to certain generative rules within a domain of knowledge.

They take from the knowledge-representation literature in AI the notion that sub-concepts may specialize a super-concept by role modification (nouns are treated as frame-like, having slots called 'roles'). They emphasise that the process is "asymmetric [and] knowledge-dependent" (1984: 51), where by 'knowledge-dependent' they mean, for example, that which role gets modified will depend on context, different contexts promoting some roles for modification more strongly than others. They argue that domain knowledge also modifies, for example, some of the role values of *pet* when it is used in *pet fish*. Thus, knowledge of the world influences the typicality structure of complex concepts which consequently, they continue, will not be predictable on the basis of the constituent concepts alone. They also note that the effect of a role value specification on typicality can depend on the setting of other roles. For example, the effect of baldness on a man's typicality depends on his age. Effects of correlational structure such as this would, then, have to be considered in evaluating typicality with respect to complex concepts, but this would entail further use of domain knowledge.

A crucial problem facing any approach to conceptual combination is that the two terms in the conjunction may not be interchanged without altering the typicality rating of an exemplar in the complex concept. Although the 'min' rule is undoubtedly wrong to consider only one of the constituents in determining exemplariness in the complex concept, it is not the case either that

each contributes equally, as would be predicted by the multiplicative rule. As Hampton (1988) has recently shown, for example, a constituent contributes more to the determination of exemplariness in a complex concept when it is playing the role of modifier than when it is playing the role of the head. His study used expressions of the type: noun plus relative clause modifier, for example, 'Games which are sports'. He found that a given instance (for example *chess*) might be rated more highly in relation to this form of the putative conjunction than the other (i.e. 'Sports which are games'). Moreover, whenever there is a 'head-modifier' relationship between the terms, the head is typically the one that determines the superordinate category.³ Thus, for example, a *cow catcher* is a type of catcher, and not a sort of cow; a *milk bottle* is a kind of bottle, and not a dairy product; and a *pancake* is pottable but not a pot. Such non-commutativity confounds the attempt to model the combination of prototypes by means of fuzzy logic.

Thus there appear to be two difficulties confronting attempts to model the combination of concepts by means of fuzzy logic: first, the knowledge-laden nature of combination, and second its asymmetry.

2. Lexical indeterminacy of types

Our second reason to doubt that lexical concepts have a prototype structure is motivated by the finding that some properties inherent in prototypes are not inherent in lexical items. We will consider two sources of evidence, one deriving from the work of Nunberg (1979), the other from Jackendoff (1983).

Nunberg's first, obvious, point is that many words can be used to refer to more than one sort of referent. He argues that this is a very general process that is not adequately handled by trying to list all the different senses. He proposes instead to list only a few senses from which other uses may be derived by general pragmatic principles. The principles may be summarised in the following way: the designatum, or basic sense of a word, can be used to refer to

3 Hampton (1988) has produced tentative evidence that entities may be treated subjectively as members of a complex category even if they are regarded by the same subjects as non-members of the constituent categories.

an entity if the cue validity of the referential function (RF) mapping from the designatum to the entity is sufficiently high to clearly and unambiguously pick out the entity in question from any alternatives. 'Cue validity' is the exactness with which cues (in this case, referential functions) pick out a given referent. Take, for example, an exam script. This object may be considered with respect to two relations: firstly, it is related to the set of examination candidates, and secondly it is related to a set of markers. Now candidates produce only one script while markers handle more than one, and indeed, more than one marker may be involved with a single script. Therefore, the referential function from the exam script to the candidate will have a higher cue validity than the referential function from the exam script to the marker. Consequently, if a speaker says, while pointing to the exam script, *He has a good grasp of the subject*, he will refer to the candidate. (Given certain assumptions about the background knowledge of speaker and hearer.)

Nunberg's argument continues as follows: Just as we can talk about relative cue validity of referential functions in cases of deferred ostension, such as the one we have just described, we can also use relative cue validity to determine which senses of a word are most basic. Consider the word form *Dylan*. This can be used to refer either to a singer, Bob Dylan, or to his work, the songs of Bob Dylan. Which of the two is more basic? In order to compute this, we must determine the cue validity of the referential function from singer to song and the inverse referential function from song to singer. The more basic sense of *Dylan* is taken to be the one in the domain of the referential function with higher cue validity. Since the cue validity of the referential function from Dylan to his work is greater than the cue validity of its inverse, *Dylan* designates the man.

Nunberg next considers the cases in which we are most interested for the purposes of this paper, cases in which it is not clear which of the two referential functions has greater cue validity. That is, there are cases in which cue validity appears to be equally high in both directions. Of his several examples we will consider just one. It concerns the choice between a token or a type reading of a phrase, e.g. *three trees*. Either this refers to tokens (the trees themselves) or it refers to types (different species of trees). In this example, cue validity is the same for the referential function from trees to species of trees as it is for the inverse function from species to trees. That is, each tree is of just one species, and each species picks out just some trees and not others.

If Nunberg's arguments about cue validity are correct, then his conclusion with respect to the type-token distinction is of great interest and relevance to our own argument. Rumelhart (1987, in a talk to the PHLING Conference, Edinburgh) has recently observed that prototypes could as easily be termed 'types'. But if that is the case, choosing prototypes as the characterization of lexical concepts would imply that the type sense is more basic, with the token sense being derived from it. However, this is precisely the asymmetry that Nunberg argues we cannot empirically justify. We reach the same conclusion from a consideration of some work by Jackendoff (1983), though we arrive at it by a different route.

Jackendoff (1983) provides some evidence that the choice of a type or a token reading for an NP is determined through the satisfaction of a number of grammatical constraints. Thus, choice of a type or a token reading for the predicate NP (in English) depends upon its syntactic position; the choice of verb, and definiteness of the article. If the verb is *be* and the article indefinite (*Margaret Thatcher is a politician*), a type reading results. The definite article, on the other hand, gives rise to an ambiguity with respect to a token or a type reading (*Margaret Thatcher is the woman drinking a martini*) (cf. Donnellan's (1966) referential-attributive distinction). A number of verbs behave similarly (e.g. *resemble* and *become*). Others do not. When the verb is *buy*, for example, all predicate NPs express tokens, regardless of definiteness. Moreover, tense and aspect present a further source of constraint. Thus, if a verb is stative all its NP arguments must be indefinite for a type reading (cf. *A policeman has a truncheon* with *A policeman has the truncheon*). However, a verb that expresses an event must appear in simple present rather than present progressive for a type reading (cf. *A spider spins a web* with *A spider is spinning a web*).

What Jackendoff's evidence suggests is that the type/token distinction is not determined by a lexical feature of the head noun of a noun phrase. His ambiguous instances, in particular, suggest that neither sense is the more basic. In these cases the distinction is both lexically and syntactically under-determined and would need to be resolved at the level of discourse representation. A type reading does not derive from information inherent in a lexical entry. There is a sense in which a type classification may be seen to be more basic in Jackendoff's analysis, since both tokens and types incorporate features of the type in which they are a member. However, there is a reciprocal relationship between types and tokens: just as tokens presuppose types of which they are instances, so types presuppose tokens that exemplify them. Another way to

interpret the sense in which types might be more basic than tokens would take us into genetic epistemology, and therefore, well beyond the scope of this paper. For the present, we take Jackendoff's evidence as further reason to doubt that lexical concepts are prototypes.

3. Coherence of prototypes

Prototype structure has usually been described with respect to some similarity metric, where similarity has been taken as both the criterion of category membership and the principle underlying conceptual coherence. Things belong together in a category if they are sufficiently similar to the prototype. The prototype reflects the central tendency of instances in relation to the dimensions along which similarity is computed. However, this position has been challenged by those who believe conceptual coherence is determined by the theories people have about their physical and social worlds, and in terms of their current goals. According to Murphy/Medin (1985) GOE ratings are 'heavily constrained' by knowledge of the world, and conceptual coherence is derived from a theory that explains why things are the way they are.

As Miller (1978: 104) once wrote:

lexical knowledge ... is not isolated from [the] general conceptual system; the lexicon has a cognitive structure only because it is an integral part of everything a person knows and believes.

Earlier, Miller/Johnson-Laird (1976) had argued that semantic fields have a conceptual core that derives from the implicit theories shared by members of a linguistic community. In a similar vein, Lakoff (1987: 68) observes that our categories are not determined with respect to properties of the real world, but in terms of 'interactional properties' related to our exchanges with that world. And this is so even in the case of basic level categories. We "structure and make sense of our experience" in terms of cognitive models (roughly, what others have called variously 'images', 'frames' and 'scripts'). 'Cognitive models' are defined "relative to idealized circumstances". Consequently, idealized cognitive models are what give rise to the prototype structure of our categorizations, not the lexical concepts that cue them.

4. Instability of prototypes

Another doubt as to whether lexical concepts are prototypes arises from recent studies which have provided evidence that typicality structure is not a fixed property of a word, and that its focus can be shifted by context.

Roth/Shoben (1983) base a study on a finding of Garrod/Sanford (1977) that the speed of anaphora resolution depends on the semantic relatedness of anaphora and antecedent, as if membership involved category membership verification (when the antecedent is more general than the anaphora). Roth/Shoben examined whether reading times were affected by a single GOE structure in differing contexts, or whether context could affect speed of resolution by changing GOE structure. The finding was that resolution for typical exemplars (typical for the category in isolation) was slower when the context was biased to make that exemplar a bad fit than when the context was neutral. Thus *robin* is a poor exemplar in the context *The bird walked across the barn-yard*.

Barsalou/Sewell (1984, reported in Barsalou 1985) found that GOE structure can change dramatically when people take different perspectives on a concept. Their subjects rated exemplars (e.g. *robin*, *ostrich*, *swan*) for typicality as *birds*, but assigned them different degrees of centrality according to the cultural perspective ('American', 'African', or 'Chinese') they had been asked to assume. These results held for both taxonomic and goal-derived categories. Thus, people can flexibly generate different concepts in different contexts.

What various findings illustrate is that prototypes are unstable; they vary with context.

5. Summary

Four arguments have been presented indicating that lexical concepts are not idealized summary representations corresponding to prototypes. What then, are lexical concepts? Dunbar (forthcoming) has put forward a view of the relationship between lexicon and cognition that attempts to solve these difficulties. It is akin to some recent ideas proposed by Barsalou (1986) in its emphasis on the dynamic nature of conceptual structure. In comparison, however, it is more radically 'minimalist' (Cruse 1988). Problems arise, however, and it is not yet clear whether this approach will be viable, so for the moment we rest content to record the observation that prototypes as defined appear to be unstable, hard to combine, too specific and not independently coherent.

REFERENCES

- Barsalou, Lawrence W. 1985. Ideals, central tendency and frequency of instantiation as determinants of graded structure in categories. *Journal of Experimental Psychology (Learning, Memory and Cognition)* 11, 629-654.
- Barsalou, Lawrence W. 1986. Are there static category representations in long-term memory? *Behavioural and Brain Sciences* 9, 651-652.
- Cohen, Benjamin, Gregory L. Murphy. 1984. Models of concepts. *Cognitive Science* 8, 27-58.
- Cruse, D. Alan. 1988. Word meaning and encyclopedic knowledge. (In this volume).
- Donnellan, Keith S. 1966. Reference and definite descriptions. *Philosophical Review* 75, 281-304.
- Dunbar, George L. 1987. Prolegomenon to a cognitive science of lexical meaning. Edinburgh Ms.
- Dunbar, George L. forthcoming. The cognitive lexicon. Diss. Edinburgh.
- Garrod, Simon, Anthony J. Sanford. 1977. Interpreting anaphoric relations: The integration of semantic information while reading. *Journal of Verbal Learning and Verbal Behavior* 16, 77-90.
- Goguen, Joseph A. 1969. The logic of inexact concepts. *Synthese* 19, 325-373.
- Halle, Morris, Joan Bresnan, George A. Miller (eds.). 1978. *Linguistic theory and psychological reality*. Cambridge, Mass., London, MIT Press.
- Hampton, James A. 1988. Overextension of conjunctive concepts: Evidence for a unitary model of typicality and class inclusion. *Journal of Experimental Psychology (Learning, Memory and Cognition)* 14,
- Jackendoff, Ray. 1983. *Semantics and cognition*. Cambridge, Mass., London, MIT Press.
- Jones, Gregory V. 1982. Stacks not fuzzy sets: An ordinal basis for a prototype theory of concepts. *Cognition* 12, 281-290.
- Lakoff, George. 1987. *Women, fire, and dangerous things. What categories reveal about the mind*. Chicago, London, The University of Chicago Press.
- Miller, George A. 1978. Semantic relations among words. In: Morris Halle, Joan Bresnan, George A. Miller (eds.), 60-118.
- Miller, George A., Philip N. Johnson-Laird. 1976. *Language and perception*. Cambridge, Mass., The Belknap of Harvard University Press.
- Murphy, Gregory L., Douglas L. Medin. 1985. The role of theories in conceptual coherence. *Psychological Review* 92, 289-316.
- Nunberg, Geoffrey. 1979. The non-uniqueness of semantic solutions: Polysemy. *Linguistics and Philosophy* 3, 143-184.

- Oden, Gregg C. 1977. Integration of fuzzy logical information. *Journal of Experimental Psychology (Human Perception and Performance)* 3, 565-575.
- Osherson, Daniel N., Edward E. Smith. 1981. On the adequacy of prototype theory as a theory of concepts. *Cognition* 9, 35-58.
- Osherson, Daniel N., Edward E. Smith. 1982. Gradedness and conceptual conjunction. *Cognition* 12, 299-318.
- Roth, Emilie M., Carolyn B. Mervis. 1983. Fuzzy set theory and class inclusion relations in semantic categories. *Journal of Verbal Learning and Verbal Behavior* 22, 509-525.
- Roth, Emilie M., Edward J. Shoben. 1983. The effect of context on the structure of categories. *Cognitive Psychology* 15, 346-378.
- Zadeh, Lotfi A. 1965. Fuzzy sets. *Information and Control* 8, 338-353.
- Zadeh, Lotfi A. 1982. A note on prototype theory and fuzzy sets. *Cognition* 12, 291-297.