Polysemy

Adam Kilgarriff
D. Phil.
University of Sussex

December, 1992

Abstract

What does it mean to say a word has several meanings? On what grounds do lexicographers make their judgments about the number of meanings a word has? How do the senses a dictionary lists relate to the full range of ways a word might get used? How might NLP systems deal with multiple meanings? These are the questions the thesis addresses.

The 'Bank Model' of lexical ambiguity, in which polysemy is treated as homonymy, is shown to be flawed. Words do not in general have a finite number of discrete meanings which an ideal dictionary would list. A word has, in addition to its dictionary senses, an indefinite range of extended uses. The lexicographer describes only the uses which occur reasonably frequently and are not entirely predictable from the word's core meanings.

Polysemy is not a natural kind. It describes the crossroads between homonymy, collocation, analogy and alternation. (An alternation is a pattern in which a number of words share the same relationship between pairs of usage-types.) Any non-basic type of use for a word can be treated as belonging in one of these four camps. For computational lexicography, putative polysemous senses should be represented in the lexicon as homonyms, or within collocations, or, implicitly, as the outcome of applying an alternation. Uses in the 'analogy' camp will not be described in the lexicon.

As others have argued, an elegant theoretical description of lexical knowledge must be inheritance-based. It must be possible to state and inherit generalisations. Within such a lexicon, regularly polysemous and other predictable usage-types can be described concisely. The thesis presents two fragments of the lexicon in which various alternations and other aspects of lexical structure are given concise formal treatments in an inheritance-based lexical representation language.

Contents

PREFACE

There are four kinds of thesis in cognitive science: formal, empirical, program-based and discursive. Which sort was mine to be?

Let us consider the thesis, briefly, inside out. I started the research on a strictly empirical tack, and the fruits are to be found in the two central chapters, the factual core of the work. But that exposed many open questions about word senses, what they were, and how they might be identified. I was drawn into the discursive mode. The outcome is the two chapters encompassing that core.

Another product of the empirical work was intrigue at the more regular parts of the domain. What exactly were these regularities? Perhaps they could be exploited in an NLP system. The lexical representation language DATR was to hand, and was well suited to the task of making them explicit. The sense of a puzzle to be solved took over. As DATR is both a formal representation language and a language one can program in, I found myself writing both a program and a formal theory. That final stage resulted in the two following chapters. Add a couple of chapters of literature review to the front, top and tail with introduction and conclusion, and we have an agreeably symmetrical thesis which, I look round in delight to find, does a little bit of all the things a cognitive science thesis might do!

ACKNOWLEDGMENTS

I should like to thank my supervisors, Gerald Gazdar and Roger Evans, for guidance throughout the project, Roger particularly for his fabulously cunning solutions to all my DATR problems, Gerald for the unfailing intelligence of his comments on my many drafts – and the meticulous promptness with which they were delivered. I would also like to thank the many people in the School of Cognitive and Computing Sciences who have helped in all manner of ways, personal and academic, with theory and in practice, along the long and winding journey to a thesis. Amongst those to whom I owe most are Lynne Cahill, Robert Gaizauskas, Geoff Goodhill, Inman Harvey and Lionel Moser. I must also thank Longman for the use of their dictionary and permission to quote extensively from it: SERC for supporting me financially while the work was undertaken: and, last but far, far from least, Gill Lamden, for being there through all the bumps and starts along the way.

Chapter 1

Introduction

Many words have many meanings. The dictionary tells us so. But what does that mean? How many meanings does a word have, and what grounds do we or the lexicographers have for saying it is not more or less? And how is it that language-users can effortlessly comprehend and generate novel uses of words? What does that tell us about lexical structure? How might natural language processing computer systems deal with multiple meanings, or novel meanings?

These are the questions that motivate the thesis. The introductory chapter will spell out these questions in greater detail and show why we need a fuller understanding of polysemy; address the question, "What is polysemy?"; defend the kinds of methods used; take a tour of the thesis, sketching the methods and results of each stage of the research; and finally draw attention to the three principal claims the thesis makes.

1.1 Why is polysemy interesting?

1.1.1 All human knowledge ...

What is the structure of human knowledge? The question demands attention but is vast – far too vast to be directly researchable. The domain must be constrained: not 'knowledge' but some specific variety of knowledge. Some have taken knowledge of geometry, or geology, or arithmetic, or medical diagnosis: others have shifted focus from the knowledge itself to the words that are used to express it. Section 8.2 argues that lexical and general knowledge may share structure in some important respects.

So what is the structure of the human lexicon? But that is still a huge question. Words relate to other words in innumerable ways, and some corner of the whole must be selected for study. Again, the area of study must be reduced and focussed. Here, a methodological consideration comes to our assistance. A proven experimental technique is to hold as many factors as possible constant, in order that any observed variation can be attributed to a limited number of sources. In studying polysemy, we hold the form of the word constant, and then observe variations in meaning and distribution. Research into polysemy is one avenue for investigating the structure of lexical knowledge and hence, indirectly, the structure of human knowledge.

1.1.2 Wilks's problem

> LDOCE[1] is a full-sized dictionary containing over 55,000 entries, most of which contain multiple sense definitions. This level of real-world detail creates special problems of

[1] Longman Dictionary of Contemporary English

> scale for language processing systems. ... [E]ven a simple-seeming sentence like
>
> > There is a huge envelope of air around the surface of the earth.
>
> –considering only traditional content words (nouns, verbs, adjectives and adverbs)– represents a big space, because LDOCE contains 11 adverbial senses for *there*, 2 adjectival senses for *huge*, 14 nominal or verbal senses for *air*, 11 senses for *around*, 7 for *surface* and 12 for *earth*. Taken all together, and at the most crude level of analysis, this sentence is a a staggering 284,592-way ambiguous.
>
> This demonstration simply serves to show that a parser for text, accessing a realistic machine-readable language-resource like LDOCE, is faced with solving a large, and hard, problem. And not a problem *created* by large on-line dictionaries, rather a problem of *language* ... (Slator & Wilks, 1987, p 4–5)

Here is a second reason for studying polysemy. What is an NLP system to do when it goes to the dictionary to find the meaning of a word and finds several? Wilks's problem is one of the great obstacles lying in the way of wide-coverage natural-language computer systems. An account of polysemy is a prerequisite to removing that obstacle.

1.1.3 Creativity in language use

The fact that language users utter sentences that have never been uttered before has long been seen as one of the central facts linguistic theory must account for. It is a *sine qua non* of syntactic theory that it account for an indefinite number of possible sentences. But novelty in language use is not constrained to syntax. Language-users also use words in novel ways, the most dramatic being metaphor and metonymy. Novel uses of words are not arbitrary: they are constrained by the system of meanings from which they emerge. People are creative in their use of words by observing the rules and relationships that, given known meanings, generate unfamiliar but possible meanings. Relationships between actual meanings will tend to observe the same rules that hold between actual and possible meanings, so the study of polysemy is a route towards understanding how novelty and creativity in language use are possible.

1.1.4 The linguist's let-out clause

An important set of issues for linguistics concerns correlations between form and meaning. For example, "the BENEFICIAL case, typically flagged by *for*, can generally take the indirect object position if and only if the meaning of the verb entails the production of something" (Hirst, 1986, p 153). Thus

> Mary baked a cake for her.

can be restated as

> Mary baked her a cake.

but

> Mary collected the potatoes for her.

cannot be restated as

> *Mary collected her the potatoes.

Here, *bake* is a verb involving creating something and *collect* is not. Some verbs can be either: consider *paint*:[2]

Mary painted a picture for her.

can be restated as

Mary painted her a picture.

but

*Mary painted her a wall.

(in the painting-and-decorating rather than the picture-of-a-wall sense) is unacceptable. One way of maintaining the hypothesis is by arguing there are different senses of the verb, needing separate treatment, involved. But when can the "different sense" argument be resorted to? An account of polysemy is required.

1.2 What is polysemy?

First, a little terminology. A 'usage' will be used to mean a particular occurrence of a word in a particular context. If we have two different sentences containing a word, or even the same sentence on two different occasions, we have two usages. Whereas usages are tokens, 'usage-types' or 'uses' are types. Thus wherever we might wish to say that, in two usages, a word is being used in the same way, we say both usages exemplify a particular usage-type. Any set of usages of a word where all the members have some aspect of meaning in common, so there is some motivation for saying they all mean the same thing, is a usage-type. 'Senses' or 'word senses' are a subset of usage-types. 'Senses' are those usage-types which are or ought to be listed in a dictionary.

1.2.1 The SFIP criterion

On what basis, then, does a usage-type merit listing in a dictionary? The first consideration is frequency. If a type of use for a word is a one-off then it is not part of the lexicon. It may have served some particular communicative goal in some particular situation but that is not sufficient for it to be considered part of the lexicon. Publishers, faced with commercial constraints on the size of their dictionaries, have to take very seriously the issue of what is a 'sufficiently frequent' usage-type to gain admittance to a given dictionary.

A second consideration, less obvious but also essential, is predictability. Along a dimension of predictability, polysemy describes the middle zone. 'Homonymy' is to be found at the end of the scale where two usage-types for a given orthographic or phonological form are entirely unrelated and thus neither could be interpreted, or have its meaning predicted, on the basis of knowledge of the other. At the other end of the scale, where one usage-type is entirely predictable from another then the predictable usage-type is not worthy of space in the dictionary. The dictionary-user can predict it and everything about it so it would be superfluous to state it. For example, every noun with a usage-type denoting something visualisable may also be used to denote a representation of that 'something'. Thus *seal, splash* or *sunset* may equally denote a seal, splash or sunset in a picture or 'in the flesh'. Entirely predictable: entirely unworthy of note.

The two considerations combine in the 'SFIP criterion': a usage-type merits listing in a dictionary when it is Sufficiently Frequent and Insufficiently Predictable. The two parts interrelate. The more common a pattern, the stronger the prediction that further words will follow it. The

[2] *Bake* can also be either, but the contrast is clearer with *paint*.

thresholds of 'sufficiency' and 'insufficiency' will depend, for paper lexicography, on the size and target audience of the dictionary. In inheritance-based computational lexicons as described in this thesis, 'predictability' will be subsumed under a more general treatment of the inheritance structure of the lexicon.

Polsemy does not form any kind of 'natural kind'. It describes, rather, a crossroads. In one direction lies homonymy, in another – as in the 'highly predictable' case discussed above — metonymy. In others again, collocation and analogy. 'Collocation' describes those usage-types which only occur in the neighbourhood of one, or a small number of, other words, so the use is best described in a dictionary by giving the meaning of a multi-word unit and the question of whether there is a distinct sense for a particular word is side-stepped. 'Analogy' is used for those usage-types which are predictable, but the predictability stems from general knowledge and the situation of use rather than a rule which might be stated in the lexicon and considered a pattern of metonymy.

For each direction, there is no natural divide between polysemy and its neighbour. *Light*, of colour and of weight, may be considered homonymous or polysemous: *whisky*, used of the liquid ("a glass of whisky") or of a glass of it ("I'll have a whisky"), polysemy or metonymy: *light* in "travel light", polysemy or a collocation.[3] Polysemy is a concept at a crossroads and an investigation into it must be an investigation of the roads leading into and out of it.

1.3 A defence of the methodology

Before proceeding to a resumé of the contents of the thesis, some comments regarding the kinds of methods used are in order. Some would argue that the lexicon is an abstraction from the mental lexicon, which is best studied using psycholinguistic methods, so here we defend our approach against that attack.

1.3.1 A cognitive science perspective

> I believe, perhaps with undue pessimism, that the mind is too complicated to be seen clearly, or to be studied with advantage, from the perspective of a single discipline. The scientific understanding of cognition depends on a synthesis; [my research] is an attempt to bring together some of the ideas and methods of experimental psychology, linguistics and artificial intelligence. (Johnson-Laird, 1983, p xi)

We take the study of the lexicon to be intimately related to the study of the mind, and hold that these considerations also apply to it. For an understanding of the lexicon, the contributing disciplines are lexicography, psycholinguistics and theoretical, computational and corpus linguistics.

1.3.2 Conceptualism and realism

In some quarters, psycholinguistics is seen as having a privileged view of the lexicon, with other insights and information sources secondary. Thus, in building a lexical entry, Ilson & Mel'čuk (1989) say:

> Now, we believe that in contemporary English, BAKE is primarily a verb of cooking. This belief is based not on frequency but on psychological salience . . . (p 336)

In a paper on 'WordNet, a lexical database organized on psycholinguistic principles', we find:

[3] Simple examples of analogy are hard to come by, for reasons discussed in section 7.4.

> Pinker (1989, p 101) claims that speakers of English decompose verbs into such semantic subpredicates as CAUSE, GO BE and PATH, which enables them to predict the verbs' idiosyncratic behavior. Such an analysis may well be part of speakers' linguistic competence, but there is no evidence that it serves to organize the mental lexicon. (Fellbaum, 1990, p 282)

We would argue that lexical structure is part of linguistic competence, and it is a sound methodological principle that "speakers' linguistic competence ... serves to organize the mental lexicon". Beckwith, Fellbaum, Gross, & Miller (1991), in presenting their rationale for WordNet, say:

> by focussing on historical (diachronic) evidence the OED [Oxford English Dictionary] like virtually all other dictionaries, ignores the vast range of evidence for the synchronic organization of the lexicon that psycholinguists have gathered in this century. (p 212)

Their selection of an explicitly historical dictionary, the OED, followed by an aside that other dictionaries are similar, does little to suggest they acknowledge the contribution lexicography makes to our understanding of the synchronic lexicon. Learners' dictionaries in particular are concerned with synchronic description, and often present detailed accounts of the lexical relations for which Beckwith *et al.* cite psycholinguistic evidence. Collins COBUILD Dictionary (hereafter COBUILD), for example, has explicit mechanisms for noting synonyms, superordinates and antonyms.

While a carefully regulated set of psycholinguistic experiments, involving a range of individuals, might be a preferred method of exploring the lexicon, the resources needed to do this for large numbers of words are not available. Experimental techniques standardly involve a specific hypothesis about lexical organization or access, and use a small number of words. By contrast lexicographers, using less structured methods but with expertise accumulated and refined over generations, have covered all aspects of the behaviour of words, across the whole lexicon, in considerable detail, many times over. Even WordNet gathers all its information on most words from dictionaries. Dictionaries are inevitably a major resource for studying the lexicon of a language.

Pollard & Sag (1987) shed some light on the psycholinguistic preference:

> What sort of thing is a natural language object, say, the English word *cookie*? Is it a mental object, something which only exists in the minds of English speakers? Or is it a part of the real world, external to minds? This is one of the fundamental questions in the philosophy of language ... (p 1)

If words are mental objects, as conceptualism holds, then psycholinguistics is particularly well-placed for determining their nature. The perspective taken in this thesis is, rather, a 'realist' one, as described more fully in Chapter 4 (see also Pollard & Sag (1987, Chapter 1), Johnson-Laird (1983, Chapter 9)). While the relevance of psycholinguistic methods is not disputed, the empirical part of this thesis addresses the lexicon through 'language in the world' rather than 'language in the mind'. One merit of this is that corpus material, unlike psycholinguists' experimental results, is available in abundance.[4] Wherever the structure of the lexicon is reflected in the distribution of words, questions of structure may be resolved by looking at text, in bulk. A further theoretical consideration is that a speaker's lexical knowledge is derived from a lifetime of hearing and reading words used in various settings. All the linguistic information that feeds into lexical competence (though not the non-linguistic information) must therefore be available in a sufficiently large and appropriately-selected corpus.

In sum, the study of the lexicon, like that of the mind in general, will benefit from the perspectives of a range of disciplines. Lexicography and corpus linguistics are first class members of the club, along with psycholinguistics and theoretical and computational linguistics.

[4] At least for written material. It is increasingly available in bulk for spoken material also.

1.3.3 The benefits of formalism

The thesis contains empirical studies of dictionaries, theoretical discussions of word meaning and computationally oriented formal theories of polysemy in two small domains. The three-pronged approach to an area of enquiry is characteristic of cognitive science. The choice of empirical approach is explained above. The theoretical discussions need no further discussion, but the role of the formal theories may be unclear.

A theory stated in a formal language is likely to be more explicit, with its claims and consequences less subject to interpretation and ambiguity, than the same theory stated informally. While there is little point in attempting to formalise a domain before a theory concerning its structure is available, once there is such a theory, formalising it can only be a benefit. The predictions the formal theory makes will be explicit, and it will be correspondingly easier to see how the theory is falsifiable.

1.3.4 The benefits of modelling linguistic competence in an NLP system

A natural language forms a large and complex system. What a theory of the system has to account for is what people say or write, and when, at least in its linguistic aspects.[5] It is no simple matter to determine whether a theory of one part of the domain — for example, polysemy — accounts for a sample of the data. The theory will have to be embedded in a framework that includes other theories of other parts of the domain. Only the whole framework can be assessed against a substantial sample of data. But then the implications of and interactions between theories will be complex, and are best explored within an integrated NLP system.

When a theory is realised in an NLP system, new opportunities arise for testing it, both as a unit and in its interactions with other units. There are many reasons why an NLP system might perform well or badly, only some of which will shed light on any of the linguistic ideas it embodies. Nonetheless a system implementing a theory is often a source of insight into the strengths and weaknesses of the theory. Given the dearth of direct ways of verifying or falsifying linguistic theories, the evidence from computer models such as NLP systems is of great value.

1.4 A tour of the thesis

1.4.1 Word sense disambiguation: Chapters 2 and 5

Wilks's problem provides a point of entry to the study of polysemy. It presents a well-defined goal: to enable NLP systems to deal appropriately with natural language inputs, where many of the words in the input have more than one sense listed in the system's lexicon. Chapter 2 chronicles the history of assaults on the problem.

These have generally assumed the domain of word sense distinctions to be relatively homogeneous. Wherever a disambiguation strategy has been developed, it has been assumed it will be relevant across the board. The hypothesis examined in Chapter 5 is that there are quite different varieties of word sense distinctions, so different kinds of disambiguation strategies will apply in different areas. While the study was limited in its success at developing a broad-brush taxonomy of distinctions, it did serve to clarify a number of issues which shape other parts of the thesis. The close reading of the dictionary displayed both strengths and shortcomings. The sheer heterogeneity of the kinds of distinctions the lexicographer makes was striking. Information about the rule-following nature of usage-types, and where one was predictable from another, was discovered in the structure of the entry. But any expectation that the dictionary provides a key to a set of discrete 'word-sense' entities in the way that it does to a set of distinct words was soon

[5] We make no claims about how, or whether, linguistic and non-linguistic aspects can be distinguished.

dispelled. Different words are different because they have different spellings and sounds. There is no comparable fact of the matter for determining what makes a word sense different.

1.4.2 Breaking the Bank Model: Chapter 6

> "Read and read until you find something that everyone seems to accept that just can't possibly be true ..." (Anthony Robins, in conversation)

In addressing Wilks's problem, all researchers have, at least until the late 1980s, implicitly adopted the 'Bank Model', which we characterise as follows:

> A word like *bank* presents a very clear case of a word with more than one meaning. It can mean the side of a river, or an institution which looks after your money for you. For any usage of the word as a noun, either a money bank or a river bank is being referred to, and the word always refers to one or the other, not both. When English speakers encounter the word in a discourse, they know instantly and effortlessly which meaning of the word applies. This knowledge is an important part of human competence in a language, and an NLP system, likewise, needs to be able to choose.
>
> The word *bank* has been used because, in it, the issues are clear to see, making it a good pedagogical example. It makes clear what sorts of issues we have to deal with in lexical disambiguation for NLP.

The model suggests that the kind of distinction between senses found in the *bank* case is the pervasive kind found on every page of the dictionary. This is an empirical claim, which the experiment described in Chapter 6 investigates, using corpus examples of words in context alongside a dictionary. The results indicate that most words are sometimes used in ways which do not fit the Bank Model. Sometimes it was impossible to say which of a set of senses applied. Sometimes different senses contributed different aspects of meaning. The varieties of divergence from the Model are described.

1.4.3 "What is polysemy?" revisited: Chapters 3, 4 and 7

If the Bank Model is not a good picture of polysemy, what is? While philosophers, literary critics and linguists have expressed interest in the matter, it has never been central to their concerns. To lexicographers, polysemy is a pressing daily problem, but lexicographers write dictionaries rather than writing about writing dictionaries so the literature is limited. The most salient contributions are described in Chapter 3. The account adopted in the thesis and sketched in the 'What is polysemy' section above is developed and argued for in Chapters 4 and 7.

1.4.4 Formal lexicography: Chapters 8 and 9

Where polysemy is regular, and can be described in terms of an alternation relating to a number of words, it is redundant to express the facts about the alternate form at each lexical entry where it applies. There is a generalisation: this only need be stated once, provided that there are explicit mechanisms for stating that words 'inherit' it. In these chapters, some facts about the polysemous behaviour of classes of words are introduced and then formalised, using the lexical representation language, DATR. In Chapter 8, the regular polysemy linking senses of words for fruit, trees, and wood is given a concise formal treatment. Arguments for using a single taxonomy to organise both encyclopedic and lexical knowledge are presented, as are ways in which the formal analysis softens the distinctions between polysemy, homonymy and metonymy.

In Chapter 9, the 'diathesis alternations' relating forms of verbs with differing subcategorisation frames receive a similar treatment. An informal account of lexical organisation for some

classes of verbs is taken from Levin & Rappoport Hovav (1991) and formalised. Here, the different usage-types have different syntax. The analysis shows how the appropriate syntax-semantics mappings are inherited for each usage-type, and gives a concise, default-based account of the relations between syntactic complements and syntactic arguments. A further consideration was that the formalism in which an NLP system is written will impose constraints on the form the lexical entries should take. The lexicon entries described by the DATR theory inherited a DAG-like structure, making them directly usable by any NLP system working within an HPSG-like unification-based formalism (Shieber, 1986; Pollard & Sag, 1987).

1.5 Claims

The thesis makes three principal claims, one empirical, one theoretical, and one formal and computational. The first is that the Bank Model is fatally flawed. The second is that polysemy is a concept at a crossroads, which must be understood in terms of its relation to homonymy, alternations, collocations and analogy. The third is that many of the phenomena falling under the name of polysemy can be given a concise formal description in a manner that elucidates the relationship to alternations, metonymy and homonymy, and which is well-suited to computational applications. Two small parts of a lexicon that meets these goals are presented.

Chapter 2

Literature Review I: Word Sense Disambiguation

2.1 Introduction

A major theme of the thesis is how polysemy has been addressed within natural language processing. It has long been evident that there is a problem. Words often have several meanings. If an NLP system is to operate at all on the meanings of words, it will have to access the right meaning where there is a choice of several. This chapter will look at the history of the problem, reviewing the literature and sketching how findings from other disciplines shed light on the arena.

In the beginning, there was the direct approach. The problem was that words were ambiguous, so the solution was to devise procedures for disambiguating them. The concern was for showing what it was possible to do with computers, in the Artificial Intelligence mould. Section 2.2 covers this work.

But a dominant feature of the lexicon is its size. There is a 'lexical acquisition bottleneck'. Writing the procedures for disambiguating words was very time-consuming. For many, a more appealing technique was to extract information from an existing source: the dictionary. Machine-readable dictionary research is chronicled in section 2.3.

2.2 The Artificial Intelligence tradition

The work in this tradition has included the Bank Model amongst its theoretical presuppositions. For all this work,

1. The author gives no justification of how he chose the sample of words to be considered. The words are selected according to the researcher's ideas of what words were interestingly ambiguous.

2. The senses to be chosen between were arrived at by the investigator. All the authors make reference to the large numbers of senses to be found in dictionaries as a major source of motivation, yet no use is made of published dictionaries thereafter.

3. The possibility that a usage might fit more than one sense is not mentioned.

4. Only a very small number of words has been studied.

Chapter 6 considers these reservations in detail; for the time being, let us put them to one side. The work that has been done has explored the task as it relates to at least one subset of cases, and effective NLP systems will certainly need to exploit some of the same techniques. I shall discuss the work of Wilks, Boguraev, Small, Hirst and Cottrell.

2.2.1 Wilks: Preference Semantics

Wilks's model (Wilks, 1975) is one in which the all-or-nothing semantic features of Katz & Fodor (1963) have been softened to become preferences. Rather than specifying what is and is not part of the language, as was Katz and Fodor's goal, the features associated with a word now lead us towards a best attempt at interpreting the input. The basic selection-restriction mechanism is this. The item in the lexicon is marked, if it is a noun, according to its semantic features. These form a set which almost always includes HUMAN, ANIMATE, ABSTRACT and a few others but can then have any number of other elements. If it is a verb, it is marked with the likely semantic features of its subject and other arguments; if a preposition or adjective, with the likely features of the noun or noun phrase that will follow it or that it will modify. Where a word has more than one sense, a different lexical entry will be provided for each sense, and assuming that there are different semantic features (on the nouns) or selection restrictions (on verbs, prepositions and adjectives) for the different senses, then sometimes only one set of choices of word senses will permit the sentence to be interpreted without selection restrictions being infringed. This will then allow us to select that set.

For Wilks, input for which selection restrictions are infringed will no longer be outlawed as ill-formed. He considers that any input a system receives will have made sense to its author, so it is incumbent on the system to produce an interpretation. If there is both an interpretation of the input in which selection restrictions are not infringed, and one in which they are, the one without infringements will be preferred, but if there are only interpretations with infringements, the one with least will be put forward as the system's interpretation. Thus the system could make sense of a limited range of metaphorical usages, such as the much-cited

> My car drinks gasoline.

(Although *drinks* expects an animate subject, it will accept a machine).

Wilks uses a set of sixty semantic primitives, and his lexical entries are templates specifying, for a noun, which primitives generally apply to it, and for a verb, which primitives are to be expected to apply to its various arguments, as well as the kind of event or state it denotes. The system operates with dictionary rather than encyclopaedic knowledge, though it is not argued that there is a principled distinction between the two. For Wilks, preferences play a central role throughout the various stages of natural language processing. He uses semantic parsing, guided by preferences, and those same preferences are the mechanism by which words are disambiguated and anaphora resolved. They also provide the language in which the eventual interpretation of the input is presented.

Both his language and his system are severely limited by using just one kind of mechanism, of such limited expressiveness. The approach eschews clues from syntax or from associations of nearby words and phrases which do not stand in head-modifier relations. It does not represent the referents of the text so where any inference about the referents is required for disambiguation, as in disambiguating between physical and debating points in

> He found a stick, got out his knife, and made a point so it would stick in the ground.

there is no possibility of the system reasoning to the 'physical' reading. Where an ambiguity can only be resolved by observing the relations between the discourse referents and these are not reflected in relations between the individual words (other than pronouns) used to identify those

referents, a preference semantics system is not equipped to make the leap. Wilks's contention was that such cases were relatively infrequent, and that the computational costs of resolving them were very great, thus making them a low priority for NLP at the time Wilks was writing.

While selection restrictions and preferences play a role in all the other systems discussed and are of great use to practical NLP systems, later systems have not used preferences to drive syntax and semantic interpretation, and have used semantic features and preferences as one tool among many for lexical disambiguation.

2.2.2 Boguraev: a not-quite-so-mini working program

Boguraev's thesis work (Boguraev, 1979) borrows Wilks's formalism for semantic representation, but purely as a pragmatic move, so that he has a formalism to work with. His work explores how semantic and syntactic factors can be brought together for purposes of disambiguation. Thus he uses syntax for parsing. Whereas Wilks dealt almost exclusively with ambiguities between two different nominal senses of words, Boguraev disambiguates both between readings where the word has different parts of speech, and between alternative senses having the same part of speech.

Boguraev's program generates alternative sentences, with the ambiguous word or words substituted by others that retain the sentence meaning. It is an existence proof that it is possible to disambiguate, between the senses listed, using the limited syntactic and semantic information available in his lexicon. It operated with a lexicon of 400 words, which though not a vast number, is large by the standards of experimental AI systems. It is large enough so that it is unlikely all the words in it are special cases, or that they have all had particularly ingenious entries written for them, so it does indicate it is likely that the approach will be quite straightforward to extend.

2.2.3 Small and Word Expert Parsing

For Boguraev the distinction between two different word senses of the same syntactic category was something that took place after parsing and, as far as possible, after the subject matter had been determined. For Small (1980), the lexicon is the location of all information relating to the syntactic roles, referents and domain of discourse of words in an input sentence so it is through the process of disambiguating that we determine the structure and meaning of the sentence. The 'word experts', as the packages of information in the lexicon connected with particular words are called, operate directly on the input text. There is no intermediate syntactic structure.

Small also argues that each word needs its own expert, and is sufficiently idiosyncratic so that word experts are, essentially, to be written on a case-by-case basis. How words differ is for Small far more impressive than how members of some subsets behave similarly. A word meaning is then both a procedure and (assuming the word has more than one sense) a discrimination net. The procedure is one that, given a context of likeminded lexical entries for neighbouring words, will access the information needed from its neighbours and use it to determine the route to be taken through the net. His principles are:

- Words are active processes for language comprehension.
- Uniform rule interpretation is incompatible with language comprehension.
- Word processes are discrimination networks. (pp 43–44)

Rather than there being general control mechanisms which apply a grammar to the input, or a set of semantic composition rules to a syntactic structure, the control information is specific to the words in the text. Word experts are thus the repository for a very large quantity of information that in other systems might be distributed between a parser, grammar, lexicon, and other modules. Moreover he is ideologically committed to each word expert not merely being a near copy of another. Thus:

> The construction of word experts requires patience, dedication, and finesse, and inherently involves far more intricate labor than ought to be expected of any person. (p 200)

Small's approach to parsing pays great regard to the idiosyncratic behaviour of words. The ubiquity of exceptions and oddities, and the extent to which much research has overlooked it, is a concern shared by the current study. Yet the benefits of what generalisations there are need not be cast aside in order to express and pay heed to exceptions. A default inheritance formalism such as DATR (Chapter 8) provides a mechanism where regular, subregular and irregular information can be stored in a single, consistent manner. Also Small insists that control information should be stored alongside declarative knowledge about the word. Systems where control information is to be found in many different places are notoriously difficult to comprehend or to modify, and the same functionality can be obtained without those drawbacks in a system which keeps the lexical information separate from the procedures for manipulating it.

His model of discrimination nets does open the following possibility: it might not be possible to reach a decision, for a particular word usage, as to which branch of the discrimination net to follow. For some such nodes it might be appropriate to say the usage is simply unspecified, as between those two senses. Beyond noting the variability of how well specified a word usage might be:

> Each word expert continues to probe the context even after creating an appropriate meaning representation for the word it represents, in order to refine the representation to reflect as much of the context as possible. (p 77)

he does not explore the question further.

2.2.4 Hirst's Polaroid Words

Hirst (1987) presents a Natural Language Processing system which incorporates a grammar, parser, lexicon, semantic interpreter, frame-based knowledge representation system (in which the final interpretations of inputs, as well as general world knowledge, will be expressed: the formalism is called FRAIL) and specific lexical and structural disambiguation processes. The whole is a working system: where the input words all fall within its small vocabulary and the topic is within its domain, its knowledge base will be amended or updated according to the meaning of an input sentence. An advantage of this many-sided system is that there are many different possible sources of information which might be relevant to word sense selection, and here they can be seen playing in consort. In contrast to Small's Word Expert Parsing, the approach is modular, and the interplay of different components is correspondingly easier to trace.

Hirst is one of several investigators who have been concerned with the cognitive science goal of modelling human lexical disambiguation processes They have wanted their systems to display the same sorts of behaviour that humans have been found to display in a number of psycholinguistic experiments. Hirst argues that, often, 'finding out how people do something and trying to copy them is a good way to get a program to do the same thing' (p 20), and states that the research will be of wider interest, if it aims for psychological reality.

Most of the relevant psycholinguistic work has used semantic priming techniques. It is well-established that, if I have just heard the word *doctor* (the prime), and then a sequence of letters is flashed up on a screen and I am asked to identify whether it is a word or not, I shall respond faster if it is a word and it is *nurse* than if it is a word but unrelated to *doctor*. (This is the 'lexical decision' task in a mixed, visual and auditory procedure. It is one of a variety of versions of semantic priming experiments. The basic effect is robust across a number of experimental strategies.) If a prime such as *bank* is given, it turns out that both *river* and *money* are primed

for. In a context which serves to make one and only one of these an appropriate reading, after something between 50 and 200 ms a choice is made and after that only one of *river* and *money* is primed for.

There has been a debate as to whether all meanings of a word are accessed, and then a choice is made (Post Decision), or whether the context determines that only the appropriate reading is ever accessed (Prior Decision). The kind of evidence alluded to above regarding *bank* is taken by many to have decided this issue in favour of Post Decision. There have been experiments aimed at distinguishing between high and low frequency readings of the prime, and here again it seems that both readings are primed for.

In the spirit of semantic priming, Hirst employs a strategy of 'marker passing'. This is a discrete model of spreading activation, which passes flags round a knowledge base to establish which senses of which words are most closely associated with which senses of which other words in the input. If there are striking results, it can be concluded that the correct word senses have been found. It involves marking the node in the FRAIL knowledge base for each sense of the word and then, in a time-stepped iterative process, marking all the nodes for word-senses that appear in lexical entries containing an already-marked node. All the senses of all the words in the sentence are marked simultaneously, and the process is set running. There is a set of rules regarding how many time steps the process runs through (if this is set too high, the whole knowledge base will be marked), what action to take when there is a 'direct hit' (with a sense for one of the words being marked in the first time step by a sense of another), what action to take for an indirect hit (when a node is marked from two different sources), and various other complications. (Small refers to 'conceptual proximity' (p 87), which is a metric for a comparable strategy for Word Expert Parsing, but it appears to be neither fully worked out nor implemented.)

Hirst calls his more specialised lexical disambiguation processes 'polaroid words'. Polaroid, because like the photographs, they develop over time and during that time they can be inspected and will give an honest, if incomplete, picture of what the finished product will be. The time in which the picture is emerging can include the time in which parsing, semantic interpretation and structural disambiguation are taking place. Unlike a polaroid photograph, the picture emerges only through interaction, and as these other processes do their work, so there will be more information available to enable the picture to develop.

There are polaroid word processes described for nouns, prepositions and verbs. There is information about a word in two places. In the packet of knowledge brought in to play by the polaroid word process the alternative senses are introduced, and in the knowledge base there is the world knowledge about what each sense of the word denotes. The mechanisms used by polaroid words are selection-restriction ones, but since there is a knowledge base already containing hierarchies of types of object, there is no need to introduce a set of semantic features specially for use with selection restrictions. The taxonomy used in the knowledge base, and which, in principle, will be used for all manner of other inferences, can be used. Moreover there is the potential for resolving the ambiguity in examples such as the *point* case described above because, in FRAIL, there are representations of objects which one can learn more about. Whereas for Wilks the reasoning could only be about words and their associations, for Hirst there is the potential for reasoning about the referents of the discourse. ABSITY, the semantic interpreter, uses model-theoretic semantics to build a representation of the entities and relations referred to, and sometimes this will provide the information the polaroid word is waiting for.

Hirst's achievement is impressive particularly because, like Winograd's SHRDLU in the early seventies, it brings together techniques from a number of different areas within AI Knowledge Representation and NLP and shows it is possible to integrate them. He does this while retaining an interest in the psycholinguistic evidence: his system is always compatible with the little we know about how people process language. Lexical disambiguation is a task that requires cues from local syntax, from selection restrictions, from word associations, and from the overall structure

and meaning of the sentence, *viz.*, what sorts of things are being referred to and into what sorts of roles do they fall. Hirst provides an architecture in which all these cues co-operate, and, as a part of the process of arriving at an interpretation of the input in a general-purpose knowledge representation language, words are disambiguated along the way.

2.2.5 Cottrell and connectionism

Connectionism, or neural networks, provides a style of computing well suited to many processes which require various constraints to be satisfied simultaneously, and also appear to have some design features in common with the brain. Word sense disambiguation is a process requiring evidence from various sources to be brought together and evaluated, so, *prima facie*, is a suitable domain for tackling with a neural network. And there again, there is the added allure that a system performing the disambiguation in this way could be argued to be doing it 'in the same way the brain does', though whether the features shared between the brain and connectionist systems are very significant for understanding higher-level mental processes is an open question.

Here, Cottrell's *A Connectionist Approach to Word Sense Disambiguation* (1989) is considered. A basic understanding of the nature of connectionist models is assumed.

Cottrell's goal is in the domain of cognitive science rather than NLP. He wants to model the processes behind the semantic priming evidence from psycholinguistics, and to do this in a way which is compatible with several language disorders to be found, on rare occasions, amongst human populations. His goal is to build a system which is able to disambiguate, but he is concerned primarily with the mechanism, not the performance. He argues:

> it is important to understand how people resolve the ambiguity problem, since whatever their approach, it appears to work rather well. (p 1)

Since word sense disambiguation is one part of the process of sentence understanding, he also needs to model the other parts of the interpretation process which send messages to and fro. Only in a network where alternative syntactic and case interpretations were being weighed alongside alternative word sense selections could the best interpretation be expected to emerge. A syntactic module was essential if the system was to choose the correct part of speech for categorially ambiguous words. A case module was essential if selection restrictions or preferences were to be brought into play, since they involve the assignation of a noun word sense to a case-slot of a verb sense. There are frequently alternatives for how nominals are assigned to cases. So the overview of his model is as shown in Fig. 2.1.

For the likelihood of these options to be represented in the model, case assignations, alongside word sense ones, must be seen as part of the puzzle to be solved. Word associations of the types Hirst modelled with marker passing are modelled here through positive feedback at the word sense level.

The lexical module is the input mechanism. As each word in the sentence is encountered, the unit representing it is fired. The words activate their word senses, and they activate the syntactic roles and cases they expect to fill or to have filled, as well as 'all related concepts', marker-passing-style. An interpretation of a sentence is reached where, for each word, just one word sense unit, just one case and case-binding unit, and just one syntactic-role and syntactic-role-binding unit is firing.

There are units for, amongst other things: each lexical item; each sense of each of these; each semantic case, in relation to the verb of a single clause sentence; each argument position of a verb; and then, the binding units. Different versions of Cottrell's system use these to different extents. The problem is, if several different NPs might be the subject of a verb, and again, each of those may take the AGENT role, it is not enough to say that the verb has a subject and an AGENT and for each NP to be disambiguated. A link must be made between just one of the NPs and

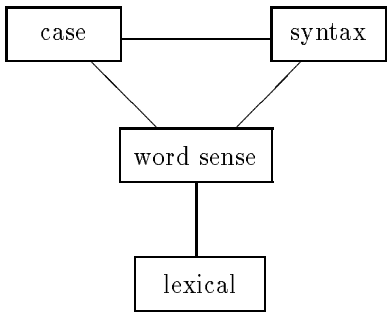


Figure 2.1: System architecture, Cottrell (1989), p 13.

the subject role, and between just one of the NPs and the AGENT case. In Cottrell's version of connectionism, any aspect of the interpretation of a sentence is to be represented by a unit firing. To allow for all the possibilities, the straightforward option is, therefore, to have a 'binding' unit, for every NP-and-syntactic-role pair, and for every NP-and-case-role pair. The unit would then fire only when that NP filled that role. The problem is, the number of these units will increase exponentially. Cottrell considers and suggests some partial solutions.

The model is clearly a complex and impressive system, and it offers some intriguing insights into, for example, different strategies for making prepositional phrase attachments. However Cottrell, like Small, is devising a system where one style of representation and processing is applied to syntactic, semantic, lexical, selection restriction and word association cues alike. His arguments for doing this are that, firstly, in our brains, all these things are going on in a similar medium, *viz.*, impulses along neurons, and secondly, the psycholinguistic evidence that the processing of all these cues is in parallel in humans. While these things are both true, they do not indicate that our ability to model or reproduce the behaviour will best be served by a connectionist system. An alternative strategy is 'divide and rule', whereby the syntactic, semantic and other aspects of the process are abstracted away, in order that they can receive specialist attention. Once these processes are better understood on their own, our chances of identifying and modelling the contribution they make to the overall problem will be much improved. Despite his claims to an interdisciplinary approach, Cottrell makes almost no reference to the linguistic literature, be it on syntax, formal semantics, verb valency or lexicology, although all of these areas have received a great deal of specialist attention which is relevant to the task. In due course, when what is specific and different about all these processes is sufficiently well understood, when, in Marr's terms (Marr, 1982), we have resolved what computations are being performed in the course of lexical disambiguation and according to what algorithms, then there will be a level of understanding which makes the question of how the algorithms are implemented in neurones a focussed and well-specified one. But until that day, a 'divide and rule' approach is preferable. It will make computer models easier to understand and modify. New findings from specialists working in particular corners of linguistics or psychology will often need incorporating into any model of disambiguation, and this will be facilitated if the design is modular. Bringing information from the various different sources together is an important matter to study, but its value is dependent on the correctness of the more specialised analyses.

2.2.6 McRoy: A state of the art system

A characteristic of Hirst's approach was the way in which it acknowledged that there would never be a 'neat' theory of word sense discrimination. The process required information from a variety

of knowledge sources, and a system would have to bring different cues together and weigh them against each other. McRoy (1992) describes a system which takes this theme and brings it up to date, adding knowledge sources which were little understood, or where the information was simply not available, when Hirst's work was done in the early eighties.

A major theme in NLP in recent years has been 'scaling up'. Toy systems have been built and have indicated what might be done: the task now is to see if the same things can be done on a large scale. Whereas all the work described above used lexicon of tens or, at most, a few hundred words, McRoy worked with a lexicon of 13,000 senses. The 'life-size' nature of her study has also meant she confronts Wilks's problem head-on: she cannot first parse a sentence to give all the possible parses and then set about the disambiguation task. There will simply be too many parses to store and assess. Disambiguation must take place as early as possible, alongside the parsing process.

The burgeoning discipline of corpus linguistics has made available collocational information (section 3.3), and McRoy uses a lexicon of collocations automatically extracted from a corpus of the same genre as the texts she aims to disambiguate. If two or more words in a collocation occur together in an input text, that provides strong evidence that the words are being used as they are in the collocation. The system would need customising if it was to be used in a different domain and part of the task would be building a new collocation, but as this is substantially an automatic process it would not be a major cost. A further theme of corpus linguistics has been the 'ecology of language' and the extent to which words and word senses occur in specific domains. McRoy uses both a core lexicon and domain-specific lexicons. This limits the disambiguation task. *Engage* has an 'attack' sense only in the military domain so this sense will only be considered when military matters are under consideration and the 'military' lexicon is active. The core lexicon, containing the general-purpose word senses of the language, is always active. However:

> by design, [it] includes only coarse distinctions between word senses. This means that, for a task such as generating databases from text, task-specific processing or inference must augment the core lexical knowledge, but problems of considering many nuances of meaning or low-frequency senses are avoided. (p 7)

It is not clear what justification there is for the apparent assumption that subtle distinctions between senses are always domain-specific and do not belong in a general-purpose lexicon.

In both core and domain-specific lexicons, McRoy distinguishes high-frequency 'preferred' senses and low-frequency 'deprecated' ones. Deprecated ones will only be selected when there is some specific reason, such as a failure of selection restrictions for all preferred senses.

She also uses a concept hierarchy, with each sense having a 'parent' stated in the lexicon. This serves several functions. One is to improve on Hirst's 'marker-passing', which required a number of *ad hoc* rules. To test whether two senses are semantically related, the system checks to find the lowest point in the taxonomy at which they have a common ancestor: the lower it is, the closer their relationship. It also greatly reduces the labour involved in lexicon-building. As far as possible McRoy uses small 'conceptual clusters', as she calls them, such as `c-published-document` (where the '`c`' specifies a conceptual cluster as distinct from a word sense of English). The basic information needed about each sense to drive the semantic aspect of the disambiguation process is, as in the other systems described, selection restrictions or 'role-related expectations'. All the senses in the same cluster tend to share this information, so these shared specifications can be stated somewhere higher than the base level of the hierarchy and inherited by the senses below. So when a sense is added to the lexicon, it will not in general be necessary to describe the specifications it shares as they can be inherited. The selection restrictions and preferences themselves can also be stated in terms of nodes as 'high' in the hierarchy as appropriate. Thus where a verb has a selection restriction on its direct object, the selection restriction can be stated at a general node in the verb hierarchy, where it applies to a class of verbs, as a requirement that

the nominal concept filling the direct object role is a concept to be found at a general node in the noun hierarchy where it dominates a class of nouns.

The system first analyses words morphologically to identify roots and affixes, then locates their entries in the various lexicons, then does part-of-speech tagging (section 3.3.3), then parsing and semantic interpretation, at each stage rejecting as many senses as possible. A complex scoring system is described for weighing the evidence for and against different senses, and the system returns the highest-scoring sense.

The system was tested using arbitrary samples of text from the *Wall Street Journal*. Results were "quite encouraging" (p 26) with the system often finding the right word senses even where a sentence could not be fully parsed. However, McRoy says:

> We have been unable to get a meaningful quantitative assessment of the accuracy of the system's sense tagging. We made an unsuccessful attempt at evaluating the accuracy of sense-tagging over a corpus. First, we discovered that a human "expert" had great difficulty identifying each sense, and that this task was far more tedious than manual part-of-speech tagging or bracketing. Second, we questioned what we would learn from the evaluation of these partial results ... (p 26)

The outcome is exactly as expected, given the account of word senses presented in this thesis.

2.3 Machine Readable Dictionary (MRD) research

Barring McRoy's very recent work, all this research has used lexicons with hand-crafted entries, one reason being that lexicons in which the required information can be accessed automatically were not available. Disambiguating word senses is one of many reasons why NLP needs large lexicons. Knowledge about words is needed across NLP, and if an application is to process free text input it will generally need knowledge of a lot of words. The benefit of extracting information from dictionaries where possible has long been recognised. As dictionaries have become available in machine-readable versions, so tools and techniques have been developed for extracting information from MRDs and for converting MRDs into lexical databases, also called machine-tractable dictionaries.

Byrd (1989) observes that "the same lexical information that serves to identify word senses ... is also the information that we want to store in CompLex [IBM's lexical knowledge base]" (p 78). This is no accident. An essential criterion for a knowledge base is that some inferential processes are defined over it, so that the knowledge it contains is not only that which is directly stated, but also that which can be inferred. The one inference mechanism used in all knowledge representation schemes is based on the IS-A link. If an *A* IS-A *B*, then any properties known to apply to all *B*s can be inferred to apply to all *A*s. The interest in genus terms is based on the observation that the relation from a word to its genus term bears a close resemblance to the IS-A link. *Pilot*, in LDOCE, has the genus term *person*, and this would correspond to an IS-A link in a knowledge-base developed form the dictionary from 'pilot' to 'person'. Ideally, then, finding genus terms from the dictionary determines the IS-A hierarchy, or taxonomy, for the denotations of words in the dictionary. But where a word is ambiguous, the plan is foiled. In Webster's Seventh, senses of both *pigsty* and *quill* have genus term *pen* which has one sense with *enclosure* as genus and another with *implement*. If inheritance is based on words, then "there is no way to block the inference that a *pigsty* is an *implement* or that a *quill* is an *enclosure*" (Klavans, Chodorow, & Wacholder, 1990, p 114). So an MRD potentially provides an inheritance hierarchy for the denotations of many of the words of the language, but the potential can only be realised if genus terms are disambiguated. Amsler (1980) employed armies of hourly-paid workers to disambiguate the words in definition-texts in order to be able to proceed with his classification of the word senses

and to arrive at a dictionary-based inheritance hierarchy. Disambiguation strategies are needed for the production of a lexical knowledge base from an MRD, and the information gatherered in that process can then also be made the content of lexical knowledge base entries, particularly since tasks the knowledge base might be used for, such as information retrieval, often involve sense disambiguation.

Klavans (1988) surveys the 1988 state of the art (from an IBM perspective). She emphasises the level of demand for theory-neutral computational lexicons and discusses how IBM's products, first UDICT and now CompLex, have fared. The demand is from the whole range of types of NLP application, from translation to query-answering and style-checking. She identifies two major problems, the polysemy problem and the mapping problem. The mapping problem is the problem of mapping the senses that one dictionary gives for a word onto the senses that another dictionary gives. Different dictionaries divide up the range of uses a word has in different ways. The problem is ubiquitous. The polysemy problem comes about when semantic features and selection restrictions are to be attached to dictionary word senses. The UDICT formalism only allowed one list of features for each word in a specific part of speech. Which features are to get onto that list: any feature that any usage of the word might reasonably have, or only those features that all usages of the word might be expected to have? Behind both major problems lurks the same ugly question: what are we to make of dictionary word senses? From Amsler's early work until now, it has sat stubbornly by.

In MRD research the problems and projects have tended to be practical, to do identifying and exploiting the structure of the published dictionary. It has not been a domain of competing theories but rather one where different research groups have developed various tools and techniques, or have achieved wide-coverage analyses, and further research has simply been able to use the results of what has been done before. What follows is a brief history of the projects completed and tools developed.

2.3.1 Amsler: *The Structure of the Merriam Webster Pocket Dictionary*

The first major work in the field was Amsler's thesis (Amsler, 1980). It was his goal to discover the conceptual hierarchy or hierarchies implicit in this particular MRD. His motivation was to provide a knowledge representation scheme, upon which inference could be performed, for a substantial part of the vocabulary of English. The method was computer-assisted, rather than automated or semi-automated. His extended example comprised the verb *move* and those verbs defined in terms of it. He worked from the assumption that all definitions complied with both the 'replaceability' criterion (that the definition should be able to replace the word without change of meaning) and a genus-and-differentiae structure. Definitions which did not fit these patterns were modified so that they did, in a pre-processing phase. Then, all the verbs with *move* as the genus terms in their definitions were collected. A search was made for all the different senses of *move* in the language, in the dictionary under scrutiny and elsewhere. Then, for each sense of *move*, the definitions in which *move* was being used in that sense were identified. This was one of the labour-intensive operations for which large numbers of paid disambiguators were brought in. He then looked through the definitions to determine what range of case relations were specified in the differentiae of the definitions. Then, with the range of cases known, the differentiae could be rewritten as sets of attribute-value pairs.

The work is important in that it showed that taxonomies could, in a systematic if not at that time automated process, be generated from dictionaries. It encounters and discusses the problems of loops in the dictionary; of definitions varying from the genus-and-differentiae structure; of hard-to-identify genus terms; and of the relation between case-relations and semantic primitives, thus setting the agenda for much subsequent research.

2.3.2 MRDs for sense disambiguation

A direct MRD approach to word sense disambiguation was adopted by Lesk (1986). His premise was that each word sense was likely to appear in proximity to a limited number of words, and a well-written dictionary would convey this information by using some of those words in the sense definition. The algorithm took all the words in an input sentence containing an ambiguous target word, and calculated the overlap between those words and the words used in the definition for each sense of the target word. The process claims 50-70% accuracy. The technique has been developed and refined, incorporating LDOCE codes for subject-areas in one study (Guthrie, Guthrie, Wilks, & Aidinejad, 1991) and using a very large neural network to encode strengths of relationships between different words in another (Veronis & Ide, 1990). This work is constrained to work with the word senses a particular dictionary offers, and treats all distinctions (with the possible exception of part-of-speech ones) as equivalent. Nonetheless it is of great value, both for practical functions such as information retrieval and for the purposes of bootstrapping sense-disambiguated lexical resources into existence.

2.3.3 Parsing the dictionary

Entries for published dictionaries have more structure and consistency to them than many other varieties of text, yet they fall short of what is required for fully automatic processing. There are different fields of information within a dictionary entry, for example syntactic information, the definition text, alternative senses, and collocations (see Chapter 5). The raw material of MRD research is generally the typesetting tape used for printing the paper dictionary, and this does not directly support automatic access to the different fields of an entry. For NLP it is essential that fields can be readily accessed, and the process of making the information available in a relational database has been described in several papers (Ahlswede & Evens, 1988; Nakamura & Nagao, 1988; Calzolari & Picchi, 1988).

While this work meets the criterion of making the content of the fields available, Neff & Boguraev (1989) point out that there is also information embedded in the often complex hierarchical structure of the entry, and this is generally lost with the conversion to relational database form. A dictionary entry, like a sentence, has a tree structure which can be recovered from the string format. They have developed a parser and framework for writing grammars for entries for particular dictionaries. As with natural language grammars, the grammar provides rules according to which an underlying structured representation can be recovered from a linear representation. For a grammar for a dictionary typesetting tape, the terminal symbols are not (for the most part) words but are opening and closing brackets, typesetting codes for change-of-font and indented-new-line, parentheses, slashes, bars, and text strings. When the grammar is written, entries can be parsed to produce a structured representation of their contents. They have also written LQL, Lexical Query Language, for accessing information and they describe the way in which the hierarchical structure, true to the dictionary, makes it easy to make complex, linguistically interesting queries. The approach moves towards an 'interlingua for dictionaries'. Entries for different dictionaries can be transformed into the same format, so opening up new vistas for exploring and exploiting the contrasts between the approaches of different dictionaries. The potential for using bilingual dictionaries alongside monolingual ones (Calzolari & Zampolli, 1988) is of particular interest, in relation to lexical structure in general, translation and word sense disambiguation, as different senses of a word sometimes have different translations.

2.3.4 From MRD to LKB

Alongside the task of making the fields in a dictionary entry accessible, is the one of making the content of the definition text available. Work at the Cambridge University Computer Labora-

tory, IBM's Thomas J. Watson Research Centre, Illinois Institute of Technology, the Computing Research Laboratory at New Mexico State University, the Istituto di Linguistica Computazionale del CNR, Pisa, amongst others, has been setting about automating various of the processes that Amsler performed manually. Byrd, Calzolari, Chodorow, Klavans, Neff, & Rizk (1987), and other articles in the same special edition of Computational Linguistics describe progress so far, as does Boguraev & Briscoe (1989), a volume specific to research on LDOCE. Wilks, Fass, Guo, McDonald, Plate, & Slator (1989), in that volume, explicitly address the question of the potential of MRDs as resources for building lexical knowledge bases, and both the introduction to the book and that paper review the history of the enterprise.

The techniques include parsers and pattern matchers for dictionary definitions. They are aimed at identifying genus terms and their inverses (hypernym relations, or a term's 'children' in the taxonomy), lexical relations (Evens, 1988), and more specific sets of words such as active and stative verbs (Byrd et al., 1987). Different articles concentrate on different parts of speech (e. g. adjectives in Ahlswede (1985)), on definitions which vary from the genus-and-differentiae format (Vossen, Meijs, & den Broeder, 1989; Guthrie, Slator, Wilks, & Bruce, 1990), or on the exploitation of thesauri, corpora (see below) and further dictionaries (Calzolari, 1989).

Crucial to much of this work is the fact that lexicographers tend to use 'defining formulae' in writing their definitions. The standard practice for nouns is for the first noun phrase of the definition to have the genus term as its head noun. Several studies (Vossen et al., 1989; Guthrie et al., 1990; Klavans et al., 1990) have looked at noun definitions which do not follow this pattern. The standard alternative is for the definition string to have an 'empty head' as the head of the first noun phrase, with *of* and another phrase following it. A way to search for the defining formulae following this pattern is to find all noun definitions with *of* near the beginning, then to find, for all those definitions, the word which preceded *of*, and to see which words occur in that list with high frequency. Using this technique on Webster's Seventh, Klavans, Chodorow and Wacholder arrived at a list of 29 words which occurred in this position in 100 or more definitions, and can be considered defining formulae. They investigated further the cases with *unit* and *group* as the empty head. On analysing the different patterns containing 'unit' they are able to isolate the patterns in which *unit* has its 'amount or measure' sense, and those where it has a 'subdivision of an organisation or structure' sense. The definition of *inch* is 'a unit of length . . . ' and this sort of analysis makes possible the automatic or semi-automatic extraction from an MRD of lists of words like *inch* which denote units of measure, of words like *length* denoting abstract quantities which are measured, and for this kind of information to be entered in a lexical knowledge base under *inch* and *length*.

2.3.5 Using the dictionary as it stands

A slightly different approach has been adopted by Jensen & Binot (1987) and Ravin (1990). These authors, like Lesk, have used published dictionaries directly, but have been addressing different kinds of cases. They have studied the ambiguities of prepositions which can mark any of a number of cases. Ravin specifies that this is part of the bootstrapping process. Dictionary definitions, as they stand in published dictionaries, are a valuable resource for automating the business of disambiguating other dictionary definitions, as a prelude to putting the disambiguated versions in a lexical knowledge base. Thus she sets herself a goal, similar to that of Jensen and Binot, of taking those verb definitions which fit the pattern 'to VERB with NP', and of disambiguating the *with* according to which of a set of five semantic relationships holds between the verb and the NP. Her example is:

angle: to fish with a hook.

The five senses were arrived at after consultation of various dictionaries, a descriptive grammar (Quirk, Greenbaum, Leech, & Svartvik, 1972) and a small corpus, and three of the five have further

subdivisions. The disambiguation process first parses the string to identify the verb and the noun which heads the NP. Each of these is then itself looked up in the dictionary (Webster's Seventh was used). The head noun, *hook* in the example, is found to have *implement* as a genus term for its nth sense, so it is marked +INSTRUMENT. This counts as evidence towards the conjunction of a USE-OF-INSTRUMENT sense of *with*, and a use of *hook* in its nth sense. The next stage is to look for evidence of each of the possible with-relations in turn. Additional evidence for the USE-OF-INSTRUMENT sense is found via the fact that the verb *catch* is found in the definitions of both the first intransitive sense of *fish* and the nth sense of *hook*.

Measuring the success rate of the program is not a straightforward matter. Sometimes it arrived at more than one interpretation, sometimes none at all. Also the experimenter's decisions about the 'correct' interpretation of some definitions will be debatable. But putting such details to one side, the program did correctly disambiguate 75 out of a test set of 132 examples, with just 22 having only incorrect interpretations. In the course of disambiguating the preposition, the noun and verb are sometimes also disambiguated. Also semantic features and selection restrictions are assigned to word senses. They could in principle be stored and re-used. The project is an interesting exploration of how the great unformalised store of information within dictionary definitions might be directly exploited. However it should be noted that Ravin's and Jensen and Binot's work has sought to disambiguate just one word, *with*. Such detailed word-by-word analysis is certainly required. To complement it, studies that look at the vocabulary of the language at large are needed. Unless such studies succeed in finding regularities across sections of the vocabulary, the prospect of producing disambiguation strategies for the whole lexicon where many words might need even a fraction of the individual attention lavished on *with*, is daunting indeed.

Chapter 3

Literature Review II: Linguistics and Lexicography

3.1 Introduction

The last chapter surveys responses to Wilks's problem. But Wilks's problem was only one of several answers to 'Why study polysemy?' This chapter surveys work concerned with other answers.

As argued in the introduction, dictionaries are a treasure-house of information on polysemy. Perhaps their authors have direct answers to the questions of how, and with what rationale, lexicographers have deemed words polysemous. Section 3.2 considers the lexicographical literature.

The credibility of dictionaries rests on their authors' consideration of the language as it is used. The language as it is used can now be studied as never before, through samples of it in machine-readable corpora. Corpus research is the topic of section 3.3.

As the quantity and sophistication of the lexical information available increases, so the question of how to structure it takes a more prominent role. The remaining sections look to theoretical, cognitive and computational linguistics and Artificial Intelligence for relevant work and insights on lexical and conceptual structure, and bring the history up to date, to the point of departure for the argument presented in Chapter 1 that "the nature of lexical knowledge should determine the characteristics of the model of the computational lexicon" (Boguraev & Levin, 1990, p 67)

3.2 The lexicographers

Lexicographers have generally written remarkably little about the nature of multiple word senses. On the one hand, this is surprising, given that it is an issue a lexicographer must struggle with for most of the dictionary entries they ever write. On the other, lexicographers are by and large concerned with writing dictionaries, not theorising about writing dictionaries. What has been written about the task of lexicography is mostly concerned with more practical matters like how to select and alphabeticise the headword list or collocations, and how to represent syntactic, phonological, or usage information. While stating that the specification of word meaning is the central task for the lexicographer (Zgusta, 1971, p 23) and the division of a word's meanings into senses is a central part of that, Zgusta's authoritative handbook gives little guidance beyond admonishments to avoid making too many, or too few, distinctions (pp 66–67).

One reason is this. Theoretical studies of dictionaries and lexicography inevitably fall within the domain of linguistics. But linguistics has proved itself of very limited use to the lexicographer.

Zgusta warns him or her "not to be too impressed by the basic uncertainty concerning the nature of lexical meaning" (p 24) and, until recently, the methodology of researchers in linguistics has almost always been too contentious and the conclusions insufficiently specific to be of practical use. More recently, particularly since Hornby's Oxford Advanced Learners Dictionary (first published 1948), lexicography has been increasingly informed by the results of work on syntax in linguistics, but word meaning has not been a particularly fruitful area of linguistic research. So there the situation remains the same. Lexical semantics research can learn from dictionaries, the lexicographer may conclude, but there is little about word meaning that lexicographers can learn from linguistics. If lexicographers were to stray into the theory of word sense distinctions, they would be straying into notoriously treacherous territory; better to stick to the practice.

Two articles in the lexicographic literature directly addressing the variety of senses a word may have are Malakhovski (1987) and Robins (1987). The former presents a taxonomy which distinguishes, amongst sense pairs which are not purely coincidental, distinctions in syntax alone, distinctions in meaning alone, and distinctions in both. But after declaring the hyperlexeme (which includes all those senses which vary from each other in either syntax alone or in a more complex pattern of syntax and meaning) the 'main structural unit of modern English vocabulary' (p 48) the analysis goes no deeper. Robins's article is a discussion of the impossibility of a rigid distinction between homonymy and polysemy. His quest for principles for distinguishing the two is not successful.

A more rewarding pair of articles are Ayto (1983) and Stock (1983). Ayto asks:

> What set of procedures do lexicographers have available to them to pin down those protean entities, 'meanings'? ... How do they decide what, for the purposes of a dictionary, constitutes the meaning of a word, and where, in the case of polysemous words, one meaning ends and the next begins? (p 89)

His response is that the basic tool for the lexicographer is the 'analytic' definition, comprising genus and differentiae. In choosing the genus term, the lexicographer must take care to neither select one that is too general —*entity* would not do as a genus term for *tiger*— nor too specific, if the specific genus term is likely to be unknown by the dictionary users. Where two meanings of a word have different genus terms, they need treating as different senses. The next task is to identify the differentiae required to separate out senses falling under the same genus term. He starts by considering the problem in relation to distinguishing between words in the same semantic field. He considers 'things for sitting on' —*seat, chair, bench* etc.—, presents a table showing the features that distinguish them, and discusses how those features provide the differentiae for the definition. He then wished to apply the same approach to the task of identifying senses for polysemous words. Here, though, the task is different. We do not have formally distinct words for which meaning distinctions are to be found. It is not clear from his account how he intends to identify where different meanings amount to "lexicographically distinct senses" (p 93). He discusses *cup*, and argues that there are three senses, one for the 'trophy' sense, one for the varieties standardly made of china or earthenware, and one for the prototypically plastic or paper varieties. But his consideration of the arguments for treating the second and third of these as distinct ends in a welter of open questions.

Stock (1983) is a response to Ayto's piece, and finds it wanting, firstly, in the circularity involved in using different genus terms to identify distinct senses —the lexicographer will only look for distinct genus terms after determining there are distinct senses— and secondly, in that the model cannot be applied to many words. She says:

> although the model seems to work very satisfactorily with respect to concrete nouns referring to fairly common objects in the real world, it is not at all clear that it would be so satisfactory with words which are more abstract, for example *degree* or *culture*,

> or with words which are highly polysemous such as *do* or *say*, or with words of other word classes than verb or noun. (p 132)

Stock was working on the COBUILD project (section 3.3.1 below), and the alternative model she puts forward sets aside genus and differentiae for identifying the senses of a word, looking instead to a corpus of citations of the word in use. In common with the view taken in this thesis, she views the task of identifying word senses as essentially one of clustering the citations into 'types'. She sketches a set of procedures for the job, based on syntactic similarity, collocations and domains. She then considers the various ways in which the word *culture* is used, displaying with citations how it flows between meanings associated with art, sophistication, or an entire society, concluding:

> It is precisely the lack of clarity in our use of the word *culture* which makes it such a handy word to have at our disposal. It offers, as it were, semantic extras just because in most uses its possible meanings are not clearly disambiguated. We use it in a rather 'vague' way. What can the dictionary maker do to reflect this state of affairs? (p 139)

She considers two published dictionary entries for the word, observes the arbitrary nature of the divisions between senses that they make, and considers they are doomed to fail by the inflexible list structure of their entries, which, she says, leaves no way of representing that "there is slippage between some of the senses that they give but not others" (p 139).

3.2.1 Mel'čuk's Explanatory Combinatory Dictionary

The matter of integrating lexicography and theoretical linguistics has more of a tradition in Russia and the former Soviet Union than elsewhere. Weinreich reviewed the 'lexicology' tradition in 1963, commenting:

> To an American observer, the strangest thing about Soviet lexicology is that it exists. No corresponding discipline is officially recognised in Western European or American linguistics ... Soviet textbooks assign to lexicology a prominence comparable to phonology and grammar. (Weinreich, 1980, p 315.)

A major theory exploiting both the lexicographical and lexicological traditions in Russia and other Soviet republics as they then were has been Mel'čuk's Meaning-Text Theory. The theory emphasises the role of a highly-structured, information-rich lexicon, so anticipating the 'lexicalism' of much recent thinking in computational linguistics. The lexicon which forms part of the Meaning-Text Model is an 'Explanatory Combinatory Dictionary' or ECD. ECDs have been written for Russian and French. MTT claims to be a formal theory, and there is a well-defined structure to an ECD entry. Each has three major zones: semantic, syntactic and 'lexical combinatorics'. The semantic zone contains a definition which "is a decomposition of the meaning of the corresponding lexeme" (Mel'čuk & Polguère, 1987, p 265). The terms in the definition are 'simpler' than the term being defined, and if pursued far enough, "the bottom level primitives are reached" (p 265). The syntactic zone contains, for verbs, subcategorisation information of the kind found in NLP lexicons, though the analysis is rather finer-grained than is attained in many NLP systems.

The lexical combinatorics zone is a area where the theory has made some major innovations. Mel'čuk and his colleagues have identified about sixty lexical functions. These are functions from lexemes to other lexemes or sets of lexemes. Their purpose is to describe all those aspects of a language where the distribution of a word cannot be determined by either a word's meaning or its syntax, but only by reference to particular, non-rule-governed facts about which words are found with which other words. In English, we *ask* a question, in French, the verb found with

the translation of *question* is *poser*, usually translated as *put* or *place*. The verb associated with *question* is not predictable from its meaning. The verb to be used must be stated in the lexicon for each language. The ECD has a lexical function, $\mathbf{Oper}_1$, for the relation from a noun to the semantically empty (or at least emptied) verb which it stands as direct object to, so we have in the ECD for English:

$\mathbf{Oper}_1$(QUESTION) = ASK

and in the ECD for French:

$\mathbf{Oper}_1$(QUESTION) = POSER

Mel'čuk and Polguère claim:

> [Lexical functions] and their combinations allow one to describe exhaustively and in a highly systematic way almost the whole of restricted lexical cooccurrence in natural languages. (p 272)

There are different varieties of lexical functions: some such as $\mathbf{Oper}_1$ capture syntagmatic relations, others, such as those linking *city* and *urban* or *top* and *bottom*, capture paradigmatic ones. Some, such as $\mathbf{Oper}_1$ seem highly 'lexical' while others such as **Syn**, linking *calling* to *vocation*, might simply be called semantic relations and others again, such as that connecting verbal *surprise* to the typical qualifier for the verb's first argument —*surprised*— could well be considered syntactic. The ECD certainly adds to our understanding of the kinds of relations existing between words, and the French and Russian dictionaries are valuable resources, though the distinctions between lexical, semantic and syntactic information remain problematic.

The ECD treatment of polysemy is disappointing. The senses for a word, or lexemes as they are called, are treated as a listable set of distinct items. A 'simpler' sense of a word may appear in the definition of another sense of the same word, but there is no acknowledgement of the possibility of usages falling between senses or of senses overlapping or being open to extension. They say:

> An ECD definition must be adequate in the sense that all possible correct usages of the lexeme defined are covered by it and all incorrect usages are excluded. [and in a footnote] By *possible* we mean 'possible in any imaginable context'— with the obvious exception of contexts involving either the phonetic form (as, e.g., in poetry) or metalinguistic use of lexical items. (p 265)

By fiat, the ECD does not provide a framework for accounting for novel senses. It invokes a highly problematic notion of 'correct' usages, and declares it to be 'obvious' where exceptions are permissible.

3.3 Introducing the corpus

An empirical study of word senses requires not only a dictionary to provide an initial set of word senses but also examples of the words in use, in order to start to see whether the lexicographers chose their senses well and to gather information that is simply not present in the dictionary.

This takes us into corpus linguistics. A corpus is, broadly, a substantial sample of a language or sublanguage. Recently, many have been gathered in machine-readable form in order that some computer-based or computer-assisted analysis can be conducted. The original item is the Brown corpus, developed at Brown University in the Sixties, one million words long and comprising

samples of different genres of written, modern American English, carefully selected to be 'representative' (Francis, 1964). Modelled on it but comprising British English is the London Oslo Bergen (LOB) corpus (Hofland & Johansson, 1982).

As the size and processing power of computers, and the bulk and availability of machine-readable text, both multiply, so do the possibilities for corpus-based studies of all language phenomena. They offer the possibility of introducing an unprecedented degree of empiricism to the study of many aspects of language.

A note of caution is required. There are many difficult and unresolved questions concerning what it is for a corpus to be representative of a particular type of language, or of language in general. Different 'types' might include written vs. spoken; spontaneous vs. prepared; variation by dialect; by level of formality; by sex of speaker/writer; by topic; (if written) by genre; and a host of other factors. A major practical problem is how to select and obtain a corpus of sufficient size and of the right language type so that frequency statistics generated from it are of any value. Corpus linguistics is a new area of study, and these questions are only beginning to be addressed. That said, let us consider what corpora offer.

To linguistics they offer a vast supply of all the patterns and structures which occur with non-negligible frequency in a language. They also open the way for quantitative claims, about relative frequencies of constructions, or for claims that some structures simply do not occur. They make it possible to test hypotheses to see whether predicted structures occur (section 3.3.4) and research questions about different genres, about how speech and written language vary, about regional and dialectal variation, using the quantitative methods of social science. As Firth said, "You shall know a word by the company it keeps" (Firth, 1957): with the aid of a corpus, it becomes possible to study the company words keep, or 'collocations' (section 3.3.6), and the 'ecology of language' as never before.

For lexicography, they offer frequency facts, to help determine what should go in the dictionary, and systematic evidence of how words are used (section 3.3.1). Statistical and user-interface tools for computer-assisted lexicography are being developed and adopted at a great pace. First there were simply concordance lines.[1] Concordancing brings all the occurrences of a word together, for comparison. The alphabetical sorting brings together all instances of some collocations, such as *live together* where the target word is *live* and the sorting is on the right contexts. This was a start, but there was much room for improvement. The lexicographer wants the computer to do as much work as possible, identifying likely collocations, (section 3.3.6), preprocessing the concordance lines and sorting them into types (sections 3.3.2, 3.3.8).

For wide-coverage NLP, a corpus of the appropriate type is a sample of what the system must process. The system will need developing alongside a corpus, training on one if it is a 'learning' system, and testing on one. It will be a source of statistics for identifying most likely parts-of-speech (section 3.3.3), syntactic structures (section 3.3.3) and word senses (section 3.3.5).

Corpus and MRD research go hand in hand. The process of bootstrapping a large lexicon for wide-coverage NLP into existence requires a corpus containing samples of the types of text the system is designed to deal with, both for gathering examples of word usages to build the lexicon for the system, and for testing it. But that information is not immediately available from the corpus; a tagged (see below) or parsed corpus is a more helpful source of data, a disambiguated one would be better still. Tagging and parsing are both processes requiring large lexicons, and an obvious source of those lexicons is MRDs.

[1] The concordance lines for a word are all 'lines' of a corpus which contain the word. The target word is generally at the middle of the selected line, or context, and the length of the line is variable but the default is to use the maximum which will fit on a line of computer screen or printout (hence the name).

3.3.1 Corpus lexicography: COBUILD

Collins COBUILD dictionary broke new ground in the 1980s through its commitment to basing dictionary entries on corpus evidence. The criterion for listing a word or phrase was, at least in part, that it occurred with sufficient frequency in the COBUILD corpus. The editors chronicled the difficulties faced and decisions made in the course of the project and published the account as a book, *Looking Up* (Sinclair, 1987). The chapters by Moon and Hanks provide a preview of many of the topics covered in this thesis, from the practical angle of 'should, or how should, this usage be represented in the dictionary'. They discuss briefly, *inter alia*, collocations, lexical fields, subcategorisation possibilities, metonymy, connotations, metaphoricity and literalness. If a word's usages fall into two classes with regard to any of these, there may well be a case for presenting alternative word senses. In COBUILD, as in most dictionaries, the presentation of an alternative word sense is not an all-or-none matter; there are a variety of strategies for showing different sorts of variation from a previously-defined sense (see also Chapter 5). The COBUILD project has been highly influential throughout lexicography, and the principle that dictionary entries should be based on corpus evidence is now widely accepted.

3.3.2 Atkins: semantic ID tags

Atkins, one of the originators of the COBUILD team, pursues the idea that "every distinct sense of a word is associated with a distinction in form" (Sinclair, 1987, p 89) in Atkins (1987, 1990). She extracts all the corpus citations for the word under scrutiny and proceeds to work through them, noting any patterns there may be in the relations between the meaning conveyed and the form of the utterance. In the 1987 article she discusses the word 'danger', and identifies a widely-used sense in which it means an unwanted possibility, as in "the danger of losing the job" which no published dictionary prior to COBUILD had spotted. She introduces 'semantic ID tags', the local, lexical and syntactic clues that lexicographers use to identify what sense a word is being used in. An example: the 'of' phrase (e.g. "of losing a job") is an ID tag for the 'unwanted possibility' sense of *danger*, though a probabilistic rather than black-and-white one. She catalogues the different sorts of ID tags, and how they might be made explicit.

Where the corpus is parsed and semantic ID tags have been established, the concordance lines with the ID tags for different senses can be grouped automatically. (It is important that the corpus is parsed as ID tags will in general make reference to words or structures in particular syntactic relations to the target word, and the syntactic relations cannot be identified in an unparsed corpus.) The lexicographer can verify the ID tags, by seeing whether those concordance lines do in fact belong together and whether other lines with the same sense have been missed. The computer has done the donkey-work of bringing together the evidence for determining what should go in the dictionary entry for the sense. OUP's new edition of the Concise Oxford Dictionary is to be prepared using ID tags and automatic sorting of concordance lines, so lexicographers can rapidly test out ideas about what the identifying characteristics of a particular sense are. The corpus they will work from will have already been parsed.

Atkins's goal is to assist the lexicographer, but the exercise is intimately related to NLP word sense disambiguation as described in the last chapter. What it adds is a role for the lexicographer. Lexicographic expertise at spotting the characteristic patterns for a word sense has not previously been given a major role in NLP systems: now the convergence of interests should benefit both parties. The production of professionally verified, sense disambiguated corpora is a matter of particular interest. They can potentially be used for automatically 'training' and improving NLP sense resolution systems (section 3.3.5).

3.3.3 Tagging and parsing

An essential early stage of language processing, for a general purpose NLP system, is syntactic parsing. But syntactic parsing is an operation, not directly on the words in a text, but on the syntactic categories of those words. The process of assigning categories, or tags, to words is called 'tagging'. Many words (a 'word' is here simply a sequence of letters surrounded by blanks or punctuation) can belong to more than one category. Also, many of the word types in a corpus are not to be found in a large dictionary.[2] So tagging is not a trivial task. The Brown and LOB corpora have been laboriously, manually tagged. For the LOB, as for all subsequent work, a set of tags derived from the Brown one has been used. That being done, there is a source of data regarding which word has which tag in a particular syntactic context. Thus the tagged corpora have been used as training sets for automatic taggers. Statistical taggers have been developed by Garside (1987), DeRose (1988) and Church (1989) and between 95% and 99% accuracy levels have been achieved. Two preference based systems, FIDDITCH (Hindle, 1989) and DILEMMA (Martin, Heymans, & Platteau, 1988) both achieve similar accuracy levels. 99% might seem very good, but it is to be borne in mind that even that success rate for words would lead us to expect to an error in the tags for every fifth twenty-word sentence.

Tagging serves some purposes in itself but is primarily a preliminary to parsing. It is large quantities of parsed text which are required to answer questions in theoretical linguistics about, for example, the circumstances in which various kinds of construction occur, for lexicography, for determining interesting co-occurrence patterns (see below) and, for NLP, as part of the bootstrapping process. Parsed text is required for automatically building grammars and training parsers which can then be used for parsing unrestricted inputs of text in NLP applications, and this is the greater goal. The parsing problem is, therefore, currently receiving a great deal of attention. Currently the two systems which come closest to providing a reliable practical parser for unrestricted input text are Hindle's FIDDITCH, and the system described in de Marcken (1990). Hindle concentrates on arriving at a parse for every sentence, whereas the priority for de Marcken is avoiding errors. Both perform tagging and parsing in tandem, and de Marcken exploits the fact that most of the errors taggers make are on those cases where there was comparable weight of evidence for two or more tags, so in those cases his parser proceeds with both possibilities and the tagging problem is generally resolved (unless it is, in fact, indeterminate) in the course of finding the more plausible parse. De Marcken's system works bottom up, and often it will not have sufficient evidence to connect up the different constituents of the sentence to give a single connected parse for the whole sentence. In these cases its output is simply a set of constituents. FIDDITCH also makes no commitment with regard to where a PP or participle phrase should attach.

Both these parsers trace their intellectual origins back to Marcus (1980) and his NLP-oriented, deterministic approach to parsing, and also to statistical techniques. In Britain a wide coverage grammar and parser have been developed as part of the Alvey Tools program, and these owe more allegiance to theoretical linguistics. The Alvey system seeks more detailed syntactic analyses, and is a better tool for addressing various questions in linguistics. Currently, it is substantially slower than the others mentioned, and this makes it impractical for bulk parsing. In the Chomskyan mould, it operates with parses being acceptable or unacceptable. It will reject ill-formed input, and will not reject parses which are merely unlikely. This again presents problems if it is viewed as a bulk parser of unrestricted input, since one would expect a certain proportion of that input to be ungrammatical. It does, however, make it possible to approach empirically such questions as, 'how fully does this particular grammar cover the language, as represented in this corpus? To what extent is the language we use 'grammatical?' (Turner, Grover, & Briscoe, 1989).

[2] Walker & Amsler (1986) found that two thirds of the word types in a large corpus of news stories were not located in Webster's Seventh. Their analysis was that one quarter were inflected forms, another quarter proper names, a third quarter hyphenated, and the residue mis-spellings or unaccounted for.

The nature of the tagging and parsing tasks and the preprocessing required are discussed in various articles in Garside, Leech, & Sampson (1987), Meijs (1987) and Aarts & Meijs (1990).

3.3.4 Corpus as testbed

For the taggers, parsers and statistical models described above corpora are needed to test the theories as well as to build the models. This raises chicken-and-egg problems. While a corpus can be used to develop a parser, according to the best linguistic practice, it is no easy matter to determine whether it is doing the job properly, as there does not currently exist a large quantity of accurately parsed text against which the parser's performance can be compared. It is the goal of the Penn Treebank Project (Marcus, Santorini, & Magerman, 1990) to produce such a test corpus. The method is pure bootstrapping: as it is found to be far quicker for people to check a candidate parse than to come up with one themselves, a parser is used on a tagged corpus and trained people check and correct the computer's output. If certain kinds of mistake are consistently made, the parser or grammar can be improved. With the aid of a special purpose interface, the production of a corpus of parsed text or 'treebank' is well under way and soon there will be treebanks for testing parsers as well as parsers for building treebanks.

That is one use of a corpus. More and more, corpus evidence is required to support theory across linguistics. The researcher must ask, is the phenomenon found in the corpus? Do the circumstances in which it occurs tally with what the theory predicts? What proportion of the occurrences does the theory elucidate, and are the exceptions damaging to it? Whittemore, Ferrara, & Brunner (1990) test how successful different strategies for attaching prepositional phrases are, using a corpus of examples. Briscoe, Copestake, & Boguraev (1990, discussed below) trawl the corpus for all occurrences of a particular family of verbs, to test their theory.

Many varieties of linguistic hypothesis could be tested using a corpus. As yet few have been. Corpus-based studies provide an opportunity for the empirical examination of many linguistic questions, and it is to be expected that a wider and wider range of issues will be broached using corpus-based methods.

3.3.5 Corpus-based sense disambiguation

As corpus analysis tools are developed, so the corpus becomes a source which can be exploited for sense disambiguation. McRoy's work (section 2.2.6) used some corpus-based strategies. Hearst (1991) is highly corpus-based. The program is restricted to homograph disambiguation: it does not attempt to discriminate between all the senses given in a dictionary, but only between those that are markedly different. The program works from a list of clues for disambiguation, which might be called 'potential semantic ID tags'. These include whether the word is capitalised, whether it modifies another item, and whether it is found in a PP headed by one of *in, on* and *of*. In the learning phase, the program is fed a substantial number of corpus citations with the target word disambiguated. By seeing which potential clues apply, with what frequencies, to citations of the different senses, the program is able to determine which items of information are salient for disambiguation. Once the relative importance of the clues has been established, new inputs can be fed to the program which will calculate and compare the evidence for each of the possible senses for the word. Hearst tests her system, and finds the results "comparable to, or better than earlier efforts using MRD's and large corpora" (p 19) such as Lesk's, Guthrie et al.'s and Veronis and Ide's. She briefly notes the difficulties of making quantitative assessments and comparisons.

3.3.6 Collocations

A collocation is a group of two or more words which are to be found in proximity to each other significantly more often than one would predict, given the frequency of occurrence of each word

taken individually. They may or may not be immediate neighbours, and the meaning of the whole may or may not be fully determined by the meanings of the parts.

Many collocations are arbitrary: there is little to be said about why we *do exercise* and *make love* rather than $*$*make exercise* and $*$*do love*. Yet using the appropriate collocations is a major part of fluency in the language. Thus they are important for language learners, and to natural language generation programs. They are also crucial for natural language analysis systems, both because collocations shade into idioms so a word in a collocation is often not being used in a way described by its dictionary senses, and because its collocates will often resolve which sense a word is being used in (as seen in section 2.2.6). Observing the language learner's needs, Benson, Benson, & Ilson (1985) and Cowie & Mackin (1975), Cowie et al. (1982) have both produced dictionaries of idioms and collocations: general-purpose learners' dictionaries also list substantial numbers, and in an ECD, the syntagmatic lexical functions specify them.

Lexicographers want to identify them from corpora, and rank them according to suitability for noting in dictionaries. Tasks for interested NLP researchers include identifying them, but also representing them in ways that make them accessible and usable to analysis and generation systems. The task is complicated by the range of ways in which collocations are subject to variation:

- Modification: Could you **do me a (great) favour**
- Agreement: I/you/he **shot my/your/his bolt**
- Syntactic variation: If you will **tempt/go tempting fate** like that
- Lexical variation: He was a great one for **tempting fate/providence**.

As the level of variation increases, so the 'setness' of a set phrase diminishes and it becomes less of a clear case of a collocation.

3.3.7 Lexical Statistics

Two tools for identifying collocations which have been proposed just in the last three years and are already widely in use are 'mutual information' (Church & Hanks, 1989) and t-test (Church, Gale, Hanks, & Hindle, 1991) statistics. The mutual information statistic is a development of the simple idea that, where two words are often found together, this is potentially of interest to the lexicographer and should be noted in the dictionary. The statistic is calculated for any two words by comparing the number of times they occur together in a corpus with the number of times they would be expected to co-occur, if the words were randomly distributed in the corpus and so only co-occurred by chance.

Where large computers are available, mutual information scores for all word-pairs for a given corpus can be calculated. Then the lexicographer can inspect those words with the highest scores, when paired with the target word. These will generally include candidates for listing in the dictionary as collocates with the target word, and will suggest other salient information such as what example sentences might be chosen to indicate typical uses of the target.

A good dictionary will not only indicate where it is appropriate to use a word. It will also indicate where it is not, although one might have thought it was. An example dating back to Halliday's early article (Halliday, 1966) is the distinction between *strong* and *powerful* in the context of *tea*, or, in a more recent empirical investigation of these words, *support* (Church et al., 1991)[3]. The intuition is that *strong tea* and *strong support* are both common collocations, whereas *powerful tea* and *powerful support* are not. Mutual information statistics can provide two lists of

[3]The authors apologise for changing the example from *tea* to *support*, explaining that their corpus was of journalistic writing, in which discussions of *tea* were rather infrequent.

co-occurring words for near-synonyms such as *strong* and *powerful*, but while the presence of a word in both lists is evidence of a type of usage that both words share, the absence of a word from one of the lists is only the weakest of evidence that the use of the two words differ in respect of the third. A sharper tool for investigating how, or whether, two words differ in respect of their frequency of co-occurrence with a third is the t-score, which looks at the difference in frequency of occurrence of, e.g., *strong tea* and *powerful tea* and counts how many standard deviations away from each other they would be if they were both part of the same population. It can serve to highlight, for lexicographers, contexts for which words of similar meaning have different distributions and can thus lead to greater definition in those parts of the dictionary entry where a word is set apart from its semantic neighbours.

With the automatic language generation goal in sight, Smajda & McKeown (1990) have developed a program, Xtract, for extracting collocations from a corpus and representing them in a form convenient for re-use in a language generation program. This use requires rather fuller accounts of the collocations than does lexicography. The generation system will need to know whether the words in the collocation must, or may, occur in a given order, with a given array of other words falling between the words in the collocation. For collocations involving verbs, it will need to know what sort of argument slots the phrase leaves, where in the string they are to be realised, and what selection restrictions apply. Xtract does not merely identify collocations; it automatically builds lexical entries for a phrasal lexicon. Zernik & Dyer (1987) describes a similar project, though not in the context of corpora.

3.3.8 Hindle: combining parsing and statistics

The mutual information and t-test statistics, when applied to unparsed text, look only at fixed two-word strings. Results of interest relied on there being a grammatical relation of interest -—usually of modifier to head— between a word and its neighbour. But for most grammatical relations, we do not expect to find the two related words next to each other. The text must be parsed for word pairs standing in particular grammatical relations to each other to be identified. Then the statistical tools can be applied to populations of grammatically-related word pairs, and the lexicographer can gain a far fuller picture of the kinds of settings the word is to be found in.

A paper pursuing this line of research is Hindle (1990). His premise is that semantically similar words tend to co-occur with the same collocates, so if we can develop a metric for similarity of behaviour according to which words occur in certain grammatical relations to which others, and can then apply this metric to large quantities of text, we can obtain an empirical base for a semantic classification scheme. He uses FIDDITCH for parsing several million words of text. The grammatical relations he studies are those relating subject, verb and object. The parser provides him with a very large number of subject-verb and verb-object pairs. In the paper he discusses the verb *drink*. He finds the nominals paired with *drink* in verb-object pairs, and for each of these, determines whether it has a high 'mutual information' score with regard to *drink*. This establishes a prima facie semantic class of 'those things which are drunk'. But that alone provides only weak evidence for semantic similarity. If two nouns, e.g. *wine* and *beer*, are similar we shall expect them to be similar not just in regard to the one verb, *drink*, but also with regard to the bulk of the other verbs in the language. So now we look at all verbs with a tendency to co-occur with the one, and ask if they co-occur with the other. Hindle defines a measure of the similarity of two nouns as follows.

1. For every verb:
   (a) For both of the nouns, count the number of times it features as object of the verb;
   (b) Calculate the mutual information score of the verb with each of the two nouns featuring

as its object[4]

(c) Take the minimum of the two mutual information scores [5]

(d) Repeating stages a–c but with 'object' replaced by 'subject';

(e) Sum the results from stages c and d.

2. Sum the results of stage (e) for all verbs, and this gives a value for the similarity between the two nouns.

If this calculation is conducted for every pair of nouns, each word can be assigned a 'nearest neighbour'. A nearest neighbour is the word for which its similarity score is highest. Where the nearest neighbour of X is Y and the nearest neighbour of Y is X, we have a pair of reciprocally nearest neighbours (RNNs).

Hindle presented the results of an experiment on a 6 million word corpus of AP newswires. He calculated similarities between all the nouns which occurred more than 100 and less than 200 times in the corpus, 319 in number. The RNN pairs for this corpus include *bomb:device* (this was a corpus of journalese); *ruling:decision; street:road; list:field; debt:deficit.*

3.4 Theoretical linguistics

We may follow Zgusta, cited above, in a measure of scepticism as to how much the theorists of language have had to offer regarding the identifying and distinguishing of word senses, yet linguistics includes a sub-discipline, lexical semantics, devoted to the meanings of words. In descriptive lexical semantics questions about multiple word senses have received little attention. In a textbook of the area, Cruse (1986), the only references to polysemy and homonymy occur in a single short paragraph where it states that these terms

> although innocuous if used circumspectly, are not entirely ideal for our purposes (p 80)

Multiple word senses have been of most interest to linguists when the relations between the senses have shown some interesting structure. Linguists have tended to see a different aspect of polysemy to NLP, for which the cases where senses were unrelated presented most problems. In an early contribution to the field, Apresjan (1974) introduces the term 'regular polysemy' and lists regular alternations to be found in the lexicon of Russian. In his textbook, Leech (1981) places 'semantic transfer rules' alongside derivation morphology as one of the productive processes whereby words may shift from one type of use to another. The phenomenon has often been observed, and almost as often, named: 'sense extension' (Pustejovsky, 1991a; Copestake & Briscoe, 1991), 'lexical implication rules' (Ostler & Atkins, 1991), 'subregularities' (Wilensky, 1990), 'systematic polysemy' (Nunberg & Zaenen, 1991) and, for verbs, 'diathesis alternations' (Atkins, Kegl, & Levin, 1986), to mention but a few. This thesis adopts 'alternation', and considers the phenomenon and selected literature in Chapters 8 and 9.

3.4.1 Ellipsis

Givon (1967a, 1967b) made an early investigation of alternations. His analyses proceeded on the basis that a wide range of apparent patterns of polysemy should be treated as the product of ellipsis followed by reinterpretation. Thus

[4] Here the population of all verb-object pairs forms the statistical base for the mutual information calculation.

[5] This stage is in fact slightly more complex as mutual information scores can be negative. Hindle's approach is to take the absolute value of the maximum of the mutual information scores if both are negative (i. e. both show a lower than expected correlation with the verb), otherwise 0.

We bought three cheeses yesterday

is to be interpreted as an ellipted version of

We bought three kinds of cheese yesterday

with *cheese* reinterpreted as a count noun in a process whereby it absorbs the COUNT feature from *kind* to render the elliptical version grammatical. This account is presented as an analysis of the alternation patterns, and it is not specified whether the claim is that the ellipsis and derivation happened or happens in the history of the word, or in the course of the child learning the word, or whenever the word is processed (or some other possibility). From an NLP perspective, it is clear that reconstructing elliptical material is at best a roundabout way of specifying a very flexible kind of co-ordination, and at worst an entirely unconstrained piece of invention.

There is a range of cases in territory indeterminate between ellipsis and polysemy. In a conversation with a carpenter about making tables from various sorts of woods, both analyses are available for:

My last customer had one which was made of apple.

Either *apple* is elliptical for 'the wood of the apple tree' or the usage-type of *apple* is one in which it means 'wood of the tree the fruit grows on'.

We might look for clarification regarding what is and is not legitimately treated as an elliptical form. Quirk et al. (1972) say:

> In a strict sense of ellipsis, words are ellipted only if they are uniquely recoverable, *ie* there is no doubt about what words are to be supplied ... What is uniquely recoverable depends on the context. (p 536)

This serves to clarify that the count sense of *cheese* was not an elliptical use (in the strict sense) since *kind* might equally have been *sort*, but only goes part of the way for the job in hand. Whether the material is uniquely recoverable may be a necessary condition, but it does not provide a sufficient condition for invoking an analysis involving ellipsis. It does not answer the question, "*Should* we go about seeking a unique, reconstructed version, or do we simply deem the usage as belonging to a distinct usage-type?" There will still be a choice between *apple* being polysemous or being an ellipted form where, at some time in the linguistic or pragmatic processing, the *wood* element must be inferred if the sentence is to be fully interpreted.

3.4.2 Verb classification

The delicate relations between syntactic structure and case have long been an intriguing topic for linguists, from Fillmore's (1968) *The Case for Case*, through the generative semantics tradition, Dowty's (1979) *Word Meaning and Montague Grammar*, Perlmutter's 'unaccusative hypothesis' and all the ensuing debates. 'Grammatical relations', 'argument structure', and 'thematic structure' are three of the many names for the field. Two recent pieces of work that make clear the links with polysemy are Atkins, Kegl and Levin (1986) and (1988). They have looked at alternations for verbs, where the mapping from grammatical roles to case roles changes, as between

John is baking the cake.

and

The cake is baking.

They identify six different possible relationships between transitive and intransitive forms of a verb. Sometimes more than one of these applies to the same verb ("John is baking."). The pattern or patterns the verb participates in depend largely though not entirely on the core meaning of the verb. Levin (1991) argues:

> As the distinctive behavior of verbs with respect to diathesis alternations arises from their lexical properties, specifically their meaning, the exploration of the ways in which diathesis alternations distinguish among verbs should reveal semantically coherent verb classes. Once identified, these classes can be examined to identify the components of meaning common to verbs participating in particular alternations. These components of meaning would be expected to figure prominently in the lexical representation of the meaning of these verbs. (p 209)

Thus she sees the study of alternations as a route toward determining the structure of an ideal dictionary. The theme is explored in detail in Chapter 9.

3.5 Cognitive linguistics

Researchers in cognitive linguistics study linguistic phenomena from the perspective that linguistic knowledge is not separate from general knowledge but integrated with it. The same principles may be expected to govern the structure, relationships and uses of non-linguistic and linguistic phenomena. Insights from language may be expected to shed light on cognitive structure, and findings from psychology may serve to explain structures in language.

Word meaning is a matter inviting study within this paradigm. A word is, to a first approximation, a linguistic realisation of a concept. Within cognitive psychology, Rosch & Lloyd (1978) and followers have developed 'prototype theory'. A range of psychological experiments have demonstrated that people view some instances of a concept as 'prototypical' of it, with others more or less marginal. Prototype theory is then a development of the idea that human concepts are very often 'prototype' concepts, and addresses topics such as how prototypes combine, how they are related to perception, and how people reason with them. The cognitive linguistics program includes taking this model and applying it to word meanings. Where there are multiple senses for a word, the hypothesis runs, they correspond to different aspects, modifications or non-standard types of instance of a 'prototypical' concept. Lakoff (1987) famously analyses the various extensions and modifications of *mother*, finding the concept to have 'radial structure', with various senses connected, as by the spokes of a wheel, to the prototypical mother at the hub.

Following the work of Lakoff & Johnson (1980), metaphor is a central topic in the field. Lakoff and Johnson sketch out a range of ways of 'seeing one thing as another', such as 'life as a journey' and 'time as space', and provide evidence that these large-scale, systematic, 'conventional' metaphors are pervasive in both everyday and poetic language. They give cross-linguistic evidence to support the claim that the most general and pervasive of the metaphors are universal. They show that many cases of polysemy, and most cases of novel uses of words, can be interpreted as the outcome of applying a conventional metaphor in a particular case. The ideas suggest how a computer system might understand metaphor: it would find which of an inventory of conventional metaphors best suited the circumstances, and then apply the mapping associated with the metaphor to the problem word. The strategy has been pursued in work described in section 3.6.1 below.

3.5.1 Polysemy or pragmatics?

Ruhl (1989) devotes a book to the topic of monosemy, arguing that many words taken by lexicographers to be polysemous should be considered to have only one sense. Variations in how they are

to be interpreted should be assigned to the pragmatic interpretation of the word in context. He presents case studies of five words which are usually treated as highly polysemous, and shows, for each, that different uses fall on a continuum, or that the word's patterns of use are shared with other words. He treats this data as evidence that a distinction is contextually specified and not a matter of polysemy. Effectively, he conflates polysemy with homonymy, as he considers any evidence that one usage-type is predictable from another as evidence for monosemy, so the only cases which might be described as polysemy are those which might also be described as homonymy. He does not seriously consider how contextual and lexical information might combine, signally failing to provide any account of either the "common core" or "single, highly abstract meaning" that apparently polysemous words share, or, but for occasional, sketchy remarks, of the contextual processes that act on that core to give rise to the various readings of a word. As he says,

> [A colleague] observes that the book is highly programmatic: it claims much more than it substantiates. ... I plead guilty ... (p 236)

Wierzbicka (1987) argues that much of what some authors have considered to be pragmatic implications are specific to certain words, word-classes and constructions, and need specifying in the lexicon. Her stalking horse is Levinson (1983), a more careful protagonist than Ruhl. Her objections are essentially as above: claims are often made for pragmatics, to the effect that variations in how a word or phrase is to be interpreted can be accounted for on the basis of general principles of reasoning and communication. However the claims are often made without a full awareness of the complexities and specificities of the behaviour of the word or phrase, and are only rarely accompanied by an account of how the general principles give the desired interpretation. She considers sentences with the superficial form of the tautology "X BE X", such as "Boys will be boys". Levinson's scheme is to note that, like any tautology, taken literally it is semantically empty; when they are spoken, the hearer will (consciously or unconsciously) realise this and seek an explanation as to why the speaker had made a remark that, superficially, communicated nothing; the appropriate interpretation of the speaker's intention in making the utterance will then unfold from the interaction of the sentence with the conventions governing conversation, Grice's maxims. However, as he admits,

> Incidentally, exactly how the appropriate implicatures in these cases are to be predicted remains quite unclear, although [Grice's] maxim of Relevance would presumably play a crucial role. (p 111)

With evidence from the different, or non-existent, roles such sentences play in other languages, and the remarkable contrasts between the implications of sentences such as the following:

> Boys will be boys
> War is war
> Promises are promises
> A promise is a promise

Wierzbicka makes a convincing case for the fact that the central role in interpreting "X BE X" sentences will always be played by semantic specifications for words, groups of words, or constructions.

3.5.2 Geeraerts on *vers*

A study located on the cusp between cognitive semantics and lexicography is Geeraerts (1990). The core of the article is an analysis of the meanings of the Dutch word *vers*, and a comparison

between this analysis and the entry in the *Woordenboek der Nederlandsche Taal*, a large scholarly dictionary following the pattern of the OED, but rather more modern and more detailed. Geeraerts's starting point is that word meanings are 'prototypical' concepts, exhibiting 'prototypical polysemy'. The various senses of a polysemous word will be overlapping and with fuzzy boundaries, though there will be some usages which unequivocally fit that sense. He presents his analysis of the word as a diagram of intersecting and nested boxes, and the various boxes and zones within them are labelled as more general and more specific types of meaning. He then presents a diagram showing the structure of the *Woordenboek* entry, not in its surface, linear form but in the hierarchical form indicated by the changes of typeface and nesting within the entry. He finds the lexicographers have in fact exploited a variety of mechanisms to indicate, within the constraints of a verbal and superficially linear entry, the kinds of inclusion relations and overlapping boundaries to be found in his diagrammatic analysis. The hierarchy of the *Woordenboek* entry is only loosely a taxonomy; *vers* is translatable most of the time as 'fresh', and the Dutch equivalent of 'fresh' as in 'not preserved' is below 'both new and in optimal condition' in the hierarchy, although fresh (i.e. 'unpreserved') fish undoubtedly can be in a less than optimal condition. He concludes that the practice of lexicography has been both taking due note of, and finding ways of representing prototypical polysemy; that there are, in this fact, the seeds of a rapprochement between lexicography and studies of cognitive structure; and there is much for cognitive science to learn from the large, scholarly dictionaries, quite apart from any benefits that might flow in the opposite direction.

3.6 Language, mind and artificial intelligence

Cognitive linguistics takes the grand picture of language and mind as its subject matter. So does artificial intelligence (AI). For an artificially intelligent agent to fully interpret linguistic inputs, it must bring conceptual knowledge of both linguistic and non-linguistic varieties to bear on the input. Two authors who have investigated the relations between lexical knowledge and the process of interpretation from an AI perspective are Martin and Wilensky.

3.6.1 Martin's metaphors and Wilensky's subregularities

A major challenge for AI is language understanding, and within that, a major challenge for the NLP component of any AI system is interpreting words used in ways which are not described in the lexicon, 'misfit' usages. But a misfit is only a misfit relative to some system or subsystem which failed to interpret it (or interpreted it wrongly). So questions about how misfits might be dealt with can be asked in a specific way only relative to a system, and the answers are only likely to be interesting when the system has a sufficiently sophisticated parser, lexicon, and semantic and pragmatic interpreter so that usages are likely to be misfits for linguistic rather than system-specific reasons.

A system which goes further than most towards meeting these requirements is the Unix Consultant, produced at Berkeley (Wilensky, Mayfield, Chin, Luria, Martin, & Wu, 1988). For the most part, it works within the sublanguage of queries about computer systems, but within that domain, when it cannot find an interpretation, the reason is often that a word is being used in a way that requires an extension to the range of its usage-types.

Martin has developed MIDAS, a metaphor interpretation system (Martin, 1990). It is premised on the idea in Lakoff & Johnson (1980) that metaphors are, by and large, conventional and systematic. Thus a misfit usage, if it is being used metaphorically, will be interpretable by reference to an already-familiar metaphor-structure. If MIDAS is already familiar with both the expression "kill a process" (on a computer) and the structure of the metaphor underlying it, it can interpret "the process died". MIDAS is an implemented, working system. For it to

have not merely this sense of *kill*, but also the metaphor structure, available in its lexicon, it requires a knowledge representation system in which the connections between core senses of *kill* and *die*, their prototypical domains, and the 'computer' domain, can all be represented and used in inference. With this behind it, MIDAS gathers up the set of possible metaphorical and other interpretations for a word, using knowledge of lexical entries, of the domain under discussion, of the prototypical domain for the word, and of conventional metaphors relating the domains. It then applies a metric to compare the different possible readings, and arrives at an interpretation.

The psycholinguistic findings of Gibbs (1983, 1984, 1986) indicate that, when processing conventional metaphor and idioms, people do not first test out a literal interpretation[6] and only then move on to non-literal ones: the time taken to process conventional metaphor and idioms is no longer than that taken to process straightforwardly compositional sentences. In the spirit of these findings, Martin would in principle evaluate all possible interpretations, 'literal' and otherwise, prior to selecting a reading. In practice, he finds this slows the system down unreasonably, and a policy of seeking metaphorical and other readings only when 'literal' ones are unacceptable is adopted.

Wilensky's work (Wilensky, 1990) is more general than Martin's. He uses the concept 'lexcon' for our usage-types, and interprets metaphoric links or 'subregularities' relating lex-cons (in our terminology: alternations) as just one of several types of linguistic subregularity. He lists a range of subregularities, and notes that there are many of them. He does not discuss how the set of subregularities might be learnt, or whether there are any constraints on what subregularities might be learnt. He considers the potential for extending the kinds of algorithms in MIDAS to a much wider range of cases. His papers, unlike Martin's, do not present an implemented theory but rather, a wide-ranging picture of the nature of word-meaning which has many implications for the kinds of ways NLP should approach lexical acquisition and representation.

3.7 The Generative Lexicon

A recent proposal which has provoked much interest is Pustejovsky's 'Generative Lexicon' (Pustejovsky, 1991a). The Generative Lexicon includes both 'core' entries for words and a system of rules for inferring further senses from basic ones. The themes are close to the topic of this thesis. Pustejovsky has looked particularly at nouns, arguing that in the past (and as indicated in the discussion in section 3.4.2 above) the semantics of verbs has been studied in details whereas nouns have tended to be seen as relatively simple objects which simply fill slots in subcategorisation frames specified by verbs. He proposes that nouns have 'qualia structures', similar to the 'frames' of knowledge representation languages (e.g. Bobrow & Winograd (1977)), containing slots for the different aspects of meaning that might be salient in different sentences. The slots will be common to a wide range of nouns, but the values filling those slots will be lexically specified (either directly or through inheritance). Some 'extended senses' of the word can then be interpreted as usages in lexical and syntactic contexts which foreground one particular role in the noun's qualia structure. He calls this process 'logical metonymy' and says that the noun is 'coerced' into being the type of entity the context demands. The ideas are presented as a formalism in Boguraev & Pustejovsky (1990).

An example of 'coercion' is the title of Briscoe, Copestake & Boguraev 1990: *Enjoy the Paper*. *Enjoy* is a verb which subcategorises for either an '-ing' verb phrase or a noun phrase. This has meant that often, in computational lexicons, the word has been given two entries, one for each syntactic pattern. But there is much in common between the two, and it is a major goal of the

[6] 'Literal' is not an ideal term. The complexities and confusions embedded within it are discussed at length in Wilensky (1989) and Way (1991). Here, a literal reading is one in which the sentence is interpreted on the basis of the basic meanings for each of the words in a string, combined according to basic syntax and semantic-composition rules.

Generative Lexicon enterprise to avoid redundancy by specifying generative mechanisms by which non-basic senses are implicit in the lexicon without being explicitly listed.

Part of what the usage-types have in common is that, in both cases, the phrase subcategorised for denotes an event. Verb phrases in -ing typically denote events, as do noun phrases headed by *event*, *party*, *flight* and the like, so *enjoy* constructions with these elements need no further explanation. However, *the paper* ordinarily denotes a physical object (or a non-count mass). When appearing as object of *enjoy*, it must be 'coerced'. The entry for the relevant sense of *paper* in the dictionary makes reference to reading (and writing), and introspection confirms that the default event associated with a paper is reading (or possibly writing) it. Composition with a verb such as *enjoy* foregrounds this information and makes available to the dictionary reader, or NLP system with a Generative Lexicon, a 'reading(/writing) the paper', event-designating sense of *the paper*.

Pustejovsky calls the purpose typically associated with a noun its 'telic role', where TELIC is one of the slots in the qualia structure. Telic roles are explored further in an experiment described in the Briscoe et al. paper. They investigate how widely the analysis given for *enjoy the paper* can be applied. They take the group of verbs, including *enjoy* and also *prefer*, *finish*, *start* and *miss*, which take both NP and infinitival or progressive VP complements. The hypothesis is that when the verb takes an NP and the head noun is not a word for an activity, it will generally be a word for an object which is either designed for or otherwise intimately associated with a particular activity, and thus has a well-defined telic role. In 'unmarked' usages, where there is no event-type supplied by the context, the original sentence needs reinterpretation if the clash between event-demanding verb and non-event-denoting object is to be resolved. The solution is to interpret the sentence as an expanded version in which the present participle of the default activity verb is inserted between the original verb and object. Hence, since *cigarette* has the telic role of being smoked and *film* of being watched, "finished his cigarette" becomes "finished smoking his cigarette" and "missed the film" becomes "missed watching the film".

They test the theory by trawling the LOB corpus for examples of the verbs under scrutiny, and then examining the examples with non-activity noun phrases as complements to see if there is a telic role associated with the head noun, and if so, whether this was what was being enjoyed, preferred or finished. Their results support the hypothesis. There was a readily available telic role for nouns appearing in this direct object slot in most cases, and the exceptions did seem to be 'marked' cases in which the context supplied an event-type.

According to Pustejovsky, coercion and corresponding logical metonymy do not apply only to this case. Adjectives such as *fast* are also taken to modify the telic role. If the sense of *fast* in which it modifies motion is taken as primary, then, in "a fast car" it is the driving that is fast (in this primary sense) and in "a fast typist" it is the typing (which is the telic role of the word if not of the typist his/herself!). Pustejovsky describes three other roles, alongside the telic one, and considers various constructions alongside the *enjoy* one which, he claims, coerce nouns into one or other of these roles. However the combination of telic role and verb of the *enjoy* class is the only construction where the case for logical metonymy has been worked out in detail and empirically tested. For *fast* and other examples presented in Pustejovsky (1989) it is not clear that either the roles which would need to be in the lexical representations for the nouns, or the categories of construction that would invoke one of those roles, are as yet sufficiently well defined for the ideas to be testable.

The theory involves certain theoretical commitments. Firstly, lexical representations for nouns will have to have a structure within which a telic role can be expressed. Secondly, the notion of 'event-designating' acquires special status. Rather than EVENT being simply one of a number of semantic features or selection restrictions which exclude, or determine preferences among, certain interpretations, here both noun phrases and verb phrases can be of the same, event-designating type, and whether a noun designates an event affects how semantic composition proceeds. The semantic category of 'event-designating' is seen as underlying, and predicting, certain syntactic

phenomena. Thirdly, semantic composition must be conceived sufficiently broadly so that it can incorporate the process of determining, from the lexical entry for a noun, which role (i.e. telic, or other) is required in a particular context.

The Generative Lexicon brings together several themes in contemporary work on lexical semantics, computational linguistics and NLP. It is still very much a set of ideas and small scale implementations in an early stage of development, and many of the ideas are still to be worked out, but does present an intriguing list of desiderata for NLP lexicons.

Chapter 4

Theory I: What is a Word Sense?

4.1 Introduction

In learning a language, the learner arrives at the meaning of a word through identifying what there is in common to the role the word plays in all the various discourses he or she has heard it in. We follow situation semantics (Barwise & Perry, 1983) in taking the meaning or meanings of a word to be an abstraction from the role(s) it systematically plays in discourses.

Pollard & Sag (1987) present the view as follows:

> According to situation semantics, the world is made up of such things as *individuals* (like Jon Barwise or the moon), *properties* (such as being a cookie or being a donkey), *relations* (such as seeing or kicking) as well as *situations*. Roughly, situations are limited parts of the world which consist of individuals having (or not having) properties, or being (or not being) in relations. An example of a situation is the particular event of Carl Pollard eating a certain orange in Office D-2 at Ventura Hall, Stanford University, at 9:42 p.m. PST, December 2, 1986. Individuals, relations, properties, and situations are real, but different groups of organisms are *attuned* to different ones in accordance with the exigencies of their ecology; as it is sometimes put, different communities of creatures "tear the universe apart along different seams."
>
> What does this have to do with meaning? According to situation semantics, meaning arises from *constraints* that hold between different kinds of situations. For example, any situation that has smoke in it is part of a situation that has fire in it. We say that smoke *means* fire; any organism that is attuned to this constant can pick up from a smokey situation the information that there is fire. So it is with language, but in that case the constraints involved are not *natural* ones; rather, they are *conventional* linguistic constraints that can be exploited by the people that are attuned to them, as when an English speaker acquires from an utterance of "Here is a cookie" the information that there is a cookie. Linguistic meaning, then, is a relation that holds between types of utterance situations and the types of things in the world that utterances describe. This view is called the *relational theory of meaning.* (pp 4–5)

This account does not yet makes any mention of what is often seen as the core issue: the relation between *cookie* and the cookies in the world. Pollard and Sag first describe de Saussure (1915)'s terminology (based on a mentalist perspective) in which:

> a *sign* is a mental associative bond between two component mental objects, the *signifiant* (the signifier) and the *signifié* (the thing signified) (p 2)

and then reinterpret it in realist terms:

> the *signifiant* is a certain type of utterance situation, namely one where "cookie" is uttered; the *signifié* is a certain property of things in the world, namely the property of being a cookie. And the sign itself is not a psychological association but the real-world linguistic-meaning relation (a constraint) between the *signifiant* and *signifié*. On this view, a natural language such as English is not a shared mental system but rather a type of linguistic-meaning situation in which certain conventional constraints are observed. The English "cookie" sign, then, is a subtype of English linguistic-meaning situation: the type of situation where "cookie" is used to mean cookie. (p 5)

A word's meaning is learnt when a speaker uses the word in accordance with the constraints, so his or her utterances of the word bear the same relations to situations as other speakers'. For lexicography, the task is to identify the constraints holding between a word and the situations where it is used. This involves looking at populations of usages, and seeing what, in addition to the occurrence of the word in question, they have in common.

We shall follow Lyons (1977) in using 'denotation' to refer to the part of the world corresponding to a word, whereas *referent* will be used for the entity a word relates to in a particular utterance.

Formal semantics generally operates with a model in which denotations are assumed as the entities to which compositional rules apply. For example the referent of a definite noun phrase such as "the cookie" is a member of the set denoted by the head noun (*cookie*). The denotation is a given, from which the analysis proceeds. However for our purposes the denotation cannot be a given. Do different senses have different denotations? The question merely reiterates the puzzle of the nature of polysemy. While the meaning/denotation distinction is of central importance for model theoretic semantics, it is of little relevance to the investigation of polysemy.

4.2 Nunberg's Referring Functions

Nunberg (1978) shares the concern that a view of meanings as denotations leaves much unsaid. He analyses the view that a lexical entry is a specification of the things and classes of things that a word can normally be used to refer to. He takes examples such as:

(1) Hearst bought a newspaper.

and points to the ambiguity between it being the copy or the company that Hearst has bought. If a lexical entry specified a class of things, then since the class of newspaper copies and the class of newspaper companies are so clearly disjoint, either a disjunctive entry or two distinct entries will be required. But one can readily imagine scenarios in which *newspaper* refers to a reporter, an editor, a particular variety of paper, the type (as distinct from the tokens, or individual copies) of a newspaper (as in 'the newspaper is printed in Wapping'), a business deal ('I know the property purchase is of more interest to you, but if we go back to the newspaper...') or an unending list of other possibilities. So the lexical entry would need to be of correspondingly indefinite length.

Moreover, the ostensive act of pointing at or otherwise indicating a newspaper carries with it an equivalent range of possible interpretations. Thus:

(2) Hearst bought that. (said when looking at a copy of a newspaper)

carries just the same ambiguity as (1). Determining the referent in a statement such as (2) is often considered part of pragmatics, whereas specifying the class that the referent of *newspaper* in (1) belongs to is considered a lexical matter. Since there is the same range of possibilities for

interpretation in both cases, a theoretical framework in which the two are treated differently is unsatisfactory. Pragmatics is an area of language-understanding requiring general-purpose inference mechanisms, whereas the lexicon is often seen as supporting only simple look-up procedures. Nunberg shows that both pointing and a choice of a word can serve the purpose of determining reference, and that in both cases, inference mechanisms are required.

The path from the word *newspaper*, or the object pointed at, to the referent can be simple or complex. It may involve any aspects of the verbal or non-verbal context, or of mutual knowledge of speaker and hearer. The task of specifying the lexical entry is doomed, unless it is accepted that there is always a further job to be done, for which inference is required, of establishing what the function is that maps the lexical entry onto the referent.

Nunberg calls this function the 'referring function' or RF. He distinguishes ostensive reference, as in (2) above where a 'demonstratum' is pointed at, from descriptive reference, where words do the work. For ostensive reference, an RF is a function from the demonstratum to the referent. For descriptive reference, the argument of the RF is the meaning of the word, an altogether less straightforward entity.

To consider ostensive reference, as in (2) above, first: the argument of the RF is just the physical copy of the newspaper. In the reading where Hearst bought the copy of the newspaper, the RF is identity. But in the other case, it is a function from newspapers to the companies that publish them. Nunberg proceeds to seek out the conditions for successful reference, by investigating what the constraints are that enable a hearer to identify the RF intended by the speaker. These are, in short, that

- The simplest RF applicable in the context is to be applied (so where identity is applicable, it applies)[1]
- "[F]or any function on any demonstratum, you have to be able to distinguish that demonstratum from other things like it for which the same function has different values; if you can't, there is no way to say what value the function yields on this demonstratum."
- Both these conditions must be manifest to both speaker and hearer, and it must be manifest to both that they are manifest to each other, and so on (in the 'Gricean regressus'). (pp 32-33)

The word-selection process could be modelled as follows: starting from an intended referent, the speaker selects a demonstratum and referring function, with the intention that, given the demonstratum, the hearer can easily and unambiguously reconstruct the RF and determine the referent. The hearer, knowing that the speaker will have chosen the demonstratum and RF with the ease of arriving at the referent in mind, can readily test hypotheses about what the referent may be by asking, 'does the demonstratum unambiguously identify the referent, in the context?'.

An RF can, under some circumstances or other, tie a demonstratum to a referent of just about any kind whatsoever.

When the reference is not ostensive but descriptive, then we are concerned with RFs, not from demonstratum to referent, but from meaning to referent. Nunberg provides a battery of examples of everyday language in which the RF, from 'meaning', however specified, to referent, is not identity unless we accept highly polysemous entries.

The thesis adopts the position put forward by Nunberg. There is no principled divide between the linguistic semantic specifications that are properly to be found in the lexical entry for a word and the pragmatic considerations that may generate new types of usages for the word on particular occasions. This may seem fatalistic. If there is no limit to what a word might be used to refer to, what should the lexical entry look like?

[1] Nunberg's ontology includes a finite repertoire of 'basic functions', thus grounding his notion of simplicity.

The answer lies in the facts of how words are used. The extent to which speakers of a language use RFs other than those which are standardly used with a particular word is an entirely empirical question. The lexicographer's task is to list all tolerably common usage-types. If he or she has done that, then usages which do not fit the lexicographer's usage-types will be relatively rare. We might say that the RF from a word to its familiar usage-types is a well-trodden path. A speaker will be more readily understood if he or she gives the hearer directions only along well-trodden paths (except in cases where short-cuts are easy to see). Thus speakers will, most of the time, exploit already-familiar routes. So, alongside the core meaning of the word in the lexical entry, there needs to be a catalogue of the RFs which are found to be used with the word.

Nunberg's arguments first suggested that lexicons listing large numbers of senses for words were doomed to failure as they could not list all usage-types. Indeed, lexical semantics cannot be separated from the pragmatics of reference and the lexicon alone will never list all usage-types, but, directly or via a referring function and an inference mechanism, it does need to represent all those used with significant frequency.

4.3 MacWhinney's Competition Model

MacWhinney (1989) presents a model of lexical acquisition which gives an account of the dynamics of RFs, sense extension, and similar. His 'Competition Model' explores the process of categorisation. It aims to show "how competition provides a way of understanding the semantic ranges of words" and "how words force each other to take on various polysemic and extended meanings" (p 195). The model is a connectionist one. The task it models is one of lexical choice: it is trained to associate different situations in which a choice is to be made, modelled as sets of input features, with different 'word' outputs. In the testing phase, it is fed a set of features and selects a word as output. It is a 'competition' model because the words that are the output options 'compete' to be the output associated with a new input.

MacWhinney contrasts his model with two other models of concepts and categorisation which have been centre-stage in recent times. Firstly, the 'Classical Theory', according to which there are necessary and sufficient conditions for an entity to fall under a concept. To know the concept is to know those conditions in order that entities can be classified as falling within it or not falling within it.

Second, 'Prototype theory'. Psychological evidence, notably that produced in a series of experiments by Rosch and others (Rosch & Lloyd, 1978), provides strong evidence that, however useful as an idealisation the classical model might be, its resemblance to human categorisation processes is limited. Humans, typically, identify a category with reference to prototypical category members or a cluster of features where any one feature need not be present for the entity to be classified under the concept. Whereas the classical theory would expect category membership to be all-or-nothing, prototype theory allows for degrees of prototypicality or centrality to the concept, and will predict that categories have fuzzy edges with some entities not being classifiable as 'in' or 'out'. Prototype theory has been closely associated with the observation that categories often have a 'family resemblance' structure (Wittgenstein, 1953), whereby, of exemplars A, B, C, ..., Z, A and B might have many characteristics in common, as might B and C and all other successive pairs in the chain, but A and Z will have no characteristics in common. Lakoff (1987) presents the case for prototype theory with great gusto.

MacWhinney acknowledges both the importance of classical theory and the criticisms of it from prototype theorists, but notes that prototype theory is purely a descriptive account. Once a concept has been identified as a prototype concept, the various ways in which entities may be more or less prototypical exemplars may be described, but the theory offers no insight into how they are learnt, represented, or used during the classification process.

The core of the problem is that prototype theory fails to define a concept in relation to other

concepts. The 'Competition Model' is so called because alternative concepts are 'in competition' for as yet unclassified objects which might or might not fall under them, so the coverage of a concept is determined by its interaction with other concepts in the neighbourhood. This spatial metaphor is pervasive in MacWhinney's work. In his words:

> The notion of semantic topography is a useful way of understanding the ways in which words compete for meanings. This topography makes distinctions not just between words, but also within words. (p 213)
>
> Major polysemic entries can be seen as corresponding to the valleys of large rivers in this topography, whereas major and minor polysemes correspond to increasingly smaller tributaries. Determining the exact reading for a given word is like tracing a a stream back to its source. Some of the decisions are easy and can be made on the basis of the words in the sentence. Other decisions require rich situational information or prior discourse cues. (p 215)

The metaphor takes usages of words in referring expressions as its starting-point and has limitations elsewhere, but does provide a useful vocabulary for describing the domain. The process MacWhinney models is that of selecting a word that an entity or situation 'falls under': metaphorically, determining for an input grid reference, which valley is most accessible. The word can then be used to denote the entity or situation in the future. The competition model suggests there are areas in semantic space, the valleys, committed to particular words, but there are other areas —high ground, plateaux — where the outcome is not obvious. When a word is needed for an entity in one of these areas, the terrain might be easier either way. But once a route is established, so it becomes the obvious route: the word acquires a basin of attraction in that area.

The metaphor falls rather short in relation to alternations. Where a pathway is established for one word, a parallel one can be invoked for a similar word. As MacWhinney says:

> One of the most remarkable aspects of human language is the way in which words can assume new meanings right in the middle of a discourse ... Sometimes these new meanings are created through processes which operate in fairly well-worn pathways. In other cases, not only the meanings but also the processes deriving the new meanings are more innovative and mark off new territory in the semantic topography. A particularly well-worn extensional pathway produces "metonymy" or the use of the part to refer to the whole.[2] For example, we can use "hands" to refer to sailors, "guns" to refer to soldiers, or "wheels" to refer to an automobile. In these three cases, the pathways have been precut into the semantic topography. However, if we refer to men as "ties" or students as "pencils", we would be using the standard extensional pathway in a very innovative way. (p 223)
>
> The more words using the particular pathway, the stronger the pathway becomes. At the same time, the closer the new word is to the words [already noted as] using the pathway, the better the spread of the extension to the new case. (p 230)

With a metaphor oddly resembling Pustejovsky's 'type coercion' (section 3.7), MacWhinney calls the process whereby a word extends its semantic range 'pushy polysemy'. Pustejovsky's concern is with the interpretation of sentences given a fully developed lexicon for English, whereas MacWhinney is concerned with lexical acquisition, so their focal phenomena are different, but both authors are concerned to show how the interaction of a word with its context can bring about changes to how the word will be interpreted.

The Competition Model is allied to a relatively well-defined account of how words are learnt, represented, and used in processing. The account is a PDP or connectionist one (Rumelhart

[2] This definition of metonymy is not the one adopted in this thesis: cf section 7.5.

& McClelland, 1986). The investigator identifies what seem to be the salient dimensions of the semantic space under consideration, the 'features'. The connectionist network is then 'trained' by presenting it with pairings of a word and a bundle of feature-values. A network can be trained in relation to a number of different words. When the training phase is complete, if the network is presented a bundle of feature-values as input, its output will correspond to the word which gives the 'best match'. Features could, in principle, relate to both linguistic context and non-linguistic settings in which the word was used: they might cover semantic and pragmatic aspects of the situation, as well as syntactic, phonetic and collocational ones. Thus the Competition Model offers an account of how words are learnt, by analogy with how the network is trained. It offers an account of how the word meaning is represented. And the network, given incomplete input which might just be the part of the input associated with semantic and pragmatic features, will give a word as output: this can be interpreted as an account of how a speaker uses the representation to select a lexical item in the process of language generation.

MacWhinney's account is detailed using PDP concepts, but is not in itself tied to any implemented PDP system. Indeed specifying the full set of features relevant to a range of the usage-types for any set of words is a forbidding task. Taraban, McDonald, & MacWhinney (1989) present a working version of the Competition Model, in a more constrained domain. The choices the system makes are not between content words but between the six possible forms of the German definite article. They show their PDP model establishing the range of uses each form has and making some of the same mistakes as a child learning the language.

For this thesis, the Competition Model offers an appealing picture of the categorisation process involved in learning, interpreting, or extending word meanings. It offers an account of the processes involved in identifying and extending the semantic ranges of words. The implementation demonstrates interactions between frequencies and lexical choice, and the abstract model points to how this may apply both at the level of words and at the level of alternations: a word is more likely to be chosen if it has frequently been used in a semantic neighbourhood before, and an RF or alternation is more likely to be used if it has often been used with other words in comparable situations before. It provides a wealth of metaphors for making sense of the question, "What is a word sense?"

4.4 How many senses does a word have?

It should be clear from the discussion above that the question has no absolute answer. There is a limitless range of RFs connecting a word with potential referents. Some are frequently used. Then that word-and-RF pair may be treated as a word sense. But frequencies vary by degrees, and vary according to where you look, and there are no hard and fast lines distinguishing one RF from another, so it will always be a matter of judgement how many senses to list for a word. However lexicographers are obliged to give each word a determinate number of senses. This section investigates a range of attempts to get to grips with the question, with tools drawn from etymology, morphology, syntax and translation, as well as semantics.

4.4.1 Polysemy and homonymy

There is a traditional distinction between polysemy, where one word has several senses, and homonymy, where two different words happen to have the same form. The problem can be stated in terms of sense pairs. For a pair of senses, the question is, do they both map to one word or each to different ones. If we could independently identify homonyms, we could at least say that a homonymous word (i.e. spelt form) received two entries in the dictionary. Can we do that?

Apresjan (1974) and Lyons (1977) both review a range of proposals for distinguishing polysemy and homonymy. For Lyons, homonymy and polysemy are characteristics of lexemes rather than

words. The lexeme is the entity the dictionary entry is about, which, in the case of the verb *drive*, includes *drive, drives, driving, drove* and *driven*. Lyons (p 22) states two preconditions for homonymy. Firstly, only lexemes of the same part of speech can be homonyms, and secondly, to be homonyms, all forms must be shared. Thus forms with the same spelling but different parts of speech are ruled out from being forms of homonymous lexemes. Identical forms having different parts of speech clearly are related in meaning in some cases (*rust* as verb and noun) and not, in others (*can* in "You can do it!" and in "tin can"). The effect of Lyons's stipulation is simply to make any such concerns a separate topic to homonymy and polysemy. The thesis takes a similar line, and does not consider at length issues about the relationship between verbal and nominal *rust* or *can* (the matter falls under derivational morphology) or the disambiguation problems they present to NLP systems (which are largely solved —see section 3.3.3). We do not adopt Lyons's terminology: while 'lexeme' is an important construct for morphology, the distinctions between the different morphological forms of a word are not central to a discussion of polysemy: we talk of 'words' not 'lexemes'.

An etymological criterion is simply that a sense pair corresponds to two different words if at an earlier point in the history of the language the two variants' precursors were different in sound or spelling. Thus *mouth* (of a person) and *mouth* (of a cave) share histories, and are polysemous, whereas *port* (the drink) and *port* (harbour) do not and are not (examples from Lyons). Preliminary problems with the criterion are that the etymology of a word may well be unknown, and it leaves open the question as to how distinct the words must have been in the past. Both versions of *port* derive, ultimately, from Latin *portus* —directly in the case of the 'harbour' sense, and via the Portuguese city of Oporto for the drink. Thus the notion of a distinct etymology is not well-defined. Moreover, there are sense pairs such as the oft-cited *ear* (for hearing) and *ear* (of corn) which happen to have different etymologies, but speakers of English tend to assume are related, so for two words to have distinct etymologies is not a sufficient condition for them to be unrelated in meaning, in the opinions of native speakers. The arguments make the underlying point that synchronic questions about the structure of word meanings cannot be resolved using diachronic criteria. If homonymy and polysemy are to be used to describe aspects of today's language, then the history of the words, while it might be a guide to the distinction, can never provide a criterion for it.

The semantic proposals are concerned to make operational the matter of being 'related in meaning'. A proposal from Smirnickij (quoted in Apresjan (op.cit., p 12)) is that a word is homonymous wherever the meaning distinction is not found for any other words in the language. The premise on which the idea is based is that the human proclivity for extending senses, or using non-identity RFs in Nunberg's terminology, is such that wherever a sense has been extended for one word the same sense extension is bound to have been applied to at least one other word, so on any occasion where the meaning-distinction between a sense pair is unique in the language, we can assume it is because the difference between the meanings does not correspond to any plausible RF or pattern of sense extensions, so the meanings are unrelated. However the argument that something that could happen once, would always have happened at least twice is unsound. Viewed as an operational criterion for distinguishing homonyms, the failing of the proposal is to be found in the difficulty of determining whether the difference between meanings is the same for two sense pairs. The problem of identifying unrelated words is replaced by the markedly more abstract one of identifying distinct meaning-distinctions, and the proposal does not take us forward.

At this point, Apresjan concludes 'polysemy and homonymy are relative concepts' (1974, p 13) and says simply that meanings with a trivial common part are less homonymous than those that do not even have any common part. Lyons takes the matter further, considering how speakers' intuitions can be clear on the topic. *Mouth*, people will say, has a literal meaning ('aperture in the face') which has been extended to cover 'mouths' of rivers and caves. But speakers' assessments run the risk of being closer to amateur etymology than introspection. One man's homonym is

another man's polyseme: one person may believe that two senses have a common etymology, and deduce that they are variations on a theme, to be treated as polysemy, while another does not imagine the meanings or histories are related and treats them as homonyms. Lyons does not rule out the possibility that speakers who consider two senses unrelated in meaning might use those senses differently from speakers who considered them related, and allows that, if that could be shown, it would provide evidence for a theoretical distinction between polysemy and homonymy. But it does not seem likely, and the theoretical status of the distinction is "left unresolved" (p 552).

Where speakers are found to agree that the two meanings for a sense pair are related, then the word can be taken to be polysemous, and where they agree they are unrelated, homonymous, but there are large numbers of sense pairs for which agreement will not be reached, and where the meanings are simply 'not very related'.

4.4.2 Vagueness and ambiguity

The homonymy/polysemy divide has traditionally been one that has concerned lexicographers. Two related distinction which has taxed linguists and philosophers are between vagueness and ambiguity, and between underspecification and ambiguity. Many of the same considerations apply to both distinctions, and 'vagueness' will sometimes be used as a cover term. The distinctions are also important for syntax, though here we shall only consider it in relation to the lexicon. To illustrate: *hand* is nonspecific between *right hand* and *left hand*; a *red* jumper is vague between *brick red* and *pillar box red*. *Bank*, on the other hand, is ambiguous.

The formulation requires some explanation. First, the homonymy/polysemy question is "Do we have one word or two?" We have been given two different senses to consider, and wish to determine whether they are senses of the same or of different words. For the vague/ambiguous distinction, the question is 'do we have one sense or two?' To state the question in terms of sense-pairs would be to prejudge it.

In asking whether a word is vague or ambiguous, we are concerned only with lexical ambiguity. "The man was here" will be ambiguous, to the hearer, if there is more than one man whom the hearer believes the speaker might be referring to. But there is nothing lexical in it.

Distinguishing vagueness and ambiguity has often been important for linguistic theories. Over the years, various methods have been proposed for sharpening and clarifying intuitions about whether a particular sentence is ambiguous or merely vague. The tests, described briefly below, are more exhaustively reviewed in ten Hacken (1990).

Morphology

A first type of test looks to morphology. According to the test, a word is ambiguous if, in the two uses, it has different inflectional or derivational morphology. Thus *hang* is ambiguous between the form with the past tense *hanged*, used for people, and the other, with past tense *hung*, used for anything non-human. Also the verb *act* is ambiguous between the use connected with the theatre, which gives *actor* and *actress*, and the general use which gives *action*. The test could only offer a sufficient condition for ambiguity, never a necessary one, as some clearly ambiguous words such as *bank* simply lack the morphology which could give a clue.

These tests are using as a diagnostic a factor which has only an accidental relation to differences of meaning. *Dream* has two past tense forms, *dreamed* and *dreamt*, so the test would suggest we have evidence for an ambiguity in the word. But the two past tense forms seem to be in free variation and tell us nothing about any ambiguity of *dream*. Even if the two forms were associated with slightly different aspects of meaning, it is not evident why this should lead us to the conclusion that the original word is ambiguous rather than nonspecific. Corpus evidence could reveal that *dreamed* and *dreamt* had slightly different patterns of usage, perhaps associated with

variation in register or dialect, but that would not provide a reason for lexicographers and linguists to say there were two different senses of *dream* even if the dreams described in one register were distinct from those described in the other. The case for seeing the verb as ambiguous would at least need supporting evidence from, for example, other verbs from the same domain —*imagine, fantasise, visualise*—, which may or may not show signs of polysemy corresponding to the putative polysemy of *dream*. While a morphological distinction may be associated with a sense distinction, and so may serve to disambiguate usages, it is neither a necessary nor a sufficient condition for one.

Substitution and translation methods

A common way of describing a word's ambiguity is by giving a paraphrase for it in its first sense, and a paraphrase for it in its second. Might this not be developed into a test? With a multilingual perspective, we might say that if the word gets two different translations in another language, then it is ambiguous.

The approach can develop our intuitions but, again, cannot provide an answer. It will fail to distinguish linguistic and nonlinguistic ambiguity. 'Left hand' and 'right hand' are distinct paraphrases for distinct putative senses of *hand*, but do not convince us of the word's ambiguity. The existence of single words which are synonyms for the target word in one and not the other meaning will not resolve the issue because, firstly, some concepts are lexicalised in a language and others are not, and it would be inappropriate to draw inferences about the target word from facts about the lexicalisation of its possible paraphrases. Secondly, there are occasions where a word has a different synonym in each use yet the intuition that the word is unspecific remains firm. *Child* is not shown to be ambiguous between males and females by the existence of the words *boy* and *girl*. The same arguments apply to the existence of one-word translations.

A psycholinguistic method

There is a substantial body of psycholinguistic evidence indicating that all the lexical entries for a word in the mental lexicon are accessed, and then all but one are rejected (see sections 2.2.4 and 2.2.5 and the account of priming experiments therein). The pattern of access and selection has a characteristic time-course. If a prime such as *bank* is given, it turns out that both *river* and *money* are primed for. In a context which serves to make one and only one of these an appropriate reading, after something between 50 and 200 milliseconds a choice is made and after that only one of *bank* and *money* is primed for. It is tempting to identify lexical ambiguity as belonging to those words for which this 'signature' holds. Experiencing a sentence as zeugmatic would then be a matter of the signature being disrupted because both senses had to be held on to. However, while psycholinguists have identified the 'signature' for disambiguation, there has not been work on identifying a contrasting time course for cases of 'vagueness', so at least for the time being, a psycholinguistic method for distinguishing the two is not available. Section 4.4.3 below describes difficulties for the hypothesis that there are distinct 'signatures' for lexical ambiguity and vagueness, though it remains a matter worthy of empirical investigation.

Quantification

A further test exploits the premise that a single occurrence of an ambiguous word can have only one reading. Quantification can provide a context in which an occurrence of a word is not tied to a particular referent, so there are the logical possibilities that the scope of the quantifier includes the denotata of a target word in just one of its senses or in both of its senses. Wiggins (1971, p 28) introduced the quantification test with

> All banks in the USA are now guaranteed by the Central Reserve Bank.

This is clearly not falsified by the failure of the Central Reserve Bank to guarantee certain river banks. However,

> Yesterday we sold all the red jumpers in the shop.

is falsified if all the pillar-box red, but not all the brick red jumpers were sold.

Against the test stands the question of what we would ordinarily expect a quantified expression to range over. Presumably the banks guaranteed by the Central Reserve Bank in the example do not include the Central Reserve Bank itself, yet that does not effectively falsify the statement once we know that central banks are special kinds of banks which need not be included in the scope of quantification for the statement. Likewise, foreign owned banks might be excluded. The scope of quantification is commonly fixed (inasmuch as it is determinate at all) pragmatically, by the context and the kind of thing being said, rather than by any purely-conceived meaning of the word. Thus the interpretation of quantified statements always involves an assessment of the scope of the quantifier. The widest scope is rarely what is called for. Yet the test assumes the widest scope possible, and that the boundaries of 'the widest scope possible' defines the point at which vagueness turns into ambiguity. A test sentence is no longer simply one in which the target word is in the scope of quantification, as many such sentences are not falsified where it might seem they should be, as in the Central Reserve Bank failing to guarantee itself. The test sentence needs to be one in which the quantification is truly universal, but these items are rarer, and our intuitive assessments of them correspondingly weaker.

Conjunction

The conjunction test works on the premise that a single occurrence of an ambiguous word can only get one reading. It sets up sentences where the single occurrence participates in two sets of relations, as specified in conjoined clauses. The sentence must use the same occurrence of the target word, yet without requiring it to have the same referent in both clauses. Depending on the syntax of the expression to be tested, the effect can be achieved with conjunction alone, or in association with anaphoric *one*, *so did* or ellipsis. The question then is, is the reading where the one clause requires one use of the target word, and the other, the other, acceptable? If it is the case is one of vagueness, not ambiguity. Where the word is ambiguous, the expression has a kind of oddity known as 'zeugma' and often used in witticisms such as:

> ? She came out in spots and a bath chair.

The zeugma is evidence of the ambiguity of *come out*. One using our old familiar *bank* is:

> ? I paid the money in at the bank and she sat down on one and watched the swans swimming along.

Where there is no disambiguating material:

> I saw a mouse in a bank and so did Joan

the 'crossed readings' in which *bank* means 'money bank' in the one clause and 'river bank' in the other, are not possible. Both *bank*s must be of the same variety. Where a word is vague rather than ambiguous, there is no possibility of zeugma where there is disambiguating material:

> My red jumper was pillar-box shade, whereas hers was brick-coloured.

All readings, crossed or otherwise, are possible where there is no disambiguating material:

> My jumper was red, and so was hers.

No pairing of pillar-box red and brick-red jumpers is ruled out. This is the probably the most successful and widely-cited test.

4.4.3 General considerations regarding ambiguity tests

The tests have all been presented with the aid of an unproblematical example of ambiguity and an unproblematical example of vagueness. This was done in order to demonstrate what the test was and what the two contrasting outcomes were. However, in those cases the tests would never be necessary. What we want of a test is that it is consistent with our intuitions, where our intuitions are clear, and that it resolves the question, where our intuitions are unclear. The conjunction and quantification tests fare well in meeting the consistency condition. But do the tests help where intuitions are unclear? There does not seem to be any evidence that they do. Two problems related to syntax, and to the wider question of finding an interpretation.

Firstly, the word in its two uses must have the same syntax. Consider the transitive and intransitive uses of *eat*, as in "John ate the apple" and "John ate". Is this a case of ambiguity or vagueness?

> $*$ Mary ate, and John, the apple.

is unacceptable, but the reason is that elided constituents must have the same syntax and subcategorisation in both their expressed and elided occurrences. It might be desirable to treat all words with alternative subcategorisation (or other syntactic) possibilities as ambiguous. But whether or not that is done, the test still fails to elucidate on the topic of a word's meaning, where the word has different syntax in different uses. The test can only be posed where the two uses are syntactically similar.

The second point is more general and theoretical. A certain amount of interpretation of an utterance must have been undertaken before an acceptability judgement can be made. Three parts of the interpretation process are lexical access, parsing, and 'pragmatic interpretation', the final stage of incorporating the new information into the discourse model. The premise behind acceptability judgements is that a subject can report on the outcome of the first, or second, irrespective of what goes on in the third. For a wide range of syntactic questions, the methodology is widely used and has proved its worth. For the lexicon, for simple cases there is, again, evidence that the premise holds.

Nunberg's arguments (see section 4.2) illustrate the hazards of the premise. Consider

> The newspaper costs 25p and sacked all its staff.

It is anomalous. We cannot place the origin of the anomaly in the lexicon unless we grant the word two lexical entries, one for the copy and one for the owner. Then the size of our lexicon will start to expand, as we list more and more of the possible kinds of referent for the word, and still it will never be complete. So the origin of the anomaly must be the interpretation process. But the anomaly seems similar to the anomaly that occurs with *bank*. In a case lying between *newspaper* and *bank*, how would we know whether the source of the anomaly was the lexicon or the interpretation process? In the general case the point at which the lexical process becomes a general-purpose interpretative one cannot be identified. There is no accessible intermediate representation in which lexical ambiguities are resolved (for acceptable sentences) but in which the content of the sentence has not been interpreted and incorporated into the hearer's knowledge of the remainder of the discourse, or general knowledge.

4.4.4 Ambiguity and vagueness: conclusion

We have encountered clear cases of lexical ambiguity, and have distinguished them from clear cases where lexical lack of specification may give rise to non-linguistic ambiguity. The best of the diagnostics may prove useful for confirming our intuitions about clear cases, but do not serve to resolve the issue in cases where intuitions are unclear.

The conclusion is not surprising. A word has more than one meaning where the process of abstracting the contribution it makes to sentence-meaning produces two clusters of contribution-types. Where there are two entirely distinct clusters, a word is ambiguous, but where it is vague or unspecified, there is a single cluster which spreads out along some dimension. There is no *a priori* reason to expect to find any clear distinction between the two types of cases.

4.5 The ideal lexicographer and the essential word sense

There are no decisive diagnostic tests for identifying word senses, yet there is clearly an important role for word senses in NLP and lexicography. Some words have a number of distinct, conventionalised patterns of use and both human dictionary-users and NLP systems need them specified. Dictionaries provides a huge store of data about them. Lexicographers clearly have intuitions or strategies which enable them to do a fair job of identifying senses. We may ask how the lexicographers go about it.

An idealisation of lexicographic practice provides a working definition of a word sense, as follows. We assume that for each word, the lexicographer

1. gathers a corpus of citations for the word;
2. divides the citations up into clusters, so that, as far as possible, all the members of each cluster have more in common with any other member of that cluster, than with any member of any other cluster;
3. for each cluster, works out what it is that makes its members belong together;
4. takes these conclusions and codes them in the highly constrained language of a dictionary definition.

The process is an idealisation of what actually happens in dictionary-making, displayed to expose 'the central core of the lexicographer's art, the analysis of the [citation] material collected' (Krishnamurthy, 1987, p 75). Now that extensive corpora are available to lexicographers (at least in English and some other languages), lexicography is moving towards the idealisation. It focuses on a process of clustering usages, performed by a lexicographer. The lexicographer was probably not explicitly aware of the criteria according to which he or she clustered at the time, and stage 3 is a fallible *post hoc* attempt to make the criteria explicit. Yet it is those criteria which determine the senses that eventually appear in the dictionary. They are a result of that process. But they are a result at several removes, and with each of these removes comes the possibility of confusion or error.

The idealisation is of use for our search for the nature of word senses. We should like to know what they are, and where one ends and the next begins. 'No entity without identity' runs Quine's test, and without identity conditions for word senses the concept remains hazardously ill-defined. The idealisation points us towards the criteria the lexicographer was using for his or her clustering, because, however quirky they may have been, they are the data that the published form of the dictionary is attempting to communicate. They answer, as well as anything can, the Quinean test. The identity test for a word sense in a particular dictionary is that two usages of the word belong to it if and only if the lexicographer would have put them in the same cluster.

4.6 Sufficient frequency, insufficient predictability

The lexicographer is acting, at this stage of the process, as a well-informed, well-trained subject, making judgements that are the lexical equivalents of a syntactician's introspective grammaticality

judgements. These judgements are a major source of evidence regarding the structure of the lexicon, but what more general account of word senses may they be subsumed under? The answer is the SFIP criterion, introduced in section 1.2.

Every dictionary has a length limit, and yet research on the number of word types found in corpora (Walker & Amsler, 1986) does not indicate any upper bound on the number of words which are candidates for inclusion. Not all will fit, so some criteria are required for selection. An obvious criterion is frequency. Words which are used only very rarely can be omitted from any dictionary but one on the scale of the OED. The same consideration applies to word senses. A word sense must be of sufficient frequency, for it to earn its place in a dictionary.

The previous section indicates another criterion. Where usages fall into a tight cluster, a sense is defined to cover the cluster. Conversely, to be exemplars of a distinct sense, usages must fall outside clusters for already-acknowledged senses. The cluster is, at a first pass, a set of usages which are similar to each other, so the criterion for a distinct sense is that it must represent a cluster of usages which are related to each other but sufficiently dissimilar to any existing senses. In the case of brick red and pillar-box red, there are clearly not two distinct, dissimilar clusters of usages. There are merely two small subsets of the usages of *red* which are indistinguishable from the complete set for *red* (as a colour word) in terms of syntax, collocational possibilities, illocutionary force and other factors, but which denote specific parts of the range of colour which *red* denotes. The only way these subsets can be identified is by specifying their denotation. There is no corroborating evidence from any other dimension of variation that these subsets form clusters. In most cases where the word *red* is used without a specific, observable red object or shade of red being referred to, be it in fiction, discussion of dreams, negated or quantified contexts, questions, or discussions of beliefs and desires, usages are unspecified in terms of the only dimension which makes any difference to the pillar-box red, brick red contrast.

A different example, where two sets of usages seem quite dissimilar yet lexicographers are agreed in granting only a single sense is *newspaper* as copy or corporation. Here the usage-types are distinct in terms of the subject area of discourse and the likely contexts, as well as disjoint in denotata. If dissimilarity was a criterion for granting distinct senses, then distinct senses would be granted here. Yet the lexicographer considers it profligate with valuable space to list a sense for each usage-type because, as he or she reasons, given the 'copy' sense the dictionary user can predict the 'corporation' usage-type. Anyone with the level of familiarity with English necessary to use the dictionary will be familiar with words like *Ford*, *Hoover* and *The Times* referring both to the company and to the product, and also with the kind of discourse —of business, mergers, price rises, redundancies— which call forth the 'corporation' reading. It is predictable and so does not need listing.

A problem with 'similarity' as a criterion is that it invites the image of a scale along which degrees of similarity can be marked off. But there rarely is a scale. There is no scale according to which the uses of *cherry* as fruit, tree, wood and colour can be ranked for similarity. The concept of predictability is, by contrast, grounded in an operational criterion. A usage-type is, at bottom, predictable if language learners or users familiar only with a core sense for the word in question could, on hearing the word in a context demanding some other reading, correctly interpret it and draw appropriate inferences.

Predictability subsumes similarity. Where, as with *red*, 'pillar-box' usages are not in any linguistically relevant way dissimilar to other types of usage, the new type of usage is entirely predictable. That is, a language learner with a knowledge of the use of the word *red*, but who had not come across it in a context where it was specific to a pillar-box hue before, would have no difficulty in interpreting it. Where two usage-types seem in many respects similar, such as between *dog* as a species and as the male of that species, but neither sense is predictable from the other, then, as lexicographers are agreed, both possibilities need specifying. Predictability, not similarity, is the relevant factor. Predictability, as a criterion, also has the advantage that it

makes sense of how a new usage-type can become understood simply because it shares a pattern with some already-known examples. Thus, since *Watergate* and *contragate* are familiar, *Irangate* and *Inkathagate* are comprehensible. The predictability criterion acknowledges the productive power of analogy for generating new usage-types, and indicates why the straightforward products of the process do not need listing in the dictionary.

Predictability is a very broad notion. In philosophy of science, predictive power is a general enough concept to serve as desideratum for theories in all fields of science. In psychological explorations of the nature of concepts, much recent thinking has focussed on concepts as 'theories' which makes it possible for a person to predict what will and what will not fall under the concept (Medin, 1989). As a recent textbook says,

> According to the theory-based view, [the alternative] positions suffer from a misconception that similarity is the basis of conceptual structure. The theory-based viewpoint maintains that similarity is both hard to define, and can be easily overridden (Best, 1992, p 436)

Where a fact is predictable, the theory or generalisation which made the prediction possible can always be made explicit, and for the lexicographer the generalisation might be one that comes under the name of metaphor, metonymy, alternation or analogy. 'Predictability' provides a framework for more specific accounts of what the generalisations are that make certain usage-types predictable for certain classes of words.

The traditional linguistic account of what goes in the lexicon is 'whatever is irregular about the words of the language'. The SFIP criterion fits well with that account, but the roundabout route we have taken to arrive at it has indicated something of the variety of dimensions the irregularity might take. Moreover, the regularities depend on generalisations. Lexicographers writing dictionaries for humans can, to a great extent, take the generalisations as given, as part of the apparatus the user must bring to his or her use of the dictionary. For computer systems, no such assumption is acceptable. The generalisations which underpin the predictability of many usage-types must be made available to the system. Arguments and methods for doing this are presented in chapter 8. For a human dictionary-user, a full range of usage-types is available as a consequence of the listed senses, the user's implicit knowledge of generalisations, and inference. The parallel for an NLP system or lexical knowledge base is that the full range of usage-types will be available, as a result of the explicitly listed senses and generalisations, which together form the axioms of a theory, and the inference mechanisms which can generate appropriate interpretations for usages as required.

4.7 Summary

The chapter has sketched the view of meaning and word senses that underpins the thesis. The basic elements of the ontology are usages –words paired with the situations in which they occur– and the meaning of a word is an abstraction of the similarities between different usages. In one important tradition in the study of semantics, the denotation is a central construct for the formalisation of meaning: here, it was not helpful. Where words are polysemous or are used metonymically, a consideration of the similarities between referents of the word on different occasions of use does not shed light on a word's meaning(s). Where denotation failed us, Nunberg's referring functions proved useful for describing the patterns of behaviour of words. MacWhinney's Competition Model provided an account of how the semantic range of a word may change according to the semantic ranges of other words in the domain.

Traditional distinctions between homonymy and polysemy, and vagueness and ambiguity, were considered. While the distinctions are important as working hypotheses for some areas of linguistics, and for the practical aspects of dictionary-writing, they were both found to be matters of

degree, for which clear answers were often not available. The ontology of usages makes this unsurprising. There is no *a priori* reason to expect usages to fall neatly into mutually exclusive groups. An account of ideal lexicographical practice was presented which worked directly with usages, and in which the decision as to which usages belonged together as exemplars of a single dictionary sense was, at root, a decision of the lexicographer working from trained intuitions. The final section starts to analyse the logic of the lexicographer's intuitions, a process continued in later chapters, and arrives at a general formulation of where a usage-type should be treated as a distinct sense in a paper dictionary, or, as we shall see, as an axiom rather than a theorem in a computational lexicon.

Chapter 5

Dictionary Study I: An Analysis of Word Sense Distinctions

5.1 Introduction

This chapter reports on an empirical study in which the distinctions between word senses were investigated. The range of mechanisms employed by the lexicographers for describing alternative usage-types was examined, and a classification scheme developed. A sample of words was selected, and for each of the words, if it had more than one sense listed in the particular dictionary chosen for the study, then the distinction(s) between those two (or more) senses became part of the population of word sense distinctions. The study succeeded in finding some distinct types of sense distinctions, but also indicated the heterogeneity of the distinctions, with the majority defying any simple classification.

Preliminaries

5.1.1 Choosing a dictionary

The dictionary chosen for the studies was LDOCE. The reasons were, firstly, it is an up-to-date dictionary compiled in the light of current ideas on lexicographic practice and secondly, as a dictionary for people for whom English is not a first language, it is constrained to make many things explicit, where a dictionary for English-first-language users would not need to do so. The first edition is available in a convenient, tidied-up, machine-readable form, with some additional information not available in the printed version, and has already been used for a variety of MRD and NLP projects. The revised and improved 1987 edition was used for these studies. It seems likely that the patterns of distinctions in LDOCE are broadly similar to those in other dictionaries, though comparative studies, particularly making comparisons with larger dictionaries, are a subject for future research.

Arguably, the COBUILD dictionary would have been a more appropriate choice as it was explicitly based on corpus evidence. It would certainly have been a suitable dictionary for the studies but I do not think it likely the results would have been markedly different. Although it makes less of the fact, LDOCE also uses a citation corpus. Any dictionary needs to choose a set of senses for a word and to express clearly what makes those senses different, and this task is central to the quality of the product. The important matter for this study was that these jobs were done well, rather than how they were done (see also section 4.5). Fillmore (1989) compares COBUILD and LDOCE and recommends both, thus indicating that either would have been suitable for this

study.

5.1.2 Limiting the domain

To constrain and focus the topic, several kinds of meaning distinction will be excluded from the population of distinctions under scrutiny (inevitably, the boundaries will sometimes be hard to place). These are:

1. Part-of-speech ambiguity

   Where there are word senses falling into different parts-of-speech for the same spelt form, LDOCE gives them separate entries[1]. Our strategy shall be the same. While part-of-speech ambiguity is a major issue for NLP, other authors have addressed it. Here we shall be concerned only with distinctions in word sense within one part-of-speech. We take as distinct parts-of-speech the major word classes, as identified in LDOCE; *n, v, adj, adv, conj, prep, det,* and *pron.*

2. Morphological variants

   Distinctions between words which are morphological variants of each other, such as 'socialist' and 'socialism', will not be addressed.

3. Set Phrases

   Many senses of words are identified in the dictionary as occurring only in idioms, compounds or collocations. They present difficulties to lexicographers in terms of whether, and where, the multi-word unit should appear in the dictionary, and where to list a particular usage-type for a word as a multi-word unit, and where to list it as a distinct sense, where that sense only appears in a specific lexical context. Issues concerning set phrases in the lexicon were considered in this study only where they could not be separated from the tasks of specifying sense distinctions and allocating usages to senses.

4. Semantically close words

   There are clearly many similarities and distinctions to be drawn between synonyms or near-synonyms, antonyms, words in the same contrast set and so forth. They will not be considered here.

5.1.3 Presentation of alternative word senses in LDOCE

LDOCE has a variety of ways of representing alternative usage-types. Eight methods are identified below. With methods 1–3 the distinction in usage-types is explicitly a sense distinction. With 4 it is marginal, and with methods 5–8 the two usage-types are not being treated as distinct senses, but as usage-types where the distinction between them is predictable, or vague, or one or other of the senses is rare, as discussed in the last chapter. Quotes are from the LDOCE User Guide or 'Front Matter'.

1. Different entries.

   "... bank(1) and bank(3) are treated separately, even though they are both nouns, because there is no historical connection between the two words and their meanings are completely different." (page F17)

[1] Unless there is no meaning difference, as between nominal and adjectival readings for colour words, and then a bold type subentry is given

2. Numbered senses within an entry.

   This is the basic level at which meanings are explicitly distinguished.

3. Subdivisions of numbered entries marked **a**, **b**, etc.

   The rationale for using subdivisions rather than main divisions is not clear. A necessary but not sufficient condition for their use is that the senses distinguished in this way are more closely related to each other than they are to other senses.

4. Bracketed optional part.

   One sense is given by including the bracketed material, and another by excluding it.

   Examples:

   marquetry (the art of making) a type of pattern in wood . . . .

   martini (a glass of) an alcoholic drink . . .

   mazurka (a piece of quick lively music for) a Polish dance.

5. (fig.) in front of an example.

   According to the User's Guide,

   > Some words are used in an imaginative or "figurative" way, to suggest a meaning that is not the literal meaning but has some similarities with it. If a word is often used like this, the examples will include a figurative use, and this is shown by the note (fig.) (page F36)

   Example:

   materialize 1 (fig.) *I'd arranged to meet him at seven, but he never materialized.*

   There will always be a continuum between figurative uses of a word and distinct senses which have their origins in figurative usage, since at any point in time many word usage patterns will be in flux between extremes of originality and conventionality. Where a lexicographer draws the line between the two types of cases will always be somewhat arbitrary.

6. The main definition contains a disjunction.

   Examples:

   masked 2 by or for people wearing masks: *a masked ball*

   melody 1 a song or tune: *a haunting melody*

   A disjunction of this kind is not necessarily signaled by 'or'. It could be signaled by 'and', or by a comma. The semi-colon is used for the similar and occasionally overlapping purpose of providing an alternative statement of the definition, often a synonym or near-synonym.

7. Disjunction in the grammatical coding

   Where a verb may be countable or non-count, or a verb transitive or intransitive, or otherwise occur as a member of more than one kind of minor word-class, a semi-colon in the grammatical code for the word or word sense will indicate the disjunction.

   Example: **marriage** [C;U]

The square-bracketed code means the word can be a count or a non-count noun. The non-count, 'marriage is a good institution' reading of marriage might be taken as a distinct sense. The alternatives indicated by the disjunctive grammar code are often both illustrated in the examples.

8. Brackets with a disjunction in them.

   As well as indicating an optional extra meaning, brackets sometimes give selection restriction-like information on how a word is normally used, indicating what sort of an entity a noun or adjective is used to describe, or for a verb, what its subject or object is likely to be. When these brackets contain a disjunction of dissimilar kinds of entity, it could be said that two senses are being conveyed.

   Example: **meek** (of people or behaviour) ...

   The use of 'or' in brackets of this type is often an indicator of a range of possibilities rather than an indicator of disjunction, and the only disjunction it is regularly used to convey is the 'of people or behaviour' one used here.

5.2 Pilot study

The purpose of the study was to establish what, if any, were the commonly occurring patterns of distinctions between word senses. Approximately 1% of the dictionary, twelve of LDOCE's 1227 pages, were examined. The study was manual.

The sample contained 427 full entries. An initial analysis is given below.

Full entries	427	
of which - capitalised		20
- compound/phrasal/hyphenated		37
leaving a base population of	370	
2 separate entries in same word class	6	
Entries divided into numbered meanings[1]	110	
of which - 2 meanings		75
- 3 meanings		17
- 4 meanings		10
- 5 or more meanings		8
nouns		63
verbs		21
adjectives		25
adverbs		1
Total of numbered word sense distinctions[2]	187	
Numbered entries subdivided using letters	10	
Brackets give another sense (as in 4 above)	23	
of which - 'cause to'		4
- 'too'		3
Another word sense given as "fig" example	10	

Notes:

[1] Excluding numbered meanings where a phrase, idiom or collocation is given, also ignoring subdivisions of numbered senses by **a, b** etc.

[2] On the assumption that, where there are more than two senses, the number of sense distinctions worth considering is the minimum possible, i.e., one less than the number of senses.

5.2.1 Clarifying dictionary definitions

There was a further task to be performed before a distinction could be classified. The entries had to be read closely and interpreted, in order to establish the exact nature of the distinction the lexicographer had in mind in deciding to identify two senses. In terms of the idealised lexicographic practice of the last chapter, the process of coding the distinctions into the highly constrained language of the the dictionary had to be undone.

An LDOCE definition can be a very complex structure. In this exercise the distinctions under consideration were those between numbered senses. A numbered sense definition can have all or any of the following.

1. syntactic code
2. prepositions subcategorised for
3. definition-text or 'd-text' with
   (a) syntactic structure
   (b) semantic content
   (c) restriction on range of application
   (d) 'espec.' specifying a central or prototypical usage
4. examples with
   (a) syntactic structure
   (b) semantic content
   (c) collocations (sometimes bold typeface).

In interpreting sense definitions, many questions arise. Three were particularly salient to the current exercise: necessity, consistency, and centrality.

Firstly, of 1, 2, 3c, and 4a-c, it can sensibly be asked whether the specified feature is a necessary condition of the word being used in this sense, or merely typical.

Secondly, it can also be asked whether the information from the different aspects of the definitions is consistent. Where a word has a sense *a* and a sense *b* it could be that examples and syntax suggest one dividing line between the usages to be classified as *a*s and *b*s, while the definition-text suggests another. An example: sense 2 of *application* is marked as non-count yet the d-text is 'the act of putting something to use', and acts are generally countable. In this case the potential inconsistency can be spotted when the definition is examined in isolation. The more common case is where the definition shows no inconsistency in isolation, but indicates that certain features, when taken together, identify a particular sense of the word, and another group of features indicate another sense; but then, on examination of citations, it becomes evident that many of the citations display some features from the one set and some from the other. The clustering of features indicated in the dictionary does not tally with the evidence for how the word is used. The dictionary distinction is thus inconsistent with the citations.

Thirdly, we must ask, where a part of a sense definition does not present a necessary condition for the word being used in that sense, how much variation away from the form given in the dictionary is acceptable with the word still being properly classified as that sense? This problem is particularly clear in relation to examples. An example sentence, like a picture, suffers a

shortcoming: it provides an instance, but does not provide any indication of how much variation away from the example presented is permissible, without the variant ceasing to exemplify what the example exemplifies. This was a pervasive problem. An example is appointment/2.

> **appointment** ... **1** an arrangement for a meeting at an agreed time and place ... **2** [U] the agreement of a time and place for meeting: *He will only see you by appointment.*

It is not clear whether any usages of the word except those in the *by appointment* collocation are to be classified under sense 2. The intended distinction between senses 1 and 2 is clearly quite subtle. The easiest route to follow would be to equate sense 2 with *by appointment* exclusively, but if that had been what the lexicographer had intended, then he or she would, in principle, have made the fact explicit by putting the collocation in boldface at the head of the sense entry.

These questions often could not be resolved in isolation. The experimenter identifies a sense by contrasting its specifications with the specifications of the other senses. In most cases the contrasts will start to answer the questions, but it will take the evidence of citations to remind the experimenter of the distinctions the lexicographer was attempting to capture, and to determine where the lexicographer's bundles of features seem not to belong together.

5.2.2 Classification scheme

Distinctions have been classified according to the following categories:

Must-Be-Theres (MBTs)

MBT distinctions are those where, if the situation is such that the one sense can be applied, then it is a logical consequence that the other can also be applied, though to some other aspect of the same reality. If the situation is thought of as described by a schema with slots and fillers, an MBT distinction is one which attends to different slots of the same schema. The presence or existence of the filler of the one slot entails the presence or existence of the filler for the other.

Examples:

> **marriage** *n* **1** the union ... **2** the state of being married.
>
> **matricide** *n* **1** the murder of one's mother **2** a person guilty of this crime
>
> **marry** *v* **1** to take (a person) in marriage ... **2** to perform the ceremony ...
>
> **married** *adj* **1** having a husband or wife ... **2** having as a husband/wife

MBTs are the kind of distinction which have received most academic attention. The emphasis on different slots of a frame is often linguistically realised as different patterns of subcategorisation for the target word. These have been widely explored, particularly in relation to verbs, and receive a detailed treatment in Chapter 9.

Many of the patterns of 'predictable' usage-types that, by virtue of their predictability, do not get listed in the dictionary, would be MBTs if they were. The type/token alternation, the distinction between information-bearing objects (*book, film, tape*) as physical and as information-items, and the 'diathesis' alternations for verbs of sound as analysed by Levin (1991) are all examples.

Generalising metaphors (Genmets)

The distinction is between one specific sense, which is unequivocally the right word for the specific situation, and a less specific sense which shares some features of the specific sense but can be applied in a wider range of situations. The less specific sense has in several cases acquired positive

or negative connotations, while the more specific one has more of a factual flavour. Amongst the verbs, a bracketed note stating the type of entity that the more specific sense applied to was common.

Examples:

martyr *n* **1** someone who is put to death or suffers for their beliefs ... **2** someone who gives up their own wishes or suffers ... **3** someone who suffers something they cannot avoid

maul *v* **1** (esp. of animals) to hurt badly ... **2** to handle roughly ...

meaty *adj* **1** full of meat **2** full of valuable ideas.

Domain shift (Domshift)

It is in the nature of words that, in learning and using them, we apply them to situations beyond where we have previously applied them. The point at which two situations in which the word is applied are sufficiently far removed from each other to constitute distinct senses will never be clear cut. A domshift distinction is one where the two situations of usage are sufficiently far apart so that the lexicographers have decided two distinct senses are called for, even though one might say, 'but the word has only been adapted as far as it had to be, given the different entity or situation to be described'.

Examples:

mastery *n* **1** full power to control or defeat something **2** great skill or knowledge in a particular subject or activity

marshal *v* **1** to arrange (esp. facts) in good or effective order **2** to lead or show (a person) ceremonially or carefully to the correct place

mellow *adj* **1** (of fruit and wine) sweet and ripe ... **2** (of a colour) soft ... **3** (of people and behaviour) wise and gentle ... **4** (feeling) pleasantly calm and friendly ...

Natural and social kinds (Kinds)

The distinctions classified under this heading are those where, owing to a non-linguistic fact, the entities or situations identified by the different word senses have distinct denotata, and although the denotata have many attributes in common, they will always remain different classes of things. Membership of both classes is possible, in the social if not the biological cases, but is then coincidental.

The sample contained only nouns in this category.

Examples:

marrow 1 also **bone marrow** ... **2** also **vegetable marrow** ...

marshal 1 an officer of the highest rank in certain armies and airforces ... **2** an official in charge of making arrangements for an important public or royal ceremony or event **3** an official in charge of making arrangements for a race ... **4** (in the US) **a** an official who carries out the judgements in a court of law ... **b** a chief officer of a police or fire-fighting force

For many of these cases, a dictionary definition is not a definition in the legal or scientific sense, as it does not determine the membership of the class defined, but rather a pointer to a class already set apart by nonlinguistic means. A complex set of facts, some of which may not be known to science, serve to define the class, and no more than a reference to these facts will be found in the dictionary. Of the basic distinction-types identified, this was the one that best fitted the Bank Model, or the conception of polysemy in which it is a variant of homonymy.

5.2.3 Results and interpretation

(From analysis of distinctions between numbered word senses only)

	Nouns	Verbs	Adjectives	Totals
Genmets	9	6	3	18
MBTs	14	6	2	22
Domshifts	8	13	7	28
Kinds	11	0	0	11

The classification scheme is neither exclusive, nor exhaustive, nor final. Firstly there will be cases where more than one of the basic distinction-types apply, so some distinctions fall into more than one classification.[2] Just 79 of the 187 distinctions received a classification. Secondly, the dictionary will always provide us with distinction-types not previously encountered, so no list will ever be exhaustive. Thirdly, the scheme presented here is based on a small sample. The evidence is accumulating, but this is certainly not yet a final product.

The fact that the majority of the distinctions defied classification (and for so many that were classified, it was far from obvious that they fitted one and only one classification) is itself interesting. It speaks for the heterogeneity of sense distinctions. For different words, quite different distinctions were salient. Generalisations were hard to come by. The observation gives substance to the position adopted for theoretical reasons in the previous chapter that lexicographers may consider an indefinite number of dimensions of variation in clustering usages into senses.

Another part of the explanation is that those distinctions between usage-types which would have been readily classifiable, because the distinction was simple, were generally not made in the dictionary, because they were predictable. Evidence of the usage-type distinction is sometimes available from methods 5–8 of section 5.1.3. The entry for *meek*

meek (of people or behaviour) ...

shows a usage-type distinction which, had it been a sense distinction, would have been a domshift. However the pattern applies not just to *meek* but to a range of words for dispositions and behaviours. By analogy with similar words, the alternation pattern is predictable for *meek*. As argued in the previous chapter, predictable senses are not listed (unless they occur with particularly high frequency). In this case the usage-type is sufficiently predictable to get some mention, but insufficiently predictable and frequent to require a distinct, numbered sense. In looking at the sense distinctions in the dictionary, rather than the usage-type distinctions in the language at large, we are tending to examine a particularly irregular subset of all usage-type distinctions.

The lexicographer's choice of method for conveying an alternative usage-type is closely related to its predictability. To idealise: where a usage-type is entirely predictable, as with the type/token distinction, nothing need be said. Where it is virtually predictable, very little need be said. All that need be done is to give the user a 'tick' to confirm that an alternate usage, which the user might well expect to find, is indeed to be found, with a more-than-negligible frequency. To simply insert a tick in the entry would be too minimal, but methods 5-8 of conveying alternate usage-types can be seen as techniques for the job which are as minimal and space-saving as is compatible with user-friendliness. Method 4, 'bracketed optional part', is one step more expressive and verbose. The column inches of a distinct new numbered sense are reserved for those usage-types where the distinction between usage-types is rather less straightforward, less predictable, and takes more of

[2] In principle, these distinctions would feature more than once in the table. However the first object of the study is to arrive at a classification scheme, and any double-classification would tend to complicate that task. So distinctions which potentially fell into two or more categories, or were otherwise complex, were simply not classified in the pilot study.

the expressive power of natural language to be explained. Part of the rationale for having such a range of mechanisms for indicating alternative usage-types is, then, to implicitly indicate the degree of regularity of a usage-type, and to avoid wasting space on the more predictable variants.

5.3 Summary and conclusion

The study examined the distinctions between word senses in a learners' dictionary. The mechanisms employed by the lexicographers for describing alternative usage-types were examined, and it was found that different techniques were used for different kinds of distinctions. Where it was predictable that an extended usage-type would be available for the word, given a basic sense, the extended usage-type was given a concise treatment. Unpredictable usage-types needed spelling out more fully.

A classification scheme for word sense distinctions was developed. A sample of words was selected, and for each of the words, if it had more than one sense listed in the particular dictionary chosen for the study, then the distinction(s) between those two (or more) senses became part of the population of word sense distinctions. The classification of these distinctions demanded a close reading of the dictionary entry. There were three particular difficulties in interpreting entries. First, it was often not clear what defined the word sense, so was a necessary condition of a word being used in that sense, and what was merely associated with the sense. Second, some parts of the sense entry were sometimes inconsistent with other parts, or with corpus evidence. Thirdly, it was often not clear from example sentences how close the fit had to be between a citation and the example sentence, for the example sentence to provide evidence that a citation was of a given sense.

The study succeeded in finding some distinct types of sense distinctions, with some potential for use in future approaches to Wilks's problem. In its original goal, it was only moderately successful. However, a number of other points emerged. Firstly, for a fuller understanding of the dictionary distinctions, the dictionary needs to be read alongside a set of examples of the word in use. Secondly, the dictionary's choice of how to present senses carries information, particularly about how predictable a non-base sense is from a base one. Thirdly, while a taxonomy of distinction-types covering large numbers of distinctions is hard to come by, there are plenty of small clusters of distinction-types where a group of words of similar meaning all exhibit exactly the same distinction or 'alternation'. Fourthly, between polysemy and various other phenomena there is a gradient rather than a divide. Where a verb had a range of subcategorisation frames, was this polysemy or purely syntactic variation? Where an example was marked as figurative, was this polysemy or metaphor? These were puzzles which the lexicographer had responded to on a case-by-case basis, but which, taken as theoretical questions, set the agenda for much of this thesis.

Chapter 6

Dictionary Study II: Do Dictionary Senses Match Corpus Usages?

6.1 Introduction

This chapter describes an empirical study investigating how widely the Bank Model holds. A sample of words was selected using an effectively random procedure. For each, a set of citations was gathered from the LOB corpus. For each citation, an assessment was made as to whether it fitted one, none, or several of the senses given for the word in a desk dictionary. It was established that, for most words, some of their usages could not be classified into one and only one of the senses the dictionary gave. Usages were often indeterminate between senses, and senses were often insufficiently clearly identified for it to be possible to classify with any confidence. The study indicates that the Bank Model is limited in its range of relevance, and hence that NLP needs alternative models of how the usages of a word relate to the kinds of senses a dictionary provides for it.

For the purposes of the study, a usage is specified by a corpus citation of about eighty characters with the word under scrutiny in the middle. Thus the only available 'context' for a usage is verbal and is given in the (on average) seven preceding and seven following words. This notion of 'usage' is clearly a very limited one.

6.2 Identifying the sample

The dictionary used was, once again, LDOCE. The source of usages was the LOB corpus. The sample of words to be investigated was arrived at in the following way. Very common words were excluded because they tend to have very large numbers of senses and to present complex and difficult cases. For this study, simpler cases were to be examined. Low frequency words were excluded because it would not be possible to see any patterns emerge unless there were a reasonable number of usages to be examined. A range meeting these constraints was 26-29. So the initial sample was chosen by taking all those words which had between 26 and 29 occurrences in the first half of the LOB corpus. Half of these, a sample of 154, were taken for further analysis.[1]

[1] The reason half the corpus, and half the sample, were taken was so that if a model were to be built on the basis of the study of the sample in relation to the usages in the first half of the corpus, then the untouched halves would provide an environment for testing the model.

From this set the following were removed; prepositions and adverbs (there would only have been two of these, not enough to make any general comments), proper names, adjectives relating to countries ('Dutch', 'Greek'), titles ('Earl', 'Congress'), and non-base forms of words ('cutting', 'created', 'directors') or forms which were base but where a non-base form occurs much more frequently than the base form. The size of the filtered sample was 83.

The rationale of the last points was that dictionaries generally list only base forms, so it was appropriate to look at all the inflectional morphological variants related to the same base form together. For the same reason, where the base form was in the initial sample the morphological variants were added; since 'chapel' was in the sample, 'chapels' was added. Since 'apply' was, 'applies', 'applied' and 'applying' were. The 83 words, complete with morphological variants, are listed in Appendix A.

The exercise was, in outline, to determine the following. For each word, for each usage of that word, which, if any, of the word's senses did it fit? Where each of the usages of a word could be straightforwardly matched against one and only one sense, the limited evidence allowed that the word fitted the Bank Model. Where they did not, the word fell outside the scope of the Bank Model and an NLP system committed to it would on occasion be unable to represent a correct understanding of the word. It was also hoped that the misfit cases might be revealing of the sorts of difficulties that lay in store for any replacement model.

Semantic rather than syntactic distinctions are our concern. For those words that were used as more than one part of speech, the nominal, adjectival and verbal uses were treated as separate sets, and the problem was to classify the nominal usages according to the nominal word senses, the verbal usages according to the verbal senses, and the adjectival usages according to the adjectival senses. Cases where the usage could not be readily classified as nominal, verbal or adjectival were not common.

The exercise had two stages: clarifying the dictionary definitions, as discussed in the previous chapter, and determining which of the senses a usage fits. Here there were several possibilities:

1. Exactly one sense fitted.

2. More than one sense was near the meaning in the citation. In these cases the definitions were examined for clues as to how the lexicographer would have intended the usage to be classified; this generally involved a close examination of how well the syntax and meaning of the citation matched that of the dictionary examples for each candidate sense. Sometimes a unique sense was established as a 'best fit', sometimes not.

3. Two or more non-exclusive senses applied, both making different contributions to the word's contribution to the citation's meaning.

4. A usage was indeterminate between different senses. For example, the usage of 'guest' in

   > the opportunity of showing a guest something of ourselves ...

   is indeterminate between guest/1, 'a person who stays in someone's home ...' and guest/2, 'a person who is invited out and paid for ...'.

5. A usage was not covered by any of the senses, perhaps because it was an unusual figurative use of the word, or a rare use, or simply because the lexicographer had left something out.

6. It seemed that the word was being used in one and only one sense but there was insufficient context to determine which.

6.3 Two worked examples

There follow two accounts of how usages were found not to fit one and only one of the senses. First:

> **image***n*
> **1** [C (**of**)] a picture formed in the mind: *She had a clear image of how she would look in twenty years time.*
> **2** [C] a picture formed of an object in front of a mirror or LENS, such as the picture formed on the film inside a camera or one's REFLECTION in a mirror
> **3** [C] the general opinion about a person, organization, etc., that has been formed or intentionally created in people's minds: *The government will have to improve its image if it wants to win the next general election.* | *The company tries to project an image of being innovative and progressive.*
> **4** [(*the*) S (**of**)] a copy: *He's the (very) image of his father.*
> **5** [*the*+S+**of**] a phrase giving an idea of something in a poetical form, esp. a METAPHOR or SIMILE
> **6** *old use* likeness; form: *According to the bible, man was made in the image of God.*
> —*see also* MIRROR IMAGE, SPITTING IMAGE

The citations included the following.

> of the Garonne, which becomes an unforgettable image. This is a very individual film, mannered,

Here the 'image' is an image/1, a picture in someone's mind (probably the author's; possibly the whole cinema-audience's). It is also the image/2 produced on a screen by means of projection equipment from the image/2 on the photographic film.

Regarding image/3, if we look only at the d-text we may think that the 'image' in the text is the opinion that film-going people will have of the Garonne (etc.) which has been formed intentionally by the film-maker. However the examples indicate that the lexicographers probably did not have this sort of thing in mind for image/3. On looking further at the d-text we shall conclude that the category of 'person, organization etc.' alluded to as what the image/3 might be of, is a category that probably does not stretch as far as whatever it was about the Garonne that the image was of. (If the film was made by the Garonne tourist board, then it is likely that more context would have supported our leaning towards image/3.) Image/4 does not seem a likely candidate. We do not have the 'the image of' expression which is typical of this sense[2], nor does there seem to be any copying, in any very obvious sense, going on.

Sense 5 would seem to fit except that, in this case, the 'the image of' expression is required[3] so sense 5, which from d-text alone might seem the best fit, is ruled out. The d-text for image/6, 'likeness; form', is so open as to seem quite plausible as a match for the text (in which the word 'image' could be exchanged for the word 'form' without dramatically changing the meaning), and the example sentence does nothing to discourage this interpretation. It is only the style label 'old usage', which might deter us from allowing that the text uses image/6, since there seems nothing old-fashioned about this citation.

Senses 4, 5 and 6 can then be rejected (though sense 5, only unwillingly; the d-text did seem the most natural match for our citation, but the grammar code disallowed it. We may wonder if we have here an inconsistency, as described in section 4.2 above. The dictionary indicates that a usage which matches the d-text of image/5 will only be found in the expression, 'the image of',

[2] As conveyed by the LDOCE grammar code [(*the*)S(**of**)].
[3] As conveyed by the LDOCE grammar code [*the*+S+**of**].

yet this citation seems to cut across that supposed correlation[4].) Sense 3 can be set to one side on the basis that we have no evidence that the film was made by the Garonne tourist board (or similar). But we would rather not have to make a choice between senses 1 and 2. The usage makes reference to both the projected image/2 and the images/1 that the projected images/2 caused in people's minds, and to make a choice would be to reject half the story.

Second:

> **exercise** *n*
> **1** [C;U] (a) use of any part of the body or mind so as to strengthen or improve it: *if you don't take/get more exercise you'll get fat.* | *She does exercises to strengthen her voice.*
> **2** [C] a question or set of questions to be answered by a student for practice: *Look at Exercise 17 in your book.*
> **3** [C] a set of actions carried out by soldiers, naval ships etc., in time of peace to practise fighting: *The soldiers are here for a NATO exercise.*
> **4** [S(**in**)] any set of actions, esp. when expected to have a particular effect: *Getting this report done in such a short time was quite a difficult exercise.* | *After the President's embarrassing remark, his staff had to stage an exercise in damage limitation* (=try to limit the damage he had done)
> **5** [S;U(**of**)]*fml* the use of a (stated) power or right: *Expelling him from the club was a legitimate exercise of the committee's authority.*

A usage which defies classification as one and only one sense is

> but at best only portions of the exercise can be significant artistic expression —

Syntactic clues offer no assistance. All senses admit a singular form with article 'the' and nothing subcategorised for. Sense 1 will be appropriate provided that a goal of the exercise was to strengthen and improve the faculty used in undertaking it, but in this citation no such goal is mentioned, nor would it necessarily be even if we had access to much fuller context. Exercise/2 could well apply, since the exercise may very well have been one set for a student for practice, for example in an art class. Setting aside the possibility that the form of art under consideration is the symmetry and elegance of a military attack, exercise/3 can be rejected. Sense 4 is so broad that the usage certainly could be said to fit, but it could be that this very general sense should be set aside on the basis that there is a more specific sense available. Sense 5 can be discarded with some confidence on the grounds that 'artistic expression' and 'the exercise of a power or right' are unlikely bedfellows, that there is no 'stated' power or right, and that there is no 'of' following 'exercise'.

So in this case we have a hierarchy of specificity. If sense 2 is appropriate then sense 1 could be set aside, although in a sense it did apply, since its applicability was already implicit in the more specific sense 2. A larger context might and might not resolve whether sense 2 applies. Likewise, if sense 1 applies, then sense 4 need not be invoked since it is already implicit in sense 1 —but we do not know, on the basis of the citation, whether sense 1 does apply.

The examples should not be read as a criticism of the dictionary. Dictionaries are not written with a view to the task of assigning usages of a word to one and only one of the senses for the word

[4] We may also be tempted to ask the deeper question, what grounds does the lexicographer ever have for claiming that a sense is always used in a particular syntactic context? Where a comparable claim is made about a word (in all its senses) the lexicographer's evidence would be that all the corpus citations for that word were in that syntactic context. But to make the equivalent claim about one of a number of senses for a word, the lexicographer would have to present as evidence a corpus with all occurrences of the word under dispute sense-disambiguated. Yet such corpora do not exist and, as the example demonstrates, for a word such as 'image' it is highly implausible that one ever could. So there is some reason to suspect that 'obligatory' grammar codes on word senses might often be better interpretated as 'expected' grammar codes.

that the dictionary provides. It is not a problem for a person, be they a native or a non-native speaker of the language, if more than one dictionary sense is able to resolve their uncertainty about what a word is contributing to the meaning of a sentence where it puzzles them. A user need only read an entry up to the point where their puzzlement is resolved[5]. Even if a word sense coming later in the entry would have fitted the usage equally well or better, such a user has no need of that further information. Hence it is not surprising, and no criticism of lexicographers, if usages can often not be assigned one and only one word sense. There is no practical or theoretical reason why they should be.

6.4 Results

Of a sample size of 83 words, 14 had just one sense for each part of speech they featured in, leaving 69 for which there was potentially a choice to be made. For 60 of these 69 words, there was at least one usage which could not with any confidence be classified into one rather than an other of the senses. Thus the sense selection task presented in the experiment could sometimes not be resolved to a single sense for 87% of words where the possibility arose.[6] Appendix A. contains specifications of which words fell in which categories.

Sceptics and advocates of the Bank Model may argue that another researcher would have succeeded in classifying all or nearly all usages. Appendix B presents definitions and concordance lines which, in this experimenter's opinion, could not be satisfactorily classified as one and only one sense. The challenge for the sceptic is then to identify how each of the usages in Appendix B should be treated.

The usages which could not be assigned one and only one sense could have been classified according to the reasons they defied classification. This was not done in this case because the range of explanations was arrived at only in the course of conducting the experiment. For a further experiment it would be a worthwhile exercise.

6.5 Observations

The first point to make is that the exercise was, much of the time, hard. In the Bank Model people select senses instantly and effortlessly. For the sample of words chosen here, the experimenter was frequently toiling laboriously.

The task was hard in the cases where more than one of the dictionary senses was near the usage in the citation. The dictionary provides only a set of clues to the nature of the senses that the lexicographer was intending to discriminate. Identifying the divisions that the lexicographer saw in the conceptual space of usages of a word is a matter of reconstruction, and the citations are essential to the task, so as experimenter I was working at clarifying the sense distinctions throughout the process of classifying usages. It was not possible to work with an unchanging conception of the distinctions. Each time a new citation neither clearly fitted one and only one sense, nor replicated a pattern already seen, a re-evaluation of the sense distinctions for the word was inevitable. Examples of awkward words are 'apply', 'image', and 'design'. Each has a collection of closely related senses many of which are fairly abstract and all of which are metaphorical or metonymic variations of one another.

[5] Where the user is using the dictionary for some purpose other than looking up words to determine their meaning, as when a non-native speaker uses it for generation, rather different considerations apply.

[6] No statistics are presented in relation to numbers of usages which could not be resolved because, firstly, the sample sizes were too small, and secondly, the question addressed is, for each word, a qualitative one; does the evidence permit, or does it not permit, that the word fits the Bank Model. Moreover statistics based on usages are highly susceptible to distortion, owing to the great contrasts between frequencies of different words, senses and usage-types.

The issues encountered varied greatly from word to word. Words where a sense could always be selected, and selected easily, included 'absence', where sense 1 is a specified person's absence, while sense 2 is the non-existence or lack of something, and 'capable', where all but one usage were in the 'capable of' construction, which is specified as capable/1.

'Football' and 'chapel' were words which could often not be classified owing to lack of context. The distinctions under consideration are between soccer and American football, and chapels within churches, stand-alone chapels, and chapels which are rooms for worship within houses or institutions. It was in these kinds of cases that non-linguistic knowledge played the clearest role. Knowing the corpus is of British English, and that there is an English soccer team called Newcastle, I am confident (though not certain) that in the citation containing 'Newcastle' soccer is the game in question (see 'chapel' in Appendix B for more examples).

A very common pattern is for one sense to be more general and another more particular, as discussed in relation to 'exercise' above. Copestake & Briscoe (1991) have developed a formalism in which different senses of a word may inherit aspects of meaning from other senses of the same word, but always with the possibility that there is some information attached to the 'inheriting' sense which overrides parts of what would otherwise be inherited. For the kinds of sense distinction where it is applicable, their formalism sets the Bank Model aside, since it builds on the idea that there will be large amounts of overlap between senses. It permits a clear expression of how a usage may fit a more specific sense, and thus, implicitly, carry some of the characteristics of a more general one.

In general, dictionary entries displayed an inverse correlation between the specificity of the d-text and the number of examples. Those senses like exercise/4 with d-text which only very loosely constrains where the sense might be suitable have examples to clarify where the sense is used. In contrast the American sense of football neither needs nor has any examples. The contrast is closely related to Halliday's distinction between more lexical and more syntactic items (Halliday, 1966). In general the items with more information in the d-text and fewer examples were easier to work with, and ones that required the incorporation of evidence from a number of examples and from syntax codes were harder. This was not a result of a lexicographer's decision to convey information through examples rather than d-text, but rather a fact about the word sense and its usage pattern: a different kind of information needed conveying.

6.5.1 The Bigger the Better?

It might be argued that LDOCE is a small dictionary, and it simply did not make sufficiently fine distinctions for each usage to find its true sense; a larger dictionary would allow the process to be more precise. My experience indicates the opposite. In general, the more possibilities there are to choose between, the more evaluating of different evidence and assessing of contrasting pairs is required, and the classification task becomes still more difficult. Where LDOCE gave only one sense for a word, no difficulties were encountered.[7]

One important exception was collocational information. Where the pattern of words in the text matched, exactly or very closely, a pattern of words found in an example or as a subentry, the usage could be classified directly. Where there were many collocations, clearly presented, the task was greatly assisted, and was turned into a process which could readily be automated. The theme will be developed in Chapter 7.

[7] There could of course have been occasions where none of the LDOCE senses fitted. None of these occurred. This was in part because LDOCE is a well-researched dictionary and the sample sizes were small, but also because there is usually potential for re-interpreting a 'misfit' usage as an original use of the one sense. Only when there are two (or more) senses, so the question, 'is the misfit nearer the first or the second?' arises, does the misfit defy classification as one and only one sense. This treatment tallies with the problem faced by NLP. Where a word has only one sense, the NLP system is not faced with the kind of problem the study addressed, even if, occasionally, the word is used in a way that only bears a remote relation to the dictionary entry.

6.6 Conclusion

The sort of variation in meaning illustrated by 'bank' is far from typical of the kinds of variation to be found in the dictionary. Where a word has more than one dictionary sense it is relatively uncommon that each of a small sample of usages of the word can be classified into one and only one of the senses. There are many occasions where a word is used in a way which is indeterminate between alternative senses of the word, or the sense distinctions are not made sufficiently clearly in the dictionary to permit classification, or a large amount of contextual or world knowledge needs to be brought to bear to make the classification with any confidence, or two or more non-exclusive senses both apply.

While research which assumes the Bank Model has been productive, the Bank Model assumption limits its range of application. Future research into how NLP systems might use dictionaries, and how they might deal with the multiple senses offered for most words, would do well to drop the Bank Model. The model which replaces it must make allowance for the indeterminacy, non-specificity, and overlap which is pervasive in the match, or mismatch, between a word's usages and its dictionary senses.

Chapter 7

Theory II: A Four-Way Analysis

7.1 Introduction

This chapter brings together the arguments of Chapter 4 and the evidence of Chapters 5 and 6 to complete the account of the nature of polysemy. First we develop the discussion in the introduction of the crossroads nature of polysemy, then illustrate it with a worked example: *at breakfast*. Of the four signposted destinations, two, homonymy and collocations, have a familiar role in dictionaries. Alternations were discovered in the dictionary in the guise of 'bracketed optional parts' in Chapter 5 and are treated extensively in the next two chapters. That leaves analogy, and the next section considers why and where this strategy is needed. Closely related to analogy are those mainstays of so many discussions of innovation in word use, metaphor and metonymy. Section 7.5 places them within the analysis of this thesis. Both analogy and alternation depend on the notion of a semantic field, and section 7.6 provides a brief account of the main difficulties relating to that notion. Finally, we consider the prospects for the 'four ways' of the analysis coming together, so arbitrary choices as to how a usage-type is treated could be avoided.

7.2 The four ways

Polysemy shades into its various neighbours in conceptual space. The situation may be drawn as in Fig. 7.1.

What does this say about the treatment of polysemy in NLP systems, and where does it leave us in relation to Wilks's problem? The position adopted here is that the four neighbours shown in the diagram each require a distinct kind of treatment, and an NLP system should treat any putative instance of polysemy as a case of one of them. Polysemy is not a natural kind, and there will be no processing strategy specific to it.

The issue focussed on here will be this: what representations and knowledge sources are needed, so that a description is available for all the uses of all the words of a language? A solution to this question is a preliminary to any satisfactory solution to Wilks's problem. Unless all the possibilities for all words are available, it will certainly not always be possible to interpret correctly.

Dictionary entries can be conceptualised as having left hand sides and right hand sides. The standard way to use a dictionary is to match a problem word or phrase against a left hand side, and thus to access the information on the right hand side. The left hand sides are expressions of the language under consideration: the right hand sides may or may not be in the same language. Where a collocation is listed in a dictionary, the multi-word unit is the left hand side. In homonymy or dictionary polysemy, there are several left hand sides with the same form. Wilks's problem

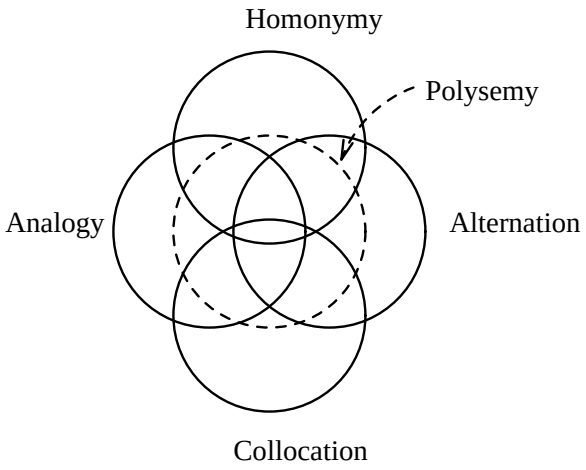


Figure 7.1: Polysemy and its neighbours

is to match the appropriate left hand side(s) for a usage in an input text. The task here is not to describe the form or content of the right hand side, nor to explain how the information might be gathered (though both tasks are closely related to the task at hand when a usage-type is implicitly rather than explicitly represented). The task at hand is simply expressing all the usage-types, that is, making available all the possible left hand sides. We shall only be concerned with any associated information in as much as it is necessary for identifying what usage-type we are expressing. The four ways of treating different cases of *prima facie* polysemy are, then, four ways of representing usage-types.

They are:

1. Homonymy. This is the straightforward case. Usage-types are expressed through being listed. They are represented in the lexicon as distinct one-word entries.

2. Alternation. A system of rules indicates how a non-basic usage-types may be inferred from a basic one.

3. Collocations. For a usage-type which co-occurs only with a limited range of words, all the collocations are listed. Thus the adjectival sense of *frontal* which means 'direct and obvious' (COBUILD) seems to occur only with *attack* and *assault*.[1] The two collocations can be stored in the lexicon. Thus there is a left hand side in the lexicon for each of the variants on the 'direct and obvious' usage-type, and no non-collocational entry for it is needed.

4. Analogy. As with alternations, the usage-type is available as an inference rather than an axiom in the lexicon. Unlike alternations, the inference cannot be made on the basis of facts and rules in the lexicon alone. It can be modelled as follows:

   > Two words, x and y, have a similar meaning in their primary senses, x_1 and y_1, and x has a familiar secondary sense, x_2. Then if y is used in the sort of context where x_2 is often used, y will be interpreted as the novel y_2, relating to y_1 in the same way that x_2 relates to x_1.

[1] The *frontal nudity* and *frontal lobotomy* senses are given separate treatment in COBUILD.

General knowledge and reasoning may be[2] required to determine how the x to draw the analogy from is to be chosen, and how the mapping from x_1 to x_2 is to be interpreted in relation to y.

Different methods will have implications for issues of representation and of processing. The leading question for NLP concerning homonyms is the processing one: how to resolve usages? Concerning collocations, key questions are: how do we automatically identify them from corpora (Church & Hanks, 1989; Smajda & McKeown, 1990), and at what points in parsing do we see whether a string of words from the input text forms an expression listed in the lexicon? Where a collocation or homonym analysis is adopted for a usage-type, the question, 'How do we represent the usage-type?' is answered trivially: we explicitly list it. For usage-types represented using alternation or analogy the representational question is an interesting one (addressed in detail, in relation to alternations, in Chapters 8 and 9).

7.3 *At breakfast*: a worked example

At breakfast[3] falls somewhere near the middle of Fig. 7.1. A lexicographer, computational or otherwise, might adopt any of the four methods in order to represent the usage-types of *at* and *breakfast* it embodies. Let us see how the four options are related to four dictionaries, LDOCE, OALDCE, COBUILD and the Concise Oxford.

1. The 'homonymy' strategy, as applied to *breakfast*, would involve listing the usage-type that occurs in the collocation as a distinct sense, at the lexical entry for *breakfast*, where it might be defined as 'the time and place of eating breakfast1' (where breakfast1 is the core sense: 'the first meal of the day'). None of the dictionaries do this. The homonymy strategy could also be applied to *at*, and this is what COBUILD does. Under *at*, there is a subentry:

   > If you are **at** breakfast, lunch, etc., you are eating your breakfast, lunch, etc. EG *I was still home at lunch*

   This is the most explicit entry that any of the dictionaries have for the construction. (It is in fact inaccurate. One can perfectly well eat nothing at breakfast.)

2. The 'alternation' strategy, in relation to *breakfast*, involves noting in the lexicon that *breakfast* is a member of a class of meal-words, and that meal-words have a time-and-place sense. All four dictionaries do this implicitly. Each gives a single sense for breakfast, with *meal* as genus term. In all four dictionaries *meal* has two senses, 'amount of food' and 'occasion', with the 'occasion' listed first three times out of four. One way of interpreting this is that, by virtue of its genus term, *breakfast* is classified as a word to which a food/occasion alternation applies.

   A snag is that, while all the meal terms can be used with *at*[4] without a determiner, *meal* itself cannot: $*$*at meal*.[5] Perhaps noting the difficulty, but still wanting to indicate, at the genus term's entry, that the pattern applies to all the subordinates, LDOCE uses *at meals* in an example.

   The alternation strategy as applied to *at* would involve noting that *at* was a member of a class of prepositions subcategorising for bare nouns with 'occasion' senses. There is no clear evidence for this treatment in the entries for *at*.

[2] Only 'may be' because statistical methods of the kind described in section 3.3.8 could point the way to solutions using simple pattern-matching. See also section 7.7 below.

[3] The example is borrowed from Wilensky (1990).

[4] And in certain other contexts: "Breakfast will sort that out!"

[5] Or $*$"Meal will sort that out".

3. The collocation strategy treats the whole collocation as a left hand side. None of the dictionaries explicitly list *at breakfast*, either as an LHS to be defined or within the body of a definition or example.

4. The analogy strategy would involve drawing an analogy between *at*, *breakfast* or *at breakfast* and a word or phrase exhibiting similar behaviour which was explicitly listed. The analogy would then support an inference from the explicitly-listed usage-type to the unlisted ones for *at* and *breakfast*. The analogy method is the most open of the four, and it is possible to see the basis for the dictionaries expressing the *at breakfast* usage-types in a variety of places, with varying degrees of conviction. *Dinner* is similar to *breakfast*, and three of the dictionaries list the expression *at dinner* somewhere in the entry for *dinner* so this is one place where one might argue that the dictionaries were conveying, by analogy, how *at breakfast* was to be interpreted.

What considerations might have driven the lexicographers' choices of method for representing the usage-types in *at breakfast*? They include:

- the size and coherence of the set of words with a sense like breakfast1 and a corresponding *at X* collocation.
- the size and coherence of the set of words and phrases that might stand in place of *at* and invoke the 'event' reading, and how varied they are in syntactic structure.
- any non-standard syntactic or pragmatic features associated with the collocation, and, if there are any, how non-standard and how narrowly-applicable they are.
- absolute and relative frequencies: if most usages of *breakfast* were in the *at breakfast* collocation, but for other meal-words (of similar overall frequency) the *at X* form was a rarity, then there would be a case for treating *at breakfast* as the core case with other meal-words drawing analogies from it.

  The same would apply if *breakfast* and other meal-words all occurred in the *at X* construction for roughly the same proportion of their total occurrences, but *breakfast* was far the highest frequency word in the set. This can be seen in relation to *drug addiction*. The relative frequency of the collocation with respect to all occurrences of *drug* is probably similar to those of *cocaine addiction* and *opium addiction* relative to *cocaine* and *opium*, yet the former has a higher absolute frequency and is the most likely to occur in a dictionary. In COBUILD, *drug addiction* is the only one of the three to occur, and it occurs twice, under *drug* and under *addiction*.

As it happens there is a small but highly coherent class of words with the same pattern as *breakfast* — the meal-words.[6] This alone suggests it would be inappropriate to list a homonym or collocate at the entry for *breakfast*. To do so would be to ignore the generalisation that *breakfast* fell under.

The frequencies of the different meal-words are of the same order of magnitude: in the LOB corpus, *breakfast* occurs 54 times, *lunch*, 66, *dinner*, 99, and *supper*, 32. The relative frequencies of occurrences in *PREP X* constructions with meanings parallel to *at breakfast* are also of the same order of magnitude, as the table shows. (The numbers of occurrences of *at X* expressions in the corpus are too low to support any conclusions if taken alone.)

[6] There are various other classes of words following the pattern, for example many institution words (*at hospital/church/court*) and all sports words (*at football/rowing/ping-pong*) but as these do not form a coherent class with the meal-words they will not be considered further. *At tea* and *at coffee* are, like *at breakfast* associated with a regular event involving ingesting, so might give cause to extend the class of meal-words to 'regular-ingesting-event' words.

Word	Freq.	*at*	*to*	*for*	*before*	*after*	others	Total	Rel. freq.(%)
breakfast	54	2	3	2	1	5	1	14	26
lunch	66	1	8	5	2	9	4	29	44
dinner	99	11	2	12	4	9	1	39	39
supper	32	0	1	5	1	4	0	11	34

Table 7.1 Frequencies of PREP MEAL-WORD expressions. The right-hand columns show the frequency of the construction occurring with any preposition and the relative frequency of the construction: its frequency as a percentage of all occurrences of the word. The central columns show the frequencies of the word in the construction with the common prepositions for the construction. 'Others' were *during, from, by, of.*

This shows it would not be appropriate to list *at X* under one of the meal-words and use that as the basis of an analogy for interpreting the others. They should all get equivalent treatment. The entry for *breakfast* should not contain specifications about the *at X* usage: such information would better be factored out to some location where it related equally to the whole homogeneous class. Membership of the class is signaled by having *meal* as a genus term.

The non-standard syntax of *at MEAL-WORD* does not apply to *at* alone. *To, for, before, after, during, from, by* and *of* share the pattern, combining with determinerless meal-words to give rise to a usage-type very similar to the *at* one. The freer the range of contexts the usage-type can occur in, the stronger the case for a representation that makes no reference to syntactic or lexical context. But the syntactic context is constrained: it is only after this small range of prepositions that the meal-words appear, determinerless and with this meaning. An 'alternation' account can capture the fact that meal-words are all subject to the food/occasion alternation: it will need pairing with syntactic specifications for determinerless uses.

As discussed so far, the SFIP criterion (see section 4.6) has only been applied to single words, one at a time. We see here that a version of the SFIP criterion applies to patterns as well: a non-negligible frequency of occurrence is a pre-condition for any linguistic fact receiving explicit treatment in grammar or lexicon. The fact can be a usage-type for a word but it can also be a collocation, an alternation, or a syntax-semantics correlation to be found in a particular construction. We also see that the two halves of the SFIP criterion are intimately linked. How predictable a usage-type is depends substantially on the frequency with which other words (both tokens and types) are found to follow the pattern the target word has followed. The choice between strategies for *at breakfast* depends on the frequency of instantiation and predictive power of the various competing generalisations it falls under.

7.4 The analogy strategy

7.4.1 Why we need it

A general-purpose NLP system must be able to deal with word uses not previously encountered. A text may always contain novel word-uses, and people are perfectly well able to process them.

The basic method for an NLP system to do this will be one in which there are rules for generating a set of available, but not necessarily attested, usage-types on the basis of a primary sense. The formal theory developed in Chapters 8 and 9) demonstrates how a lexicon can contain implicit information about the range of possibilities for a word, so a compact lexicon can express a very large number of rule-following possibilities many of which will never have been encountered or explicitly anticipated.

If an exhaustive list of alternations could be drawn up, along with one classification scheme for words and another specifying which alternations applied to which classes of words, then the

alternation approach could be used to represent all possible usage-types for words. It would be the strategy to account for all occasions where words were being used in ways unfamiliar to the system.

There are domains where the alternation strategy is appropriate, but there are both practical and theoretical difficulties to the idea of making exhaustive lists. The practical difficulty relates to the twin problems of specifying alternations and specifying the classes of words they apply to. The cataloguing of alternations will, for difficult cases, have all the characteristics of the cataloguing of usages (see Chapter 6). Alternations are abstractions of a higher order than word senses, and their identity conditions are correspondingly harder to determine. The study described in Chapter 5 had as its goal, a classification of the distinctions between the different dictionary senses for words, and one of its conclusions was that the distinctions are not, for the most part, readily classifiable. It is true that dictionary senses will in general be less predictable than those possible usage-types for a word that are not listed in dictionaries —this much the SFIP criterion tells us— but nonetheless the study suggests it would be unrealistic to expect a process of classifying usage-types into 'basic' and 'alternation' components to be without large numbers of approximations and arbitrary decisions. For issues relating to the identification of 'semantic fields', or the classes of words that alternations apply to, see section 7.6.

The theoretical difficulty underlying the practical ones is the lack of identity criteria for alternations or semantic fields. Sometimes our intuitions are clear: it is uncontentious that the same alternation, from tree ("We chopped down the ash") to wood ("We want some ash for the shelves"), applies to *ash*, *beech* and *oak*. But our intuitions are no longer so clear when we wonder whether this is the same alternation that applies to *bamboo*. Intuitions are often unclear, and then the researcher or lexicographer is faced with a tangled web of interrelated meanings. There are no criteria for answering "Does the wood/tree alternation apply to *bamboo*?"

In general, the intuition that a single alternation applies to a number of words is clearest when all the words are core cases of the same semantic field. The problem with *bamboo* is that, since bamboo is not a typical tree, the word is marginal to the semantic field. In general, the alternation module will map out the usage-types for words in a clear and principled manner in domains where the usage-types fall squarely into regular semantic fields.

How then shall we analyse word usages where the usage fits none of the known senses, nor any well-defined alternations from those senses, but where the relation to the known sense can be attributed only to an ill-defined alternation relating to members of an ill-defined semantic field? The proposal underlying the analogy strategy is simply that the behaviour of a word can be modelled on the behaviour of its 'near neighbours'. (For issues to do with identifying near neighbours, see sections 3.3.8 and 7.6.)

7.4.2 An example, and why they are hard to come by

Consider the recently attested expression, "I want to keep my finger in", a variant of the idiomatic "I want to keep my hand in". It is clearly not familiar or lexicalised. Its frequency is negligible (if it has ever occurred at all before). Its interpretation, in context, was transparent. The semantic relatedness of *finger* and *hand*, together with the lexicalised *keep one's hand in*, left no doubt about what was meant. The other participants in the conversation did not note anything abnormal.

The situations where the analogy module would be used are where

1. A sense matching the usage is not in the lexicon.
2. The word is not readily accounted for by the alternation module: that is:
   (a) it is not a member of a well-behaved semantic field to which
   (b) a regular alternation applies.

3. A near neighbour has a sense which will suit the context.

4. The near neighbour's sense can be analysed as a basic sense (corresponding to a basic sense of the target word) and a sense-transfer component, and the sense-transfer can be applied to the target-word's basic sense.

5. The analysis provides more information than the context alone would, or than the context combined with an underspecified sense for the target word would.

"I wanted to keep my finger in" meets all the conditions. But the conditions combine to present certain difficulties in presenting an example not based on an idiom. In section 4.3 and again in the matching study (Chapter 6) we noted the theoretical possibility that a usage fits none of the dictionary senses for the word. However in the study, excluding odd cases such as the word being used as a name or in a quote, this did not occur. This was because a sense definition could always be interpreted as sufficiently open to admit usages 'close' to a prototypical usage for the sense. Until another sense focussed the question, changing it from, 'Is this usage near this sense?' to, 'Which of these senses is the usage nearer?', there were never sufficient grounds for saying a usage was so anomalous it did not fit under a loose reading of any of the definitions. The goal of the analogy module would be, as condition 1 states, to cater for some of the usage-types not listed in the lexicon. But human readers are almost always able to interpret usages as fitting one or other dictionary sense. The 'analogy modules' in our heads are inevitably brought into play as we consider whether a usage fits a sense. We stretch the imprecise boundaries of a sense definition so it does. Even where an example of a would-be analogy usage-type is not interpreted as fitting a listed sense, it is very likely to be considered an alternation. An analogy, once named, is a pattern relating usage-types— so invites treatment as an alternation. Once spotted, analogies turn into something else — but there will always be further analogies, not yet spotted.

There will only be a population of misfit usages, fitting neither listed senses nor an alternation treatment, when the limits of dictionary senses and alternations are well-defined. For humans they are not.

There are not yet computer systems of sufficient sophistication to be able to make subtle sense distinctions outside limited domains. Within some domains, such as Unix Consultancy (Wilensky, 1990), there are systems for which the notion of a usage falling outside the scope of the lexical entry is well-defined, though such lexical entries are dependent on a particular state of development of the grammars, lexicons and discourse modules of those systems so will not provide convincing examples outside that context. In terms of this thesis, the work done on the UNIX Consultant and discussed in section 3.6.1 represents the state of the art in relation to the realisation of the analogy strategy.

In sum, an analogy module is required in an NLP system to stand in for some of the knowledge a human dictionary-user brings to the dictionary, in his or her tacit knowledge of how to flexibly interpret usages given knowledge of senses. The analogy module will cater for the less structured part of that knowledge —the alternation module will cater for the more structured part. The need for the analogy component only becomes evident when the cognitive flexibility with which humans cover the same territory is unavailable.

7.4.3 Analogy: lexicon or pragmatics?

The question arises, what relation to the lexicon does the analogy module bear? The analogy strategy, as sketched, is a pragmatic rather than a lexical matter, clarifying how lexical entries may be reinterpreted in the course of language understanding. All it requires of the lexicon is that it stores information about the word's location in lexical structure, perhaps through listing its near neighbours or placing the word in a taxonomy.

The position adopted throughout the thesis is that usage-types vary in their frequency and predictability, that amongst all the possible usage-types for a word there is only a subset worthy of explicit mention in the dictionary or computational lexicon, and that polysemy is a somewhat arbitrary label for a subset of the usage-types specified by frequency and predictability thresholds. It is not then surprising if, for a further subset of usage-types which might intersect with the 'polysemy' subset, the possibility of the word being used in that way is predictable so very little if anything at all needs adding to the lexicon for them to be interpretable. As we follow the road from the polysemy crossroads towards analogy, we leave the lexicon behind.

7.5 Metaphor and metonymy

These terrible twins loom large in discussions of 'nonliteral' language or of novel uses of words. Readers might be surprised not to have encountered them earlier. The purpose of this section is to show that, in relation to polsemy and the lexicon, they do not raise any issues not already covered in the analysis.

The literature dedicated to defining the terms is extensive and will not be reviewed here (Martin (1990) and Fass (1991) do this quite satisfactorily). Fass notes, "On the whole, the two remain vague, poorly defined notions in the literature" (p 49). He does little to improve the situation: although he claims his computer system identifies both metaphor and metonymy and distinguishes the two, he offers no definitions of what they are or how they differ, saying only that his approach

> is consistent with the view that metaphor is based on similarity, whereas metonymy is based on contiguity (cf. Jakobson and Halle 1956). Contiguity, readers may recall, refers to being connected or touching whereas similarity refers to being alike in essentials or having characteristics in common. ...
>
> In metaphor, an *aspect* of one concept is similar to an *aspect* of another concept ... in metonymy, a whole *concept* is related to an *aspect* of another concept. ...(p 83)

Without an account of where one concept ends and the next begins, or what distinguishes an aspect of a concept from a new concept, or what the limits of 'being connected' and 'touching' are when we move away from the domain of physical objects, this leaves us little the wiser.

Alongside the problematic difference in "kind of relationship" (p 59), Fass follows Lakoff & Johnson (1980) in noting the functional difference between the two, quoting their observation that:

> metaphor is "principally a way of conceiving one thing in terms of another, and its primary function is understanding" whereas metonymy "has primarily a referential function, that is, it allows us to use one entity to *stand for* another" ...(Fass, 1991, p 58)

Given the difficulties encountered by Fass (and others) in giving a clear picture of metaphor and metonymy in terms of kinds of relationships between words and referents, this is interesting in that it suggests part of the answer may lie elsewhere, in the different purposes to which a word is put.

The four-way analysis presented in this chapter has been driven by NLP considerations: there are four kinds of ways in which NLP might address those phenomena that might be called polysemy. We may then ask, do metaphor and metonymy have any features calling for a distinctive NLP treatment?

The question of whether a prototypically metaphoric or metonymic usages should be listed in the lexicon is as for any other usage: yes, if its frequency is great enough and its predictability

low enough. If metonymy is a device used in order to refer, then, the more obvious the RF (see section 4.2) from word to referent, the more effective it will be. For successful reference, it is essential that the hearer chooses the RF intended by the speaker. So the speaker must predict the hearer's choice, and the hearer, retrospectively, the speaker's. To the extent that the basis for prediction is a familiar pattern in the language, the RF will be represented in the lexicon as an alternation: metonymic usages will be interpreted by the same mechanisms as alternations. To the extent that the basis for the prediction is to be found in the particular situation which supplies a one-off RF, the prediction can only be made on the basis of that context and the relevant usage-type can only be made available by analogy.

Martin's work on metaphor is discussed in section 3.6.1. He specifically — and wisely — addresses only conventional metaphor. Conventional metaphors follow conventions, so are fairly predictable. Martin treats conventional metaphor as a set of rules which maps out a search space of possible metaphoric readings. Since there are rules to map out that search space, there are rules which render the readings predictable.

Many prototypically metaphoric uses, like that of *pig*, are sufficiently institutionalised to be listed in paper dictionaries. LDOCE gives,

> **pig** ... **2** *infml derog a* an unpleasant person, esp. who eats too much, behaves in an offensive way, or refuses to consider others

There is something paradoxical about this. The metaphoric use of *pig* is one in which a person is conceived of as an animal. But the corresponding sense is listed in the dictionary with genus term 'person', so if a dictionary user retrieves that sense from the dictionary, there is not so much as a mention of the animal so there is no 'conceiving of one thing as another' to be done. We might posit that the mental lexicon, like the dictionary, stores common usage-types so does not need to reinterpret them on each occasion of use. Conventional metaphor is often in this trap. To the extent that it is conventionalised, it is no longer, in the most interesting, pragmatic sense, metaphor.

In Romeo's

> It is the East, and Juliet is the sun

the use of *East* and *sun* are clearly metaphorical in that, typically, when first encountered, they provokes a new understanding of love, Juliet, dawn and the east. The usage-types will not be listed in the lexicon, or available by alternation. The role of the lexicon is to represent patterns of use of words. But the power of the novel, striking metaphor is dependent on it defying the normal pattern of use for the word. It requires the hearer to go beyond the familiar uses and patterns which are coded in the lexicon, and to bring their imaginative and intellectual capacities to bear. To the extent that a metaphor is novel and striking, it is no longer, in any interesting sense, represented in the lexicon. It is still theoretically acounted for by analogy, but usage-types expressed 'by analogy' may call on any amount of non-lexical knowledge for a full interpretation. A substantive account of any novel metaphor involves a description of the non-lexical knowledge and processes involved, and is well outside the scope of this thesis.

7.6 Semantic fields

The two productive methods of expressing non-basic usage-types both presuppose a notion of semantic similarity. Alternation and analogy both base inferences on the principle that words of similar meaning behave similarly. In areas of discourse where there are sets of words which behave similarly and can be identified as semantic fields, a much-used example being cooking verbs, introspection will serve to identify the fields. Questions of degrees of similarity can, for

some purposes, be put aside. But in general, similarity is a matter of degree, and is a complex notion on which it is unlikely that people will have strong or consistent intuitions.

Homonymy and polysemy are major issues for any similarity measure. *Bishop*, *knight* and *rook* all feature in a 'chess' semantic field, but, outside the chess domain, they are not in the same field. If we could first resolve matters of polysemy and homonymy, that would assist the identification of semantic fields. But semantic fields are needed for the resolution of polysemy and homonymy.

The topic of semantic classification is a large and important one. It has been looked at from many angles, some statistical, some seeking taxonomies of concepts (Sparck Jones, 1986; Berlin, 1978; Dahlgren, 1988; Beckwith et al., 1991, see also sections 3.3.8 and 8.2). The reasons given at the beginning of the thesis for studying polysemy could readily be revised to give reasons for studying semantic fields. The two topics are closely related. Both are central to lexical structure. The point here is simply that all the strategies for analysing unfamiliar word-uses depend on semantic classification, but there remain many unanswered questions as to how this might best be done.

7.7 Prospects for a unified approach

In the best of all NLP systems, there would be seamless joins between the four strategies for representing putative polysemy. The whole would be described in one all-encompassing language and the question, "Which strategy is most appropriate here?", would make no more sense, in relation to a case like *at breakfast*, than "Is orange red or yellow?"

The next chapter points the way to such a seamless join between homonymy and alternations. A smooth join between analogy and collocation methods can also be sketched: an analogy for *finger*, in "keep one's finger in", is identified by searching a lexicon for collocations with a similar form to "keep one's finger in" but with one of the words switched for some other word in the same semantic field. An analogy is treated as a modified version of a collocation. An algorithm for interpreting 'analogy' usage-types would require data about semantic fields or near neighbours and a lexicon containing many collocations. It would then seek the best fit, involving as near neighbours as possible, between a collocation in the lexicon and the input string.

What, then, are the chances for bringing these two pairs of approaches under a common roof?

The prognosis on this question is mixed. Alternation provides a rule-based approach to interpreting unfamiliar usages. Analogy deals with rule-defying cases, so cannot use a yes/no criterion for whether a use-pattern is available for a word since the answer by the rules would always be 'no'. It must use measures of closeness and goodness of fit. The rule-based method presupposes that what is possible in the language has been determined by the lexicon-builder and is implicit in the lexicon. For the analogy module, everything is possible: some things are only more likely than others. Numbers representing frequencies of words and word-uses, both absolute and relative to syntactic contexts, collocates, and other members of a semantic field, are essential to assessing likelihoods.

MacWhinney's Competition Model (see section 4.3) provides a picture of how apparently rule-bound patterns are emergent properties of a process with underlying statistical structure, which might suggest that the alternation module should be reinterpreted as a special case of the analogy one. But at this point two different objectives must be distinguished. MacWhinney's goal was to model the acquisition of word-meaning, and to do this he idealised away all the complexity of the context in which the learning took place, into a set of discrete, simple semantic features. For purposes of modelling a particular process, the methodology is quite appropriate. But the model cannot then be expected to be informative regarding the interactions of the process with the context when that context is no longer idealised. An NLP system needs a model of those interactions. The only frameworks available for the task are rule-based ones. While the

mental substrate in which lexical choice, syntax, semantics and discourse processes all interact might, speaking loosely, be 'connectionist', its complexity is such that it is beyond our current understanding to directly model it as a neural network. We can, on the other hand, identify a set of distinct approaches to the task, identify the rules that best describe the workings of each approach, and, so armed, build processing systems that model the whole language-understanding process.

We need both rule-based and likelihood-based strategies, neither succumbs to reinterpretation as the other, and it is hard to see how a framework could reconcile the 'everything's possible' assumption with the picture in which a use either is or is not within the range of possibilities for a word. The theme is taken up in the final chapter under 'Further Work'.

7.8 Summary

In this chapter, we developed the theme of the crossroads nature of polysemy. We argued that an NLP system needs to treat every usage-type for a word either as a homonym, or as part of a collocation, or as the result of applying an alternation to a basic sense, or as a product of analogy. Through examining the treatment of the expression *at breakfast* in four dictionaries we saw how the four ways related to lexicographic practice. We also sketched the considerations that determine which of the four ways is most appropriate for a given usage-type.

Whereas homonyms, collocations and (as shown in Chapter 5) alternations have a familiar place in the lexicon, analogy is a less obvious member of the coterie. Section 7.4 explained why there will always be usage-types not represented by other methods, and an analogy strategy is needed to complete the analysis. Section 7.5 considered metaphor and metonymy, two phenomena often discussed in relation to innovative language use, and showed how they fitted within the analysis. Metaphorical and metonymic usage-types, like all others, fall under the SFIP criterion: they would be listed in a paper dictionary where they were sufficiently frequent and insufficiently predictable. Many conventional metaphors meet this criterion, and should receive homonym treatment. Metonymy often involves a familiar mapping from a word's base sense to what it is being used to refer to, so invites alternation treatment. Original and unusual metonyms and metaphors will be treated as analogies.

Both alternation and analogy treatments make use of a semantic classification scheme. Section 7.6 makes the point that the concept of a semantic field, while convenient for expository purposes, is deeply problematic. Section 7.7 considers the prospects for a description of the ways words are used in the language in which the joins between the four ways do not show, and finds a topic for further research in the join between rule-based and statistical approaches.

Chapter 8

Formal Lexicography I: Noun Alternations

8.1 Introduction

This chapter takes those cases of polysemy that are susceptible to treatment as alternations, and presents a formal method for concisely capturing the generalisations and, potentially, making them available for exploitation for lexicography and NLP.

To recap: alternations, or 'regular polysemy' (Apresjan, 1974) occur where two or more words each have two senses, and all the words exhibit the same relationship between the two senses. An example, taken direct from LDOCE, is:

> **gin** (a glass of) a colourless strong alcoholic drink ...
>
> **martini** (a glass of) an alcoholic drink ...[1]

In each case, there are two senses referred to, one with the 'bracketed optional part' included in the definition and the other with it omitted, and the relation between the two is the same in both cases.

Recent work on lexical description has stressed the need for the structure of a lexical knowledge base (LKB) to reflect the structure of the lexicon and for the LKB to incorporate productive rules, so the rulebound ways in which words may be used are captured without the lexicon needing to list all options for all words (Boguraev & Levin, 1990; Gazdar, 1987; Pustejovsky, 1991a). The generalisations regarding regular polysemy should be expressed in the LKB, and the formalism in which the LKB is written should be such that, once the generalisation is stated, the specific cases follow as consequences of the inference rules of the formalism.

As 'lexicalism', the doctrine that the bulk of the information about the behaviour of words should be located in the lexicon, has become popular amongst unification grammarians, so formalisms for expressing lexical information have been developed. Some part of the syntax, semantics and morphology of most words is shared with that of many others, so the first desideratum for any such formalism is to provide a mechanism for stating information just once in such a way that it is defined for large numbers of words. Inheritance networks serve this purpose. The next requirement is that exceptions and subregularities can be expressed. It must be possible to describe concisely the situation where a word or class of words are members of some superclass, and share the regular characteristics of the superclass in most respects, but have different values for some feature or cluster of features. Several lexical representation formalisms addressing

[1] As the LDOCE entry for *glass* notes, a receptacle need not be made of glass to be a glass.

these desiderata have been proposed (DATR: (Evans & Gazdar, 1989a, 1989b, 1990), Russell, Ballim, Carroll, & Armstrong-Warwick (1991), Copestake (1991)). While the generalisations to be formalised are better understood for morphology and syntax, the theoretical gains, of capturing generalisations and eliminating redundancy, and the practical benefits, in terms of lexicon acquisition and maintenance, apply also to regular polysemy.

The work described here will take the DATR formalism and use it to represent a collection of facts and generalisations regarding polysemy. This chapter and the next use DATR, and a brief description will be given below, but the thesis neither presupposes a knowledge of it, nor gives a formal description. Evidence regarding regular polysemy will be introduced, in stages, with the proposed DATR account of the evidence worked through at each stage. The sense that results from excluding the bracketed part, or which is listed first in the dictionary, or which is the only one listed, will be deemed the 'primary' sense, with others 'secondary'.

8.2 Taxonomies of words and of their denotations

In the fragment presented here, facts about both the word and its denotation are accessed through the same node in the inheritance network. Thus a query regarding the syntax of the word *beech*, and a query asking what type of thing a beech tree is, will both be made at the same node. It might be argued that this is to confuse two different kinds of information. The position taken here is that there is much to be gained from holding the two types of information together, and to keep them separate is to forgo opportunities for expressing and exploiting generalisations, and to force a wide range of arbitrary decisions and duplication. The position is related to the central tenet of cognitive linguistics (see section 3.5) that linguistic meaning must be studied in the context of the overall cognitive system, though the claim here is clearly of much narrower scope.

As described in Chapter 2, Amsler (1980) and many others have shown the dictionary embodies a taxonomy. The taxonomy is primarily a taxonomy of denotations, and unearthing a rudimentary structure for human general knowledge, for use in AI knowledge representation, was Amsler's goal. The non-linguistic knowledge in a monolingual English dictionary is stated in English and the labels for the nodes in the taxonomy are English words, so in the course of expressing non-linguistic facts about beeches and trees, the dictionary provides *tree* as the genus term for *beech*, thus alluding to the potential for inheritance between the words as a side-effect. The example sentences given for *tree* (sense 1) in LDOCE include *to climb a tree; to plant a tree; to cut down/chop down a tree*, and a dictionary user would be correct in interpreting these as collocations in which *beech* can substitute for *tree*. Whether this is at all a consequence of linguistic as opposed to encyclopedic facts is possibly an unanswerable question, but it is not a matter which need concern the dictionary-user since the encyclopedic cargo and the linguistic vessel are both making equivalent journeys.

Collocational information is one kind of linguistic information which is, to a substantial degree, predictable from word meaning. The subject of this chapter, regular polysemy, is another. *Martini* participates in the glass-of/drink alternation because martini is a drink, and if we discover a new drink called foobaz we know we can order a 'count' foobaz as well as drink a lot of 'mass' foobaz. Alternations, by definition, apply to classes of words. The classes are formed according to the words' denotations, and words will generally participate in the same alternations if their meanings or denotations are similar. Taxonomies, whether in biology, dictionaries, or AI knowledge representation group similar things together, and non-linguistic taxonomies will often identify the classes of words to which alternations apply.

LDOCE commonly uses the taxonomy defined by genus terms to express alternations applicable to both the genus term and its subordinates. Thus in

> **daffodil** a very common bell-shaped pale yellow flower ...

dahlia a big brightly-coloured garden flower ...

flower 1 the part of a plant, often beautiful and coloured, that produces seeds or fruit ... **2** a plant that is grown for the beauty of this part ...

LDOCE encodes implicitly that both *daffodil* and *dahlia* can be used in both the first and second senses of *flower*. The dictionary depends on the reader's ability to spot that senses 1 and 2 are intimately related so, all else being equal, they will be inherited as a pair. One or other of the senses for *dahlia* and *daffodil* has failed to meet the SFIP criterion (section 4.6), in that it is insufficiently frequent and unpredictable to be separately listed in the dictionary, but with an intelligent reading of the dictionary's taxonomy, both senses can be seen as implicitly present.

Building general-purpose taxonomies for NLP, thesauri such as Roget's, AI knowledge representation or other purposes is a difficult and lengthy process, in which there are various different sources of evidence to be brought to bear in attempting to reach principled decisions regarding the overall shape of the taxonomy (see. *e.g.*, Dahlgren (1988)). The lexicons of one or more languages inevitably play a large role in determining what the structure of the taxonomy should be. It is often impossible to say whether the words used in such a process are simply acting as carriers of non-linguistic, conceptual information, or are colouring the conceptual information with the particularities of a word, language or family of languages (Quine, 1960). The process of building a taxonomy of words is, thus, inextricable from the process of building a taxonomy of things. Evidence from all sources is needed for a single taxonomy. So a further argument for attaching lexical information to a general-purpose taxonomy is that both emerge from substantially the same evidence, and dividing the evidence between two structures will amount to weakening the empirical support for both.

A related point is that, where an alternation applies to a class of words in a general-purpose taxonomy, the situations where the alternation applies are likely to be motivated as well as described. If the lexical taxonomy only contained information about words, then the fact that all drink-words participate in the glass-of/drink alternation could be described, but could not be related to the 'explanation' that this was because they all denoted drinks, since the relation between the denotation of *martini* and that of *gin* would not be represented. A framework in which lexical and general knowledge are held together offers a better chance of relating the classes of words to which alternations apply, to classes identified for independent reasons, which in turn improves the chances of predicting what alternations apply to a word from a rudimentary knowledge of its use.

An argument against holding lexical and general knowledge in the same taxonomy is that the lexicon contains idiosyncrasies in a way in which the non-linguistic world does not. Thus *oats* and *wheat* have similar denotations, occupying similar locations in a taxonomy of the natural world, but in the lexicon, one is singular while the other is plural (Palmer, 1976, p 119). For simple inheritance systems this may well present difficulties, but that merely indicates the inadequacy of such simple systems. DATR is designed around the need to express exceptions as well as regularities. It was devised with lexical representation particularly in mind, and has not yet been used for representing general knowledge beyond toy examples, so its potential in this direction is as yet unexplored, but it makes clear that the exception-ridden nature of the lexicon is not a reason for regarding it as outside the scope of knowledge representation languages.

There are, then, several arguments for using a single taxonomy for words and for things. First, it will avoid 'territorial' issues: "Is this fact lexical or general?". Second, it will avoid the need to duplicate information and structure between two parallel taxonomies. Third, the sources of evidence for building the two in any case overlap. Fourth, a single taxonomy will facilitate the comparison of classes identified on lexical and non-lexical grounds, giving more scope for predictive power. And finally, the counter-argument based on lexical idiosyncrasies is irrelevant once a knowledge representation scheme which allows for exceptions and subregularities

is adopted.

Once the argument is granted, a technical question arises: should words of English be associated only with terminal nodes of the taxonomy or should they also be allowed to occur as non-terminals. The motivation for the terminals-only proposal is that it would make a formal distinction between word-nodes, which are similar to dictionary entries and are the only nodes at which queries would ordinarily be made, and non-word nodes which are purely theoretical constructs used to organise information about words. If the proposal was adopted, the consequences would be that a word which occurs as a genus term in the dictionary, such as *tree*, would effectively acquire two nodes: one for the information that either applies to all trees or is shared between *tree* and all words for which it is a genus term, and the other for information which is specific to the word *tree* or otherwise is not to be inherited from *ash, beech* etc. The latter would be the terminal, 'dictionary entry' node which would inherit all of its non-linguistic and most of its linguistic information from its twin, the non-terminal tree node. The basic taxonomy provided by the dictionary then acquires a further set of nodes, one for each word that occurred as a genus term, one hanging off each non-terminal node of the basic structure.

The terminals-only policy has been found to clarify the theory. Without it, where a genus term has characteristics which apply to the genus but not its subordinates, care must always be taken to block the inheritance. We also adopt a syntactic convention for distinguishing terminal 'word' nodes from non-terminals: the former have the first letter in upper case and the remainder in lower case, whereas non-terminals are all capitals.

8.3 Overview of DATR

Evans & Gazdar (1989a, 1989b) presented the basic features of DATR. Here we briefly review those features: more detailed discussion accompanies the formalisation developed below.[2] DATR is a declarative network representation language with two principal mechanisms: orthogonal multiple inheritance and nonmonotonic definition by default. The primary unit of a DATR network description is called a **node** and consists of a set of *path/definition* pairs where *path* is an ordered sequence of arbitrary atoms (enclosed in angle brackets), and *definition* is either a value, an inheritance specification or a sequence of definitions. **Nodes** are syntactically distinct from other atoms: they start with a capital letter. The primary operation on a DATR description is the evaluation of a query, namely the determination of a value associated with a given *path* at a given **node**. Such a value is either (a) defined directly for *path* at **node** or (b) obtained via an inheritance specification for *path* at **node** or (c) determined from the definition for the longest leading subpath of *path* defined at **node**, when *path* itself is not defined at **node**.

Inheritance specifications provide a new node, new path or both to seek a value from. The simplest form of inheritance, called 'local' inheritance, just changes the node and/or path specification in the current context. To specify that `<path1>` at `Node1` inherits locally, we use one of the following.

```
Node1:<path1>   == Node2.
```

specifies that we inherit the value from `<path1>` at `Node2`.

```
Node1:<path1>   == <path2>.
```

specifies that we inherit the value from `<path2>` at `Node1`.

```
Node1:<path1>   == Node2:<path2>.
```

[2] The section borrows heavily from Cahill & Evans (1990)

specifies that we inherit the value from <path2> at Node2.

As well as local inheritance, there is global inheritance. DATR stores a 'global context' node and path, and where a DATR sentence specifies global inheritance, then information is inherited from the global context. Double-quotes are used. Thus:

```
Node1:<path3>   == "<path4>".
```

specifies that the value for path3 at Node1 is inherited from the value of path4 at the global context node (and the global context path is changed to path4). When a query is made to a DATR theory, global contexts are initialised to their values in that initial query, and 'quoted paths' such as "<path4>" are often used to say, "go back to the node the query was initially made at and see if there is a value for this path there".

When a requested path is not defined at a node, the longest subpath (starting from the left) is used to provide a definition, with all the paths (if any) in the definition specification extended by the extra requested atoms. Thus if paths <a b c> and <a b c d> are not defined at Node1, a definition such as:

```
Node1:<a b>     == Node2: <x>.
```

implicitly defines both the following:

```
Node1:<a b c>   == Node2:<x c>.
Node1:<a b c d> == Node2:<x c d>.
```

This 'definition by default' (in the absence of any more specific path definition) gives DATR its nonmonotonic character: add a definition to a node and some of the theorems which were previously valid, but derived by this default mechanism, may cease to hold.

DATR has to date been used as a formalism for expressing syntactic, morphological, phonological and a limited amount of semantic lexical information (Evans & Gazdar, 1990; Cahill, 1990; Cahill & Evans, 1990; Gibbon, 1990). Polysemy has been addressed only briefly, in Cahill & Evans (1990), and that account makes no mention of the generalisations to be made regarding polysemy.

8.4 Trees, wood, fruit: a DATR fragment

The data we shall consider will concern trees, wood and fruit. Firstly, consider the following definitions, from LDOCE.

> **ash** (the hard wood of) a forest tree ...
>
> **beech** (the wood of) a large forest tree ...

The bracketed optional part mechanism, combined with the near-identical form of words within the brackets, suggests an alternation, and indeed the tree/wood alternation applies to most if not all trees. In the basic taxonomy of the domain Ash and Beech inherit from TREE, which in turn inherits from PLANT which, in the fragment offered here, inherits directly from ENTITY.

```
PLANT: <>       == ENTITY.
TREE: <>        == PLANT.
Ash: <>         == TREE.
Beech: <>       == TREE.
```

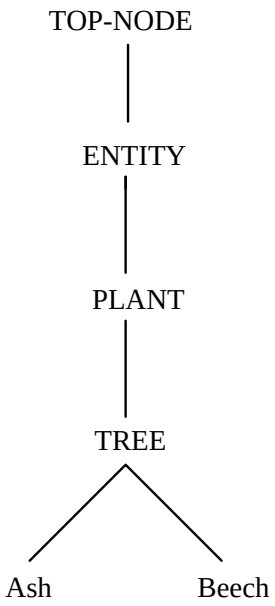


Figure 8.1: A simple taxonomy

This is shown as a taxonomy in Fig. 8.1.

To that basic structure, we wish to add a generalisation about 'wood' senses. Once we have established that *ash* is being used in its wood sense, we wish to treat the word as we would *teak* and *mahogany*. We need to distinguish secondary senses from primary ones in such a way that the paths for accessing information about them are different. We do this by prefixing the path with `alt` (for alternation). There might be several alternations, so we identify the alternation by the path element following `alt`, the 'label', for which we shall use the genus term of the alternate sense, here `wood`. Let us also add some flesh to the bare bones of the taxonomy, and state some genus terms, `word` values (i.e. the word associated with the node), and `collocates`, words commonly found as near neighbours of the target word, at various low-level nodes. The next version of the theory, to be explained below, is:[3]

```
TOP-NODE: <collocates>    == .
ENTITY: <>                == TOP-NODE.
PLANT: <>                 == ENTITY.
TREE: <>                  == PLANT
    <collocates>          == plant grow chop-down PLANT
    <genus>               == tree
    <alt wood>            == WOOD:<>.
Ash: <>                   == TREE
    <word>                == ash
    <alt wood collocates> == black TREE.
Beech: <>                 == TREE
    <word>                == beech.
```

[3] In accounts of DATR published to date, sequences are enclosed in round brackets. However the brackets are redundant in that they can be omitted without ambiguity, and future definitions of the language will not include them, so they have not been included here.

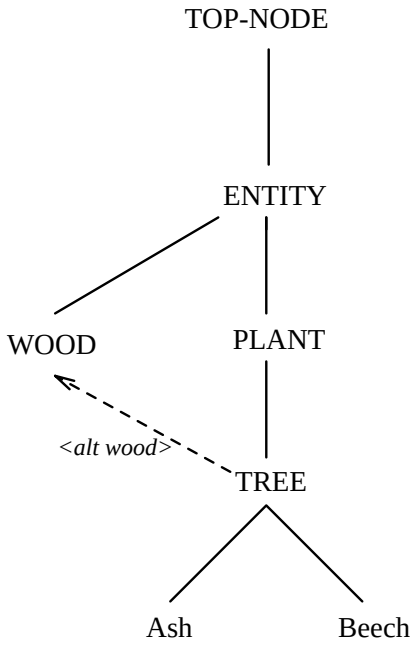


Figure 8.2: Taxonomy showing tree/wood alternation

```
WOOD:<>                  == ENTITY
    <collocates>         == burn saw plane ENTITY
    <genus>              == wood.
```

The network this describes is shown in Fig. 8.2. Now, if we query the system to obtain a value for

`Ash:<genus>`

we do not find `Ash:<genus>` explicitly defined at the `Ash` node, so, as specified in the line

`Ash:<> == TREE.`

we inherit by default from `TREE`. The path is still `<genus>`, so the query is now

`TREE:<genus>`

and this is explicitly defined, in the line

`TREE:<genus> == tree.`

so the value for the query is `tree`. Informally, we have looked 'up' the inheritance hierarchy until we have found a place where the path we are interested in is defined. For

`Ash:<alt wood genus>`

the `alt wood` path prefix diverts the inheritance to `WOOD`. The effect of the empty path on the right hand side of the equation

`TREE: <alt wood>  ==  WOOD:<>.`

is to direct the inheritance to the `WOOD` node with the path prefix replaced by the null path. In this case, that leaves the path `<genus>`, which is evaluated at `WOOD` to give `wood`. This 'prefix-stripping' action gives the desired behaviour in that, once we have specified that we have the 'wood' sense of a word, it behaves as if its primary sense were a 'wood' sense.

The axioms involving `collocates` use DATR's sequence facility. The rule for evaluating sequences is that each sequence member is treated as if it were alone on the right hand side of

the equation, is evaluated, and the value is placed back in the sequence. If we wish to find the collocates of the 'wood' sense of *ash*, we need to evaluate

```
Ash:<alt wood collocates>.
```

At the `Ash` node, we find an equation for which the left hand side matches, and the right hand side is a sequence. The first element of the sequence is an atom, `black`, so that becomes the first element of the sequence that is returned. The second element is not an atom, but a node, so for the remainder of the sequence we need to evaluate

```
TREE:<alt wood collocates>.
```

The prefix is stripped and the query diverted to `WOOD`, as we saw above, so we next need to evaluate

```
WOOD:<collocates>
```

which is again specified as a sequence. The first three elements are atoms so are returned unchanged. The last element is a node, `ENTITY`. We pass straight through `ENTITY` to `TOP-NODE`, and find that there `<collocates>` is defined as the null sequence. We now have all the members of the sequence that forms the value for the original query. The empty sequence disappears and the value returned is

```
black burn saw plane
```

as desired.

There is also a syntactic distinction to be noted. The 'tree' senses are count nouns whereas the 'wood' senses are mass nouns. So let us add a little syntactic information to the fragment.

```
ENTITY: <syntax>        == NOUN:<>.
WOOD:   <syntax>        == MASS-NOUN:<>.
NOUN:   <cat>           == noun
    <concrete>          == yes
    <count>             == yes.

MASS-NOUN:<>            == NOUN
    <count>             == no.
```

We use `syntax` as a path prefix for syntactic information, which is stripped off when we pass from general-purpose nodes to ones which are specific to syntactic information. Entity-words are nouns unless there is a stipulation to the contrary, and nouns by default are concrete, count nouns so these values are stated at `NOUN`, the highest node in the hierarchy for nouns. Mass nouns differ from this paradigm in various ways, which we shall gesture towards in this fragment simply by specifying a `no` value for the `count` feature. In other respects they are regular nouns, so `MASS-NOUN` otherwise inherits from `NOUN`. A more complex set of interactions between alternations and syntax will be addressed in Chapter 9.

The 'bracketed optional parts' indicating the 'wood' senses for *ash* and for *beech* were not identical. For *ash*, there was a further specification that the wood was "hard". Regular polysemy is often not entirely regular, and in general it will often be necessary to overrule inherited values, or to add specifications that are not inherited to an inherited sense. This is easily done in DATR. Dictionary definitions comprise genus and differentiae (which might be numbered), and the hardness of ash wood is one of the differentiae, so the addition to the theory is:

```
Ash: <alt wood differentia-1>       == hard.
```

In general, any number of further specifications may be added to an inherited sense in this way.

In LDOCE we have:

> **teak** (a large tree from India, Malaysia and Burma that gives) a very hard yellowish brown wood ...

and this illustrates there is a wood/tree as well as a tree/wood alternation. Lexicographers have not used the same formula for *teak* as for *ash* and *beech*, and this corresponds to the fact that the 'wood' sense is the more salient for the former, the 'tree' sense for the latter. To represent the two patterns as the same would be to throw away a principled distinction made by the lexicographer.[4]

In this fragment, all alternations are represented as directional links and the relation between the two alternations is not expressed. The following code adds *teak*, and the wood/tree alternation, to the fragment.

```
WOOD:<>                 == ENTITY
    <genus>             == wood
    <collocates>        == table desk ENTITY
    <alt tree>          == TREE:<>.

Teak:<>                 == WOOD
    <word>              == teak.
```

8.5 Transitive alternations

> **cherry 1** a small soft fleshy red, yellow or black round fruit ... **2** (the wood of) the tree on which this fruit grows

The definition displays two alternations involving three senses. The primary sense is the fruit. Then there is the tree on which it grows, and then the tree/wood alternation applies to the secondary, 'tree' sense giving the 'wood' sense. The definition might have used nested bracketed-optional-parts, thus:

> **cherry** ((the wood of) a tree that produces) a small soft fleshy red, yellow or black round fruit ...

but for the fact that user-friendliness is an overriding concern of lexicographers, and a recursive metalanguage falls fatally at that hurdle.

The pattern is productive. The 'wood' senses of *pear*, *orange* or *mango* are, like the glass-of sense of *bourbon*, too predictable and rare to be mentioned in LDOCE, yet, in the appropriate, carpenter's yard context, the use of the words to denote kinds of wood will be unexceptional. In DATR, we have

```
FRUIT:<>                    == ENTITY
    <word>                  == fruit
    <collocates>            == eat pick rot
    <genus>                 == fruit
    <alt tree>              == TREE:<>
    <alt tree collocates>   == blossom TREE:<collocates>.

Cherry:<>                   == FRUIT
    <collocates>            == morello FRUIT
    <word>                  == cherry.
```

as shown in Fig. 8.3.

The basic mechanism for 'transitive' alternations is to use as many `<alt x>` prefixes (where x is the identifier for the alternation) as required. Thus

[4] We do not address the vexed question of how to determine which is a more salient, or primary, sense. As far as possible we do as the dictionary suggests.

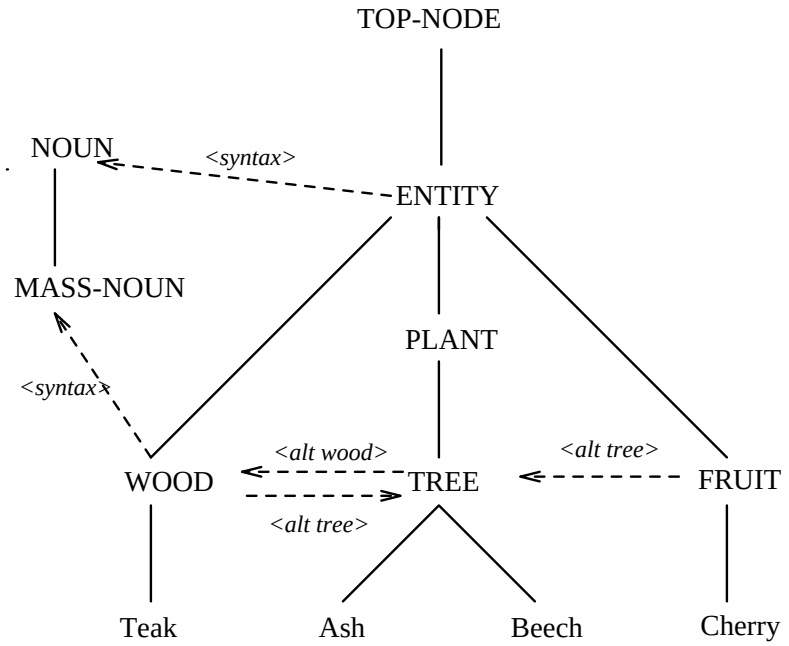


Figure 8.3: Taxonomy showing transitive alternations

```
Cherry:<alt tree alt wood genus>           = wood.
Cherry:<alt tree syntax count>             = yes.
Cherry:<alt tree alt wood syntax count>    = no.
```

There may be any number of `<alt` *x*`>` prefixes, and a query may be redirected any number of times. In a larger DATR lexicon, there would often be several alternations specified at nodes, each time redirecting inheritance to another node and stripping off a prefix. Thus, as the number of `<alt` *x*`>` prefixes grows, so the number of potential usage-types which the theory is describing for the word increases exponentially. A search space of possible uses fans out. All the alternations directly available to the primary sense of the word form a set of possibilities at depth 1. For some of the nodes which the query has been redirected to, further 'one-step' alternations are defined, as the alternation to `WOOD` is defined from `TREE` for *cherry*, and they form the possibilities at depth 2, and so on recursively.

This kind of behaviour is an outcome of the productivity of alternations, combined with the fact that alternations can apply to secondary usage-types as well as to primary ones. In modelling alternations in this way, we are adopting a position similar to that argued in Nunberg (1978) (see section 4.2) and implying that words have an indefinite number of potential senses.

8.6 Trees and plants, cherries and strawberries

We now have an axiom for which the first-pass interpretation is that fruit grows on trees — but not all fruit do. Correspondingly, we have some unwanted theorems. Also, we do not yet have at our disposal the machinery to express relationships between alternations.

Consider the following definition.

strawberry 1 (a plant that grows near the ground and has) a soft red juicy fruit

Here we have a fruit/plant alternation, a more general variant of the fruit/tree pattern. If a `Strawberry` node which inherits by default from `FRUIT` is added to the theory, then, if we query the node with a path commencing `<alt tree>` or `<alt tree alt wood>`, the inheritance follows exactly the same course as for `Cherry`. If there are values for syntax, collocations or anything else which `Cherry` will pick up from higher up the `TREE` or `WOOD` parts of the taxonomy, then `Strawberry` will also pick them up. The theory above, when supplemented with

```
Strawberry:<> == FRUIT
    <word>    == strawberry.
```

has theorems

```
Strawberry:<alt tree genus>                    = tree.
Strawberry:<alt tree syntax count>             = yes.
Strawberry:<alt tree alt wood genus>           = wood.
Strawberry:<alt tree alt wood syntax count>    = no.
```

One direct approach to this problem would be to insert a `TREE-FRUIT` node between nodes for tree-growing fruit and `FRUIT`, and for that to be the point at which paths starting `<alt tree>` were redirected to `TREE`. However, the motivation for such a node is weak. 'Tree-growing fruit' is not used in the dictionary as a genus term, so the node is not sanctioned by the taxonomy described by the LDOCE lexicographers. Also, there seems no reason for introducing a node for tree-growing fruit that does not also apply to distinctively-coloured fruit (and potentially other subsets of fruit), since both give rise to alternations. But then if there were two more intermediate nodes on the default inheritance path for `Cherry` (and other tree-growing, distinctively-coloured, fruit) they would have to be ordered. Either ordering would be unsatisfactory, since tree-growing fruit are not a subset of distinctively-coloured fruit and distinctively-coloured fruit are not a subset of tree-growing fruit.

The strategy adopted here accepts that `Strawberry:<alt tree>` will give unwanted results, but also asserts that only some `<alt` x`>` prefixes at a node receive interpretations, and are thus interesting to query. DATR's default inference ensures that any interesting theory has an infinite number of 'uninteresting' theorems following from it, since, if `<path>` at `Node` evaluates to `val`, then so do `Node:<path foo>`, `Node:<path baz>`[5], and all other paths starting with `path` which are not otherwise defined. While any DATR theory has an infinite set of uninteresting paths, though, some — for example, those including atoms not to be found in any axioms — are uninteresting at all nodes (see also section 9.6.2), while others such as `<alt tree>` are interesting at some nodes and unintersting at others, and the cases must be distinguished.

This is a 'second order' task. Rather than providing information about words directly, we must provide information about DATR paths. This need not be given in the same language or the same theory as the first order information. There are however similar arguments for using DATR in the second order theory as for the first order one, as words of similar meaning tend to share alternations in a pattern dotted with exceptions and subregularities. And since the same basic taxonomy of words holds good for the first and second order information, the mechanism for distinguishing interesting and uninteresting paths at a node will simply be added, in DATR, to the existing DATR fragment.

The mechanism is this. For each node associated with a word, a path called `<altlist>` is defined, and the value for this path is a list of all the alternation-labels for one-step alternations from the node. We add to the theory

```
TOP-NODE:<altlist>              == .
```

[5] Assuming `foo` and `baz` do not otherwise occur in relevant axioms.

```
TREE:<altlist>                  == wood PLANT.
FRUIT:<altlist>                 == "<grows-on>" ENTITY.
WOOD:<altlist>                  == tree WOOD.
FRUIT :<grows-on>               == plant.
Cherry:<grows-on>               == tree.
Strawberry:<>                   == FRUIT.
```

The mechanism employed is as for collecting up lists of collocates. The value is defined as a sequence with two parts, locally-specified items and items specified by inheritance. Where there are no locally specified items, no equations need adding to the code as the inheritance goes through by default. The double-quotes around `<grows-on>` denote DATR's global inheritance. For each DATR query, a 'global' context comprising a node and a path is initialised to the node and path of the original query, and then stored. When a quoted path is encountered, as here, the value is inherited from the stated path at the global-context node (and the global-context path is changed to the one quoted). So, in evaluating `Cherry: <altlist>`, we inherit a sequence at `FRUIT`, the first element of which is `"<grows-on>"`, which is evaluated at the query node, `Cherry`, where `Cherry: <grows-on> == tree.` so the first element of the sequence is `tree`.

The approach enables us to capture, in a limited way, the relation between the 'plant' sense of *strawberry* and the 'tree' sense of *cherry*, and to do this by referring to the facts about cherries and strawberries which will be needed in a full lexical entry independently of considerations relating to alternations. The encyclopedic information is simply stated as the value of the `<grows-on>` path. By default (specified for all fruit at `FRUIT`) fruit grows on plants, and in the case of *cherry* this is overridden locally.[6] The relation between the two alternations comes in in the definition of `<altlist>` at `FRUIT`. The first element of the sequence is `"<grows-on>"`. For `Cherry`, it evaluates to `tree`, for `Strawberry`, to `plant`, so the two alternations fill the same spot in their respective `altlist`s and the fact that the two alternations are in a sense the same has been captured.

A method of access that the dictionary permits, but which has not yet been provided for the DATR fragment, is 'blind listing'. If we look up a word in the dictionary without any further knowledge about its possible primary or other senses, we access a list of all the dictionary senses. As yet, to access information in the DATR theory about any senses but the primary one, we must already know the alternation identifiers. The `altlist` path makes blind listing possible. Blind breadth-first or depth-first search will then generate all the alternations for a word. All the labels for alternations at depth 1 in the search space are listed with *Word*`:<altlist>`. The labels for depth 2 are listed with *Word*`:<alt` *mem1* `altlist>` where *mem1* takes any of the values from depth 1. The labels for depth 3 are listed with *Word*`:<alt` *mem1* `alt` *mem2* `altlist>`, where, for each value of *mem1*, *mem2* takes in turn each value of the `altlist` generated with that value for *mem1*, and so on.

A complete listing will be impossible unless loops are excluded, and impractical even without loops in a substantial lexicon where multiple alternations are defined at a large number of nodes. This is to say no more than that there would be little point in listing all the plausible ways a word might be used, as this is a large set. A listing of all alternations of depth less than some fixed number is more likely to be practical.

The issue of loops is already present in the fragment. The following formulae

```
Ash:<alt wood>
Ash:<alt wood alt tree alt wood>
Ash:<alt wood alt tree alt wood alt tree alt wood>
```

[6] It is arguable that most fruit grow on trees so it would be more concise to set the default to `tree` than to plant. However, setting the default to `plant` serves a further function. Since anything that grows on a tree grows on a plant, the resulting theorems will be true for tree-growing fruit even where the lexical entry omits to mention what the fruit grows on.

all take us to the same location in the inheritance network. It cannot be assumed that the 'looping' path will never be used. A recent innovation of the fish-marketing industry is to reconstitute the flesh of the fish into fish-shapes (and coat it in breadcrumbs for a quick-fried family meal). When the parent asks the child "how many fish do you want?" there is clearly one alternation from animal to food in evidence, and another which re-converts the mass, 'food' sense into a countable, 'fish' sense, yet the characteristics of the breadcrumbed version accessed through `Fish:<alt food alt fish>` are clearly not all the same as those of the original, and we might expect to find specifications such as

```
Fish: <alt food alt fish manufacturer> == Bird's_Eye.
```

even though the default case is that `Fish:<alt food alt fish` *path*`>` inherits from `Fish: <` *path*`>` in a loop-like structure. So apparently looping paths may occasionally give rise to interesting theorems, though usually they will not.[7]

8.7 Polysemy and homonymy in DATR

As discussed in Chapters 1 and 4, 'homonymy' and 'polysemy' are both useful concepts for the description of the lexicon of a language, although it is impossible to cleanly distinguish them. How might this state of affairs be modelled in DATR? A simple model of a DATR lexicon is one in which there is a node for each word or lexeme. Then, a natural treatment for homonymy is to carry information, and make queries, about different words at different nodes. But if polysemy often cannot be distinguished from homonymy, should polysemy not be treated similarly, with distinct polysemous senses each having their own DATR node?

The idea has some appeal, but it would mean that information about a word sense was never defined in the theory unless a node for the sense had explicitly been added. The fragment presented has said nothing explicitly about the 'tree' or 'wood' senses of *cherry* yet it represents facts about their syntax, genus terms and collocations. The senses are predictable and follow from generalisations and should not need explicit mention. It seems likely DATR theories will need to continue working with the homonymy/polysemy distinction, with homonyms treated as distinct nodes, and polysemous senses as sets of node-path pairs with a distinct path prefix. This offers the kind of flexibility required. To turn a sense treated as polysemous into one treated as homonymous is trivial. If we add

```
Cherry/tree: <>       == Cherry:<alt tree>
    <word>            == cherry.
```

to the theory, we now retrieve the same values for theorems of the form

`Cherry2: <`*path*`>`

as for ones of the form

`Cherry: <alt tree` *path*`>`.[8]

All the polysemous senses described in the fragment above can be treated as homonyms, each with a distinct node, in this way.

If distinct nodes were wanted for a pre-defined range of regularly-polysemous or homonymous senses, they could be generated automatically. (The node names would then be automatically generated so we might expect them to take a form `Word1, Word2` etc. rather than the mnemonic `Cherry2`. Since node names are only ever mnemonic, and never appear as values in DATR theorems, this is not an issue.) We may even say that these anonymous node-definitions would be

[7] They will remain more 'visible' then other uninteresting theorems, since any simple mechanism for listing all alternations will include them in its listing.

[8] See section 9.6.1 for a more succinct way of obtaining the same effect.

part of the DATR theory, although they need never actually be typed in. A practical application of the theory may then generate and query anonymous nodes as and when they are needed in the course of language processing. Evans and Gazdar use this technique widely in The DATR Papers. The catch, of course, lies in 'pre-defining' the range of words and `<alt x>` prefixes that the automatic-generation process should apply to. Evans and Gazdar automatically generate nodes for forms such as present participles of verbs: since all verbs have a present participle, there is no risk of overgeneration. Regarding polysemy, unless the cases where automatic node generation is permissible are specified using a mechanism such as `altlist` or '`known`' (see below), there is always a risk of overgeneration.

In general, the cases where there are generalisations to be made regarding polysemy are not the cases which are indeterminate between polysemy and homonymy. So regular polysemy will be treated with path prefixes, unequivocal homonymy with distinct nodes, and for indeterminate cases we may readily switch between the two kinds of treatment.

8.8 Polysemy and metonymy in DATR

As between polysemy and homonymy, so between polysemy and metonymy. The arguments of Chapter 4 show that the two phenomena have a large overlap, and that there is nothing to be gained from treating them as different in kind. The DATR theory does not draw a distinction. There is a population of words, and a population of alternations (in DATR, `<alt x>` prefixes) and an alternation applied to one word will give a polysemous sense, while applied to another, it will give a usage-type better described as an instance of metonymy. The criterion according to which we call one and not the other polysemous will not be based on any difference of kind, generally being a matter of frequency (see section 4.6). In the spirit of Nunberg, the possibilities for how a word may be used will fan out without regard for the lexicographer's finite lists.

The `altlist` mechanism described above is one device for specifying regions of the search space which are more salient than others. All else being equal, we are relatively likely to encounter a word in a usage-type as described on the `altlist` of depth 1 or 2, and less likely to encounter one at depth 3 or 4. The altlist mechanism has been used above to convey negative information: an alternation not appearing on it for a given word will not apply to that word. But to distinguish those word-alternation pairs which occur significantly often from those of negligible frequency, a positive mechanism is wanted. While the 'wood' sense of *mango*, to be queried as

```
Mango:  <alt tree alt wood>
```

is perfectly usable and understandable (in the right context), this is a separate kind of fact to the bald, statistical one that (in England) it almost never occurs. Among all the possibilities defined in the search space, there are many that are never used, many that are very rarely used, and a few that are commonly used. Setting aside the serious questions of what 'almost never' and 'commonly' mean, the minimal distinction between those word senses which are attested and those which are not may be expressed in DATR as follows. If only the 'wood' sense of *ash* and the 'tree' sense of cherry are attested, we add the following statements to the theory:

```
Ash:    <alt wood known>          == yes.
Cherry:  <alt tree known>         == yes.
TOP-NODE:  <known>                == no.
```

The default value for the question, "Is there any empirical evidence for any given alternation having ever occurred in the language?", is "no". The information is stated at `Top-node`. Then for every sense for which there is evidence of its occurrence, a statement overriding the default and stating that the alternation is known is added. If a DATR lexicon were being used in an NLP application, such statements could be added wherever the system concluded that a word

was being used in a previously-unfamiliar sense, so the theory would develop as it 'learnt' from its input.

8.9 Summary

A lexical knowledge base needs inference mechanisms, and a structure which reflects the structure of the lexical knowledge it conveys. DATR is a default inheritance formalism designed specifically for lexical representation. Regular polysemy is one level of structure in the lexicon, about which a desk dictionary provides an ample supply of facts. In this chapter we have examined and formalised the regular polysemy of a very small fragment of English. We have been able to exploit a number of generalisations about the domain to make the theory compact and productive. The formalisation both presents a theory of the operation of regular polysemy in one corner of the lexicon, and is a model for how regular polysemy might be used to structure a lexical knowledge base.

Chapter 9

Formal Lexicography II: Verb Alternations

9.1 Introduction

The previous chapter presented a formalisation of the polysemous behaviour of some nouns. The domain chosen there displayed no very interesting syntactic behaviour and did not address how lexical entries might meet the constraints imposed by a grammar formalism which makes parsing and semantic interpretation possible. This chapter formalises the alternations apparent in a part of the verb lexicon, taking these constraints into consideration.

The kind of phenomenon to be captured is the relation between *bake* in "John is baking the cake", "John is baking", and "The cake is baking".[1] Here are three different usage-types for the verb. It should not be necessary to introduce three different primitives into the lexicon. The alternations are regular and are shared with other verbs, so should be described at some general node in the taxonomy, and inherited.

One part of the exercise is to make the relations between the senses explicit. This is one way in which this chapter is more ambitious than the last. There, no attempt was made to give a complete account of the meaning of 'wood' *ash* in terms of 'tree' *ash*: indeed, the difference between the two cannot be explicated without reference to a large amount of biological and general knowledge, so beyond a superficial analysis involving an *ad hoc* WOOD-OF predicate, it is hard to see what can be done without recourse to the encyclopedia. For transitivity alternations, on the other hand, a full account of the relations between meanings of alternative senses can sometimes be succinctly given.

Much has been written about verb alternations and their syntactic corollaries. Here we will not add to the evidence or construct new theory but simply formalise other people's accounts. Any exercise in formalising a previously informal theory is likely to involve making it more explicit, and also the choice of formalism and formal structure will force certain choices which may have been left open in the informal version. This has been done. The benefits of formalisation are described in section 1.3.1. The formalism will, once again, be DATR. The analyses we shall be formalising are from two articles, Atkins et al. (1986, hereafter AKL) and Levin & Rappoport Hovav (1991, hereafter LR).

AKL is an investigation of the range of alternations between transitive and intransitive forms of verbs . Their particular concern is the difficulties the alternations present to lexicographers.

[1] The morphosyntactic distinctions between, for example, *bake* and *is baking* are not addressed here. Extensive DATR treatments of morphology are provided in various papers in Evans & Gazdar (1990).

Simply stating that a verb such as *bake* can be transitive or intransitive, and then giving example sentences, is not an adequate treatment in a learner's dictionary. How, AKL ask, is the learner to discover that "The girl washed" means the same as "The girl washed herself" rather than "The girl washed something". They show that the problem is a difficult one, treatment in existing learners' dictionaries is inadequate, and that:

> Making the implicit knowledge encoded in a dictionary explicit is only possible in the context of a theory of lexical organisation. Linguists can contribute to lexicography by providing such a theory. (p 61)

The current chapter contributes a formal account of lexical organisation for some of the cases they discuss. AKL list six different alternations that may hold between transitive and intransitive. The three we shall formalise are the ergative ("the cake is baking"), unspecified-object ("John is baking"), and reflexive ("Mary washed").[2]

The second article shares one author with the first and has similar goals. It aims at 'uncovering syntactically relevant components of verb meaning' (p 123). A component of meaning counts as syntactically relevant if all or most of the verbs with that component to their meaning exhibit a particular subcategorisation pattern. We shall follow LR in taking a distinct subcategorisation frame as identifying a distinct usage-type for the verb, so, as with the nouns in the last chapter, we have classes (and subclasses) of verbs classified in terms of components of meaning, with the alternation behaviour following that same classification scheme. The title of the article is 'Wiping the slate clean', and LR present a taxonomy of 'wipe' verbs, 'clean' verbs, and other related groupings. Their taxonomy becomes the basis for the DATR taxonomy.

9.2 HPSG-style lexical entries

The form in which relations between syntax and semantics are to be stated is also a matter on which much has been written. Here, we adopt ideas from unification grammars, particularly HPSG (Pollard & Sag, 1987, 1992). The reader is referred to Shieber (1986) for an introduction to unification-based grammars, and Pollard & Sag (1987) for an account of HPSG. HPSG lexical entries are in a formalism in which syntactic and semantic phenomena are also described, and the formalism provides a unified account of language which can serve as the basis for automatic analysis and generation. Thus the lexical entries are constrained to be structured as they are by the roles they play in the language as a whole.

The lexical entry shown in Fig. 9.1, adapted from Pollard & Sag (1987, p 98), serves to illustrate several points of the structure of HPSG lexical entries which are central to our analysis.[3]

It shows the attribute-value matrix (AVM) notation for the entry. It describes a directed acyclic graph (DAG) or feature structure, and is equivalent to a set of PATR equations (Shieber, 1986). The equivalence is not immediately apparent and the correspondences are spelt out below. As the two notations are equivalent, they will be used interchangeably, with 'AVM' preferred for referring to the whole DAG, and 'PATR path' or 'PATR equation' for referring to its parts.

First, NP [NOM] and NP [ACC] are abbreviated forms for the AVMs in Fig. 9.2 and Fig. 9.3.

[2] AKL's other three alternations relate to issues of aspect and group noun-phrase semantics which we do not address. They are the 'symmetric alternation (as between "Ann met Cathy" and Ann and Cathy met"); the 'characteristic property alternation' (as between "That dog bites people" and "That dog bites"); and the instrumental alternation (between "The scissors won't cut the denim" and "The scissors won't cut").

[3] There are a number of attributes which are crucial to the integrated HPSG account of language but are not needed to describe the alternations we consider. For clarity of presentation, these are omitted. Thus our SYN is HPSG's SYN|LOC, our SEM is their SEM|CONT, and we omit the HEAD attribute. Also, Pollard & Sag (1987) adopt a convention whereby the first item on the subcat list is the most oblique argument. The opposite order is adopted here, as indeed it is by Pollard & Sag (1992). Thus the first item on a verb's subcat list is its subject.

$$\begin{bmatrix} \text{WORD} & \text{bake} \\ \text{SYN} & \begin{bmatrix} \text{MAJ} & \text{V} \\ \text{SUBCAT} & \langle \ \text{NP[NOM]} \ \ \text{SEM} \ \boxed{1} \ , \\ & \quad\ \ \text{NP[ACC]} \ \ \text{SEM} \ \boxed{2} \ \ \rangle \end{bmatrix} \\ \text{SEM} & \begin{bmatrix} \text{RELN} & \text{BAKE} \\ \text{BAKER} & \boxed{1} \\ \text{BAKED} & \boxed{2} \end{bmatrix} \end{bmatrix}$$

Figure 9.1: HPSG-like AVM for *bake*

$$\begin{bmatrix} \text{SYN} & \begin{bmatrix} \text{MAJ} & \text{N} \\ \text{CASE} & \text{NOM} \end{bmatrix} \end{bmatrix}$$

Figure 9.2: AVM for NP [NOM]

$$\begin{bmatrix} \text{SYN} & \begin{bmatrix} \text{MAJ} & \text{N} \\ \text{CASE} & \text{ACC} \end{bmatrix} \end{bmatrix}$$

Figure 9.3: AVM for NP [ACC]

Angle brackets are used for lists. List members are separated by commas and the components comprising a list member, here NP[NOM] and SEM $\boxed{1}$ in the first case and NP[ACC] and SEM $\boxed{2}$ in the second, are to be understood conjunctively, each being a partial description of the list member. To spell the lists out as feature structures, we adopt the standard technique (Shieber, 1986, p 29) of treating the list as a pair comprising a FIRST and a REST, where the FIRST is the first item of the list and the REST is a list comprising all but the first element. The second element is then the FIRST of the REST, the third element, the FIRST of the REST of the REST, and so on down the list until the REST value is a special symbol — we shall use NIL — which marks the end of the list. Thus ⟨a b c⟩ becomes the AVM in Fig. 9.4.

$$\begin{bmatrix} \text{FIRST} & \text{a} \\ \text{REST} & \begin{bmatrix} \text{FIRST} & \text{b} \\ \text{REST} & \begin{bmatrix} \text{FIRST} & \text{c} \\ \text{REST} & \text{NIL} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Figure 9.4: List structure

The rewritten, abbreviation-free AVM is shown in Fig. 9.5.

The DAG is drawn in Fig. 9.6. Arcs are labelled by attributes and some terminals, with values. The re-entrancies marked by the boxed variables in the matrix notation are illustrated directly, by arcs leading to the same node.

Finally, the thesis is not concerned with spoken language, and the formalisation in this chapter uses the spelt form of a word simply as an identifier, so WORD has been substituted for PHON.

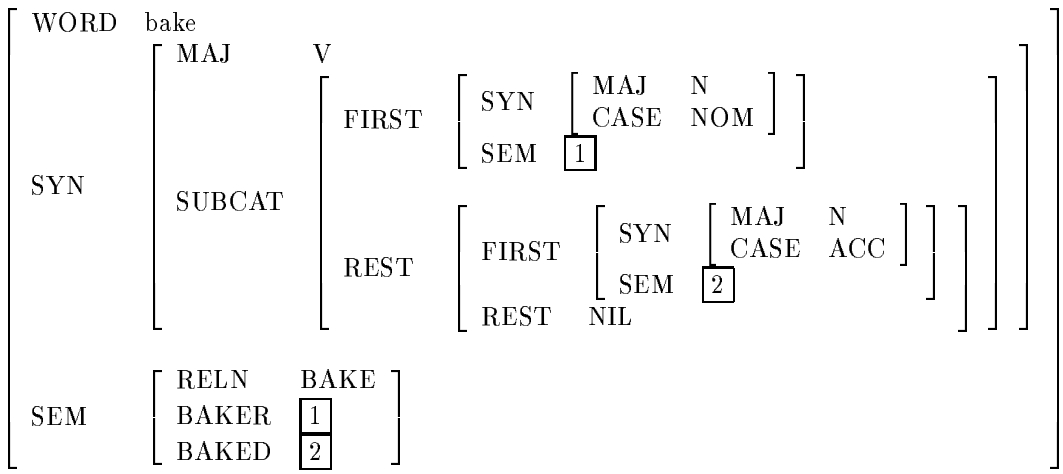


Figure 9.5: AVM for transitive *bake.*

The matrix is also equivalent to the following set of equations written in PATR:

```
Bake: <WORD> = bake
      <SYN MAJ> = V
      <SYN SUBCAT FIRST SYN MAJ> = N
      <SYN SUBCAT FIRST SYN CASE> = NOM
      <SYN SUBCAT FIRST SEM> = <SEM BAKER>
      <SYN SUBCAT REST FIRST SYN MAJ> = N
      <SYN SUBCAT REST FIRST SYN CASE> = ACC
      <SYN SUBCAT REST FIRST SEM> = <SEM BAKED>
      <SYN SUBCAT REST REST> = NIL
      <SEM RELN> = BAKE
```

Where the AVM contains a pair of boxed variables, the two AVM paths involved are contracted into one PATR equation. The equals sign for these re-entrancies has a different semantics to the basic case where the item on the right hand side of the equation is a value. It entails that, for any attribute lists A, B and C and value V, if we have A = B and AC = V (where AC is the concatenation of A and C), then we have a theorem BC = V.

Figure 9.6: DAG for transitive *bake*.

9.3 Alternations as manipulations of AVMs

9.3.1 *Bake* alternations

The difference between transitive and ergative *bake* can now be represented. Part of the goal of the chapter is to represent non-base senses without introducing new primitives, so, if we take the two-place relation BAKE as primitive, we want that rather than a new primitive to occur in the lexical entry for ergative *bake*. And this is straightforward to state. The same relation holds, but now it holds between an unspecified 'BAKER' and a 'BAKED' which is the subject of the sentence. The unspecified role filler is not 'bound' to a complement (as we shall call any item on the SUBCAT list). Rather, in predicate calculus terms, it is existentially quantified (EX-Q), so for ergative *bake* we have:

Bake: <SEM BAKED BINDING> = EX-Q

The ergative form is intransitive so will have only one item, its subject, on its SUBCAT list and the SEM of that item will unify with the BAKED, so the AVM for ergative *bake* will be as in Fig. 9.7. (The NP abbreviations and list notation have been re-introduced to enhance readability.)

$$\begin{bmatrix} \text{WORD} & \text{bake} \\ \text{SYN} & \begin{bmatrix} \text{MAJ} & \text{V} \\ \text{SUBCAT} & \langle\ \text{NP[NOM]} \quad \text{SEM } \boxed{1}\ \rangle \end{bmatrix} \\ \text{SEM} & \begin{bmatrix} \text{RELN} & \text{BAKE} \\ \text{BAKER} & \text{BINDING EX-Q} \\ \text{BAKED} & \boxed{1} \end{bmatrix} \end{bmatrix}$$

Figure 9.7: AVM for ergative *bake*.

For the *bake* in "John was baking", the subject is matched to the BAKER and it is the BAKED which is unspecified, so existentially quantified, as in Fig. 9.8.

$$\begin{bmatrix} \text{WORD} & \text{bake} \\ \text{SYN} & \begin{bmatrix} \text{MAJ} & \text{V} \\ \text{SUBCAT} & \langle\ \text{NP[NOM] SEM } \boxed{1}\ \rangle \end{bmatrix} \\ \text{SEM} & \begin{bmatrix} \text{RELN} & \text{BAKE} \\ \text{BAKER} & \boxed{1} \\ \text{BAKED} & [\text{BINDING EX-Q}] \end{bmatrix} \end{bmatrix}$$

Figure 9.8: AVM for unspecified-object *bake*.

9.3.2 *Clear* alternations

As LR observe, *clear* has at least the following subcategorisation frames: "clear X off Y" ("clear dishes off the table"), "clear Y of X" ("clear the table of dishes"), and simple transitive "clear Y" ("clear the table"). LR treat the simple transitive with an optional 'of' complement as basic. It is not obvious how the notion of 'optional complement' is best formalised. Predicates in logical languages do not normally have an optional number of arguments. One approach to modelling

optional arguments is for two predicates to have the same name but different arity, with the lower-arity predicate being defined as the higher-arity predicate with one of its arguments set, and this is the strategy we adopt. The argument is 'set' through existential quantification, so the alternation is as for the relation between transitive and unspecified-object *bake*. Thus we shall take the base form to be the "clear X of Y" one, represented as an AVM in Fig. 9.9.

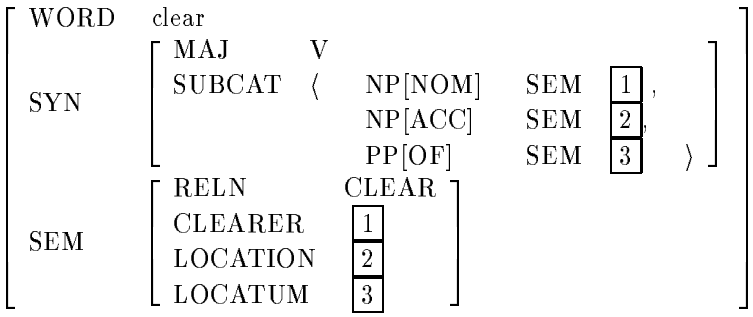


Figure 9.9: AVM for ditransitive 'of' *clear*.

A prepositional phrase abbreviation, PP[OF], has been introduced to stand for the AVM in Fig. 9.10. The PFORM feature is borrowed from GPSG (Gazdar, Klein, Pullum, & Sag, 1985) and

Figure 9.10: AVM for PP[OF]

takes the value of the preposition introducing the PP. PP[FROM], PP[AT] *etc.* may be constructed in the same way. LOCATION and LOCATUM are terms used by LR for "the table" and "dishes" respectively in both "cleared the table of dishes" and "cleared dishes from the table". There is a further complication here. "Clear X from Y" and "Clear X off Y" are not treated by LR as distinct uses of *clear*. They set aside the distinctions in meaning that accompany the choice of preposition, saying:

> We use "FROM" to represent the set of prepositions used to indicate sources in English. (p 127)

We shall follow their analysis and use PP[FROM] similarly. (The difficulties are taken up in section 9.8.)

The simple transitive has the entry in Fig. 9.11. For the "clear X from Y" variant, we first change the preposition from *of* to *from*, and then swap the mappings from SUBCAT list to arguments as in Fig. 9.12. "The skies cleared" illustrates a further possibility for *clear*. The simple transitive form can undergo the ergative alternation we saw above, to give the AVM in Fig. 9.13.

As we saw in the last chapter, alternations can apply to forms that are themselves products of alternations, and this one will be a two-step transformation, first from the 'of' variant to the simple transitive, and then to the ergative.

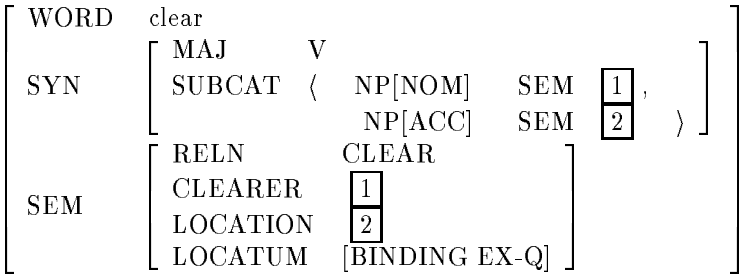


Figure 9.11: AVM for transitive *clear.*

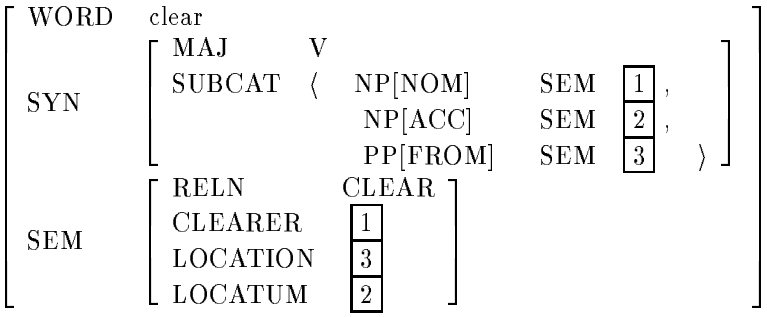


Figure 9.12: AVM for ditransitive 'from' *clear.*

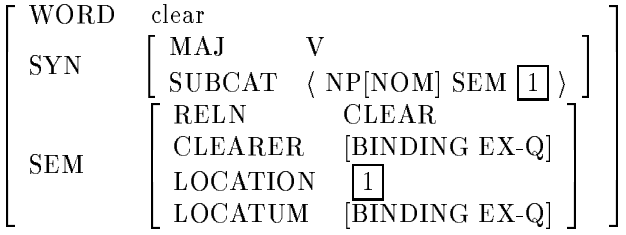


Figure 9.13: AVM for ergative *clear.*

9.3.3 Lambda calculus and AVMs

Lambda calculus provides a notation for making explicit the relations between the ordered arguments of a function and the term defining the function. Thus, we might give the following semantics for transitive *bake*:

$$\lambda y \lambda x BAKE/2(BAKER : x, BAKED : y).$$

The λy tells us that the argument that this function combines with first is y, the BAKED. The argument might be *BREAD*:

$$\lambda y \lambda x BAKE/2(BAKER : x, BAKED : y)(BREAD)$$

and then the substitutions for the lambda-term, 'lambda-reduction', will give

$$\lambda x BAKE/2(BAKER : x, BAKED : BREAD)$$

Setting aside agreement, quantification, tense and aspect and assuming the obvious mappings between words of English and upper-case semantic terms, this is a suitable semantics for the verb phrase "bakes bread". (For an introduction to lambda calculus see Chierchia & McConnell-Ginet (1990, chapter 7).) This VP semantics will then combine with its next argument, say $JOHN$

$$\lambda x BAKE/2(BAKER : x, BAKED : BREAD)(JOHN)$$

in a similar way to give

$$BAKE/2(BAKER : JOHN, BAKED : BREAD)$$

as semantics for "John bakes bread".

The information that lambda terms are used to express is:

- Which items in the semantics of the verb relate to the arguments the verb takes (where the verb is considered a function);
- The order in which arguments are supplied to the functor.

An AVM analysis also needs to specify these things, but does not need lambda terms. The reentrancies between the subcat list and semantic arguments fill the first role. Provided we take the subcat list to display the order in which arguments are supplied to the verbal functor (which is the obvious rationale for having a subcat list rather than a subcat set and which will, by default, correspond to the linear ordering of constituents in a sentence), then the order of the subcat list fills the second role. So to translate between the semantics and syntax-semantics mappings in an AVM, as presented above, and lambda calculus formulae is a simple matter. Ergative and unspecified-object *bake* become

$$\lambda y \exists x BAKE/2(BAKER : x, BAKED : y)$$

and

$$\lambda x \exists y BAKE/2(BAKER : x, BAKED : y)$$

The matter is given a fuller discussion in Moore (1989).

The lambda notation also makes apparent that 'λ' and '$\exists$' are exclusive alternatives regarding how a variable may be bound, a fact to be exploited below.

9.3.4 *Melt* alternations as manipulations of AVMs

With *bake* and *clear*, we were able to represent the extended senses directly in terms of the same predicate that applied in the base sense. But, for *melt*, the intransitive ("The ice melted") is basic and the transitive ("Maria melted the ice") is extended, and it is not possible to define the extended sense directly in terms of the basic.

The transitive can be paraphrased using *cause*, "Maria caused the ice to melt"; the alternation is called the 'causative'. It is clearly closely related to the ergative, and it would be possible to treat the transitive form as basic, with the ergative alternation applying. That route has not been followed for two reasons. Firstly, *melt* is a member of a class of physical-process verbs, also including *evaporate, freeze, dissolve, sublime* and *coalesce*. They all clearly have intransitive senses. They all might, in the right setting, be used transitively, but in cases such as *coalesce* the transitive is not a standard use and it would patently be inappropriate for it to be treated as a base form. If we are to stand by the intuition that these verbs form a class, and all participate in the same alternation, then all must have an intransitive base form.

Secondly, transitive *melt* introduces an aspect of meaning, call it CAUSE, which is not in any sense present in the intransitive. For *bake*, CAUSE is already a component of the meaning, whether or not the verb is being used ergatively. A default entailment of CAUSE is that its first argument, the CAUSER, has proto-agent properties (Dowty, 1991). If intransitive *melt* were treated like ergative *bake*, CAUSE would be a component of the meaning of intransitive *melt*. Its semantics would have an existentially quantified MELTER argument, which would be a CAUSER and which we would expect to have agent-like properties. In ergative uses of *bake*, the baking scenario still includes an agent who is doing the baking and fills the BAKER role, even though they are not mentioned. (We concern ourselves here only with cooking *bake*, not "The stones baked in the sun" and other usage-types where *bake* is behaving as a physical process verb.) In "The ice melted" there is usually no agent involved. While it might always be possible to assign a filler to the MELTER slot, perhaps "the hot temperature" or "the warm climate", they do not fit readily into the agent, CAUSER role. So we do not treat causatives as ergatives.

A standard analysis of causatives after Dowty (1979) as presented by Chierchia & McConnell-Ginet (1990, chapter 8), is

$$\lambda y \lambda x MELT/2(x, y) = \lambda y \lambda x CAUSE(x, MELT/1(y)).$$

The semantics of the causative has the predicate CAUSE, with MELT/1 re-appearing as its second argument. In addition to intransitive *melt* as shown in Fig. 9.14 we have causative *melt* as shown in Fig. 9.15.

$$\begin{bmatrix} \text{WORD} & \text{melt} \\ \text{SYN} & \begin{bmatrix} \text{MAJ} & \text{V} \\ \text{SUBCAT} & \langle \text{NP[NOM] SEM } \boxed{1} \rangle \end{bmatrix} \\ \text{SEM} & \begin{bmatrix} \text{RELN} & \text{MELT/1} \\ \text{MELTED} & \boxed{1} \end{bmatrix} \end{bmatrix}$$

Figure 9.14: AVM for intransitive *melt*.

9.4 Translations into DATR

This chapter has two goals: first, to demonstrate how non-basic senses could be expressed in terms of basic senses, and second, to demonstrate the inheritance structure for these alternations.

We have seen how the former may be addressed: let us move on to the latter.

We now need to translate AVMs into DATR. As far as possible, for each PATR equation we shall have a DATR equation which looks very similar. Let us return to *bake*. We translate it as follows.

```
Bake:<word> = bake
     <syn maj> = v
     <syn subcat fi syn maj> = n
     <syn subcat fi syn case> = nom
     <syn subcat re fi syn maj> = n
     <syn subcat re fi syn case> = acc
     <syn subcat re re> = nil
     <sem reln> = bake
     <syn subcat fi sem binding> = v1
     <sem baker binding> == v1
     <syn subcat re fi sem binding> == v2
     <sem baked binding> == v2.
```

DATR paths must be associated with nodes, so a node for the paths to be located at has been introduced. FIRST and REST have been shortened to `fi` and `re`. Upper case has been changed to lower case and PATR or AVM paths have become DATR paths.

DATR is not a unification formalism, and all the theory will do in relation to re-entrancies will be mark them with matched pairs of variables. Another module working on DATR output will be needed to interpret the matched pairs as re-entrancies. We introduce the feature `binding` for the variables to be the value of. If there were no such feature, so we had:

```
Bake: <syn subcat fi sem>    == v1
      <sem baker>            == v1.
```

then the inheritance of values for other paths starting `<syn subcat fi sem>` or `<sem baker>` from points above `Bake` in the inheritance hierarchy would be overridden. Also, as we shall see, the `binding` feature makes it possible to use the fact that a semantic argument has an existential-quantification binding to override the default that it is bound to a complement. The only kind of re-entrancy which occurs in the following fragment serves to unify complements with semantic arguments occurring along paths starting SEM. Thus all re-entrancy will be to the first second, third *etc.* member of the subcat list, or complement. The atoms denoting these positions will always be `v1, v2, v3` *etc.*, respectively.

$$\begin{bmatrix} \text{WORD} & \text{melt} \\ \text{SYN} & \begin{bmatrix} \text{MAJ} & \text{V} \\ \text{SUBCAT} & \langle \quad \text{NP[NOM]} \quad \text{SEM} \quad \boxed{1}, \\ & \quad\quad \text{NP[ACC]} \quad \text{SEM} \quad \boxed{2} \quad \rangle \end{bmatrix} \\ \text{SEM} & \begin{bmatrix} \text{RELN} & \text{CAUSE} \\ \text{CAUSER} & \boxed{1} \\ \text{CAUSED} & \begin{bmatrix} \text{RELN} & \text{MELT/1} \\ \text{MELTED} & \boxed{2} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Figure 9.15: AVM for causative *melt*.

9.4.1 Predicates and argument lists

In an inheritance lexicon we often wish to state generalisations about the holder of a particular semantic role across a number of verbs, for example that BAKERs, COOKERs and FRYERs are all to be unified with the subjects of the transitive forms of their respective verbs. If we use verb-specific names such as BAKER for the relations, the generalisation cannot be stated. We shall instead use an argument list, `args`, with members accessed through the `re fi` mechanism. This does not represent a change in the semantics: the first member of the argument list of *bake* will continue to be the BAKER whatever lexical entry it occurs in. But it makes it possible to express generalisations.

Also, we shall use PRED for the predicate rather than RELN, which is used in HPSG. All predicates will have a '/n' suffix where n is the arity of the predicate. The suffixes will not play a role in DATR inheritance but will help the user to keep track of derivations and will also serve to distinguish between words (which are part of English) from predicates (part of a semantic representation language).

9.5 A taxonomy of verb classes

We are now in a position to describe the skeleton of the inheritance structure. At the top of the tree is `WORD-CLASS`, from where all lexical entries eventually inherit by default. At the next level down is `VERB`. All verbs have a subject, and by default this unifies with the first item on the `args` list. There will be no call for an `INTRANSITIVE` node because all the positive information that might be stated there is true of all verbs so can be stated at the `VERB` node, and the negative information that intransitive verbs do not have direct objects is expressed using the list termination mechanism described below. `TRANSITIVE` inherits from `VERB`, adding the default binding between second complement and second argument.

```
WORD-CLASS: <lexical> == true.

VERB: <> == WORD-CLASS
    <syn maj> == verb
    <syn subcat fi sem binding> == UNI:<first>
    <sem args fi binding> == UNI:<first>.

TRANSITIVE: <> == VERB
    <syn subcat re fi sem binding> == UNI:<second>
    <sem args re fi binding> == UNI:<second>.
```

Rather than having the binding variables `v1, v2, v3` *etc.* scattered around the theory, they are, for tidiness, gathered together at a 'unification' node:

```
UNI:<first> == v1
    <second> == v2
    <third> == v3.
```

9.5.1 List termination in DATR

List termination involves a measure of ingenuity.[4] At a first pass we may terminate lists with:

```
VERB: <syn subcat re> == nil
    <sem args re> == nil.
```

[4] The treatment is due to Roger Evans.

and:

```
TRANSITIVE: <syn subcat re re> == nil
    <sem args re re> == nil.
```

There are two problems with this. Firstly, these paths can be extended indefinitely and the path will still be `nil`. Hence we have theorems such as:

```
VERB: <syn subcat re re re re re> = nil.
TRANSITIVE: <sem args re re re> = nil.
```

These should both be undefined. Secondly, as `TRANSITIVE` inherits from `VERB` we also have:

```
TRANSITIVE: <syn subcat re> == nil.
TRANSITIVE: <sem args re> == nil.
```

Again, these should both be undefined. `nil` must occur where a list terminates and nowhere else. [5] The first problem is resolved by replacing `nil` with the node-path pair `NIL:<>`, and defining `NIL` as follows:

```
NIL:<> == nil
    <fi> == UNDEF
    <re> == UNDEF.
```

The two kinds of extensions which can follow `re` are `fi` and `re`. Nothing else will give an interesting query. We exploit this fact and say that if the path containing `re` and inheriting from `NIL` has further elements, the first of them will be either `re` or `fi`, and in either case the value is not defined. (The eponymous `UNDEF` is undefined. Queries directed to it always fail.) The only other case is where the path is not further extended, and in that case the value is `nil`.

To solve the second problem, we divert the whole list to inherit from a node for a default list member, `ARG` for the semantics list and `COMP` for the syntax list:

```
VERB: <sem args> == ARG: <>
    <syn subcat> == COMP:<>.
    <syn subcat fi syn case> == nom
    <sem args fi semfeats> == AGENT:<>.

TRANSITIVE: <syn subcat re> == COMP:<>
    <sem args re> == ARG:<>.

ARG:<fi semfeats> == PATIENT:<>
    <re> == NIL:<>.

COMP:<fi syn> == NP:<>
    <re> == NIL:<>.
```

[5] At least not as the value of a theorem we may have reason to query. The notion of an 'interesting query' is formalised in section 9.6.2. It will remain true that

```
TRANSITIVE: <syn subcat re re foo> == nil.
```

but this is not an interesting query.

Now the only **re**-terminated paths to have values defined will be the ones that should, and no other interesting queries will have **nil** as their value. The **COMP** and **ARG** nodes are also a location for default information about list members. Complements are, by default, accusative noun phrases. Following Dowty (1991), we have a default expectation that subjects will have 'proto-agent' semantic features and objects, 'proto-patient' ones. **AGENT** and **PATIENT** set up the expected values for four of the characteristics Dowty discusses.[6]

```
NP:<maj> == n
   <case> == acc.

AGENT:<volition> == yes
    <sentient> == yes.

PATIENT:<changes-state> == yes
    <causally-affected> == yes.
```

The default accusative case and proto-patient semantic features must be overridden in the case of the subject:

```
VERB:<syn subcat fi syn case> == nom
     <sem args fi semfeats> == AGENT:<>.
```

It is to be noted that the motivation for the FIRST-REST treatment of lists is not to be found within this lexicon. Nothing in our analysis uses the indefinite extensibility that the treatment provides, and the lexicon could simply use **first**, **second** and **third** rather than **fi**, **re fi** and **re re fi**, and omit terminators. However the fragment is designed to supply an HPSG lexicon, and HPSG makes use of lists which are extensible and terminated. The goal of our treatment of lists is to provide lexical entries which are directly usable as HPSG lexical entries.

9.5.2 *Bake* and cooking verbs

To this skeleton, we can begin to add some smaller classes based on meanings, and once we introduce them we can start expressing generalisations about alternation behaviour. As in the last chapter, alternations will be referenced with an **alt** *x* prefix, and all-upper-case nodenames are used for nodes for classes of words, such as the 'clear' verbs, while lexical nodes have only initial letters capitalised.

```
Bake:<> == COOKING-VB
     <word> == bake
     <sem pred> == bake /2.

COOKING-VB:<> == C-OF-S
    <sem args re fi semfeats edible> == yes.

C-OF-S:<> == TRANSITIVE
    <alt erg> == PHYS-PROC:<>
    <alt erg sem> == "<sem>"
    <alt erg sem args fi binding>  == ex-q
    <alt erg sem args re fi binding> == UNI:<first>.
```

[6] Clearly, we have here only the most rudimentary of proto-agent and proto-patient descriptions. In a fuller theory, different individual verbs and different classes of verbs would override any of a number of features which, together, defined a proto-agent or proto-patient.

Bake is a cooking verb, and cooking verbs are, in the base case, transitive change-of-state verbs. Thus `Bake` inherits, by default, from `COOKING-VB` which inherits from `C-OF-S` (for 'change of state') and then from `TRANSITIVE`, so acquiring the default specifications for semantic features for its subject and object, and the re-entrancies between subject and first argument, and object and second argument. The DATR fragment now represents all the information in the DATR lexical entry for *bake* presented above, and proto-agent and proto-patient specifications in addition.

Many change-of-state transitives, *bake* amongst them, can undergo the ergative alternation to become change-of-state intransitives, or 'physical process' verbs. Queries regarding the ergative forms will have paths beginning `alt erg`. The semantics of the ergative will be as for the base form. The third line of the `C-OF-S` node tells us, with the double-quotes, to inherit the ergative's semantics from the semantics of the global-context node, which will be the node for the base form of the verb. The two further specifications are that the first argument is existentially quantified, and the second unifies with the first complement via `UNI:<first>`.

In all other matters, the ergative form is diverted to inherit from a node for physical-process intransitives:

```
PHYS-PROC:<> ==  VERB
    <sem args fi semfeats> == PATIENT:<>.
```

The first semantic argument of a physical-process intransitive typically has proto-patient semantic features and otherwise inherits from `VERB`. We now have almost all the information needed to build the lexical entry for ergative *bake*. One item we do not yet have is the intuitively obvious fact that the `word` for the alternate form is the `word` for the original. This is true by definition for all alternate forms. All alternate forms will eventually have all their `alt` *x* prefixes stripped and inherit from `WORD-CLASS` at the top of the tree. So we add the following line:

```
WORD-CLASS:<word> == "<word>".
```

Now all alternate forms will inherit their `word` from the `word` at the global context node, which will always be the node for the base form.

Many cooking verbs undergo the 'unspecified object' alternation, for which we shall use the label `unspec`. All information relating to this form is gathered at an `UNSPEC` node:

```
UNSPEC:<>   ==  VERB
    <sem> == "<sem>"
    <sem args re fi binding>  == ex-q.
```

This simply states that the form is a standard intransitive, with the semantics of the base form except that the second argument is existentially quantified. Cooking verbs with `alt unspec` prefixes are diverted here with:

```
COOKING-VB:<alt unspec>  == UNSPEC:<>.
```

9.5.3 *Melt* and causatives

Melt is a physical-process verb which has a causative form. The ergative alternation led from `C-OF-S` to `PHYS-PROC`. This makes a similar journey in the opposite direction, from `PHYS-PROC` to `CAUSE` and then `TRANSITIVE`. The alternation label is `cause`.

```
Melt:<> == PHYS-PROC
    <sem pred> == melt /1
    <word> == melt.
```

```
PHYS-PROC:<> ==  VERB
     <alt cause> == CAUSE:<>
     <alt cause sem args re fi> == "<sem>"
     <alt cause sem args re fi args fi binding> == UNI:<second>.

CAUSE:<>   ==  TRANSITIVE
    <sem pred> == cause /2.
```

Causative *melt*, with the `alt cause` prefix, is a regular verb of causing, and inherits its syntax and most of its semantics including the predicate `cause/2` from `CAUSE`. Its first argument will have the usual characteristics of a CAUSER, and its second, the predicate-argument structure of the base form of the verb.

9.5.4 *Wash* and reflexives

AKL list six ways in which transitives can have intransitive senses: we undertook to formalise three of them. Ergatives and unspecified-objects we have seen: the third is reflexives, as in the relation between transitive and intransitive *wash* ("I washed something" and "I washed"). The alternation applies to a class of verbs about 'caring', *wash, shave, dress,* etc. It is very simply expressed in DATR:

```
CARE-VB:<> == TRANSITIVE
    <alt refl> == REFL:<>.

REFL:<>   ==  VERB
    <sem> == "<sem>"
    <sem args re fi binding>  == UNI:<first>.

Wash:<> == CARE-VB
    <pred> == wash /2
    <word> == wash.
```

Wash is a verb of caring, and we expect these to undergo the reflexive alternation, in which the second semantic argument unifies with the subject and the first argument. The analysis implies that, in "John washed", John is both agent and patient, and proto-agent and proto-patient characteristics are unified with each other. If there were features of agent-hood that clashed with features of patient-hood, this would be a problem since the DATR lexical entries would be implying that two DAGs were unified, when any attempt to unify them would fail. However, Dowty's lists of contributing properties for the agent and patient proto-roles do not clash so the intuition that John is both agent and patient, with the properties of both, stands.

9.6 Query-only nodes and a query grammar

9.6.1 Query-only nodes

The information needed to form the full-listing lexical entry will be accessed by querying the DATR `Bake` node, and prefixing paths with `alt erg`. The interface will be neater if queries can be made at a distinct node and the interfacing system or user does not need to know about the essentially theory-internal mechanism of the `alt` *x* prefixes. If a node specific to ergative *bake* is required, it is:

```
Bake/erg:  <> == "Bake:<alt erg>".
```

This says not only that ergative *bake* inherits all its values from the base form of *bake* with `alt erg` prefixed, but also that the global context is changed so that the stored node is `Bake`, rather than `Bake/erg`, so the behaviour is exactly as if the query had been made at `Bake`. For every alternation, it is technically a trivial matter to produce such an alternation-specific node and it could be done automatically (though see section 8.7 for the pitfalls); in this chapter, such nodes will be freely created (with the *word/alternation-label* naming convention as here), and queried, without further discussion.

9.6.2 Which paths to query? A query grammar

A system using the lexicon will need some way of knowing which paths to query to build a lexical entry. Kilbury, Naerger, & Renz (1991) study this question, and argue that there is a major drawback with an approach to a DATR lexicon in which DATR paths correspond directly to PATR paths. They say all equations will effectively need specifying twice over: once, in order that the interfacing system knows which paths to query for a given lexical entry, and again, where the value that the path takes is stated.

It is necessary to state both kinds of information, but this need not involve redundancy. As shown below, both kinds of information can be stated in a non-redundant manner, and without there being any obvious overlap between the two descriptions.

In a theory of the kind presented in this chapter, there is a substantial but finite number of paths which it might be of interest to query at a given node. Kilbury et al. mention the possibility of 'typing', and this technique has been adopted for the purposes of generating the paths for the `show` declaration. To 'type' a feature structure is to specify, for each feature, what path extensions may be defined for it and what the values for those extensions may range across (Carpenter, 1990). In our DATR theory, it is apparent that interesting paths are a very small subset of all the paths that might be formed by combining the atoms of the theory. Many short atom sequences will never even be subsequences of interesting paths. Thus `syn` will never follow `sem`, nor will anything except `re` ever occur twice in succession, nor will `maj, case, lexical, word, pred` or others occur anywhere other than as the last element of a path. We can specify, for each element, what may follow it, to give a finite state transition network describing all the paths it would ever be meaningful to query. The basic structure of the FSTN is drawn in Fig. 9.16.

The formalism for representing the FSTN is taken from Hopcroft & Ullman (1979). A legal DATR path in the grammar is one that starts at 'Start', 'consumes' the label on the arc with each transition it makes, and terminates at an 'end' state, drawn as a double circle. (As mentioned above, `alt` *x* prefixes are considered theory-internal. They do not feature here. The nodes it will be appropriate to query will be either basic lexical nodes such as `Bake` and `Melt`, or, for non-base forms, the ones generated as needed with *word/alternation-label* e.g. `Bake/erg`, `Melt/cause`.)

As this FSTN is iterative it produces an infinite number of paths. In this fragment lists are never more than three items long and there are never more than two lists in an interesting query. Taking this and other constraints into account, the FSTN can be modified so that it accepts only those paths which will give interesting theorems at some node or other in this fragment. Then the complete set of paths it accepts can be generated. This was done, and 95 paths were generated. With all those paths loaded and DATR in a node-querying mode so that, when given a node, it attempts to evaluate all loaded paths at that node (and prints nothing where a query fails),[7] the simple node-query:

```
Bake/erg?
```

[7] The Sussex DATR implementation is supplemented with devices for 'node-querying' and the `show` declaration, which 'loads' a set of paths in the manner required here. See Jenkins (1990).

Figure 9.16: FSTN for interesting paths

gives the following output:

```
Bake: <lexical> = true.
Bake: <word> = bake.
Bake: <syn maj> = verb.
Bake: <syn subcat fi syn maj> = n.
Bake: <syn subcat fi syn case> = nom.
Bake: <syn subcat fi sem binding> = v1.
Bake: <syn subcat re fi syn maj> = n.
Bake: <syn subcat re fi syn case> = acc.
Bake: <syn subcat re fi sem binding> = v2.
Bake: <syn subcat re re> = nil.
Bake: <sem pred> = bake /2.
Bake: <sem args fi binding> = v1.
Bake: <sem args fi semfeats volition> = yes.
Bake: <sem args fi semfeats sentient> = yes.
Bake: <sem args re fi binding> = v2.
Bake: <sem args re fi semfeats changes-state> = yes.
Bake: <sem args re fi semfeats causally-affected> = yes.
Bake: <sem args re re> = nil.
```

9.6.3 Predicate names

Lexical entries can be kept smaller if we observe the redundancy associated with the `word` and `pred` paths. The predicate is built, in the default case, by taking the word and adding `/1` for an intransitive and `/2` for a transitive. The default applies to all forms seen so far except causative change-of-states. So we add the following lines:

```
VERB:       <sem pred> == "<word>" /1.
TRANSITIVE: <sem pred> == "<word>" /2.
```

It is no longer necessary to state a `sem pred` equation at any lexical node.[8]

Predicates will be DATR sequences. This is not, of course, to say they should be treated as anything other than atoms by a parser or other program which uses the output of the lexicon, and it may be desirable to delete the white space between the elements of these DATR sequences prior to using them as DAG values, but that is an issue for the client system rather than the DATR lexicon.

It is to be noted that, while this is a saving on the amount of typing involved in inputting new lexical entries, it is not of theoretical interest. No generalisation relating words, which are linguistic entities, with predicates, which are semantic ones, has been captured. The generalisation is a trivial one relating words to *names* of predicates.

9.7 The LR verb class taxonomy

From a starting point in which 'wipe' verbs and 'clear' verbs are seen as two subspecies of 'remove' verbs, LR's analysis moves on to find the 'clear's to be a subspecies of change-of-states and the 'wipe's a subspecies of activity-verbs. LR say that both kinds have extended senses which are subspecies of 'remove's.

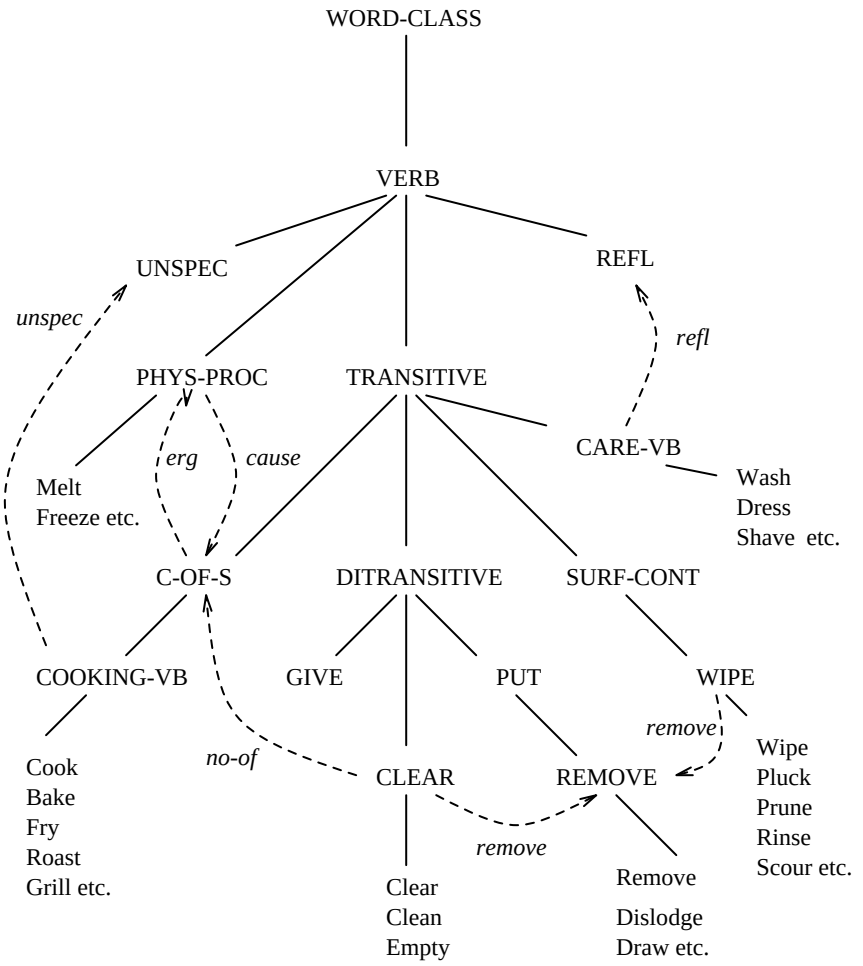
The analysis presented here does not address the 'Vendler-classification' of verbs into states, activities, accomplishments and achievements (see, *inter alia*, Vendler (1967), Dowty (1979), Pustejovsky (1991b)), and that classification seems orthogonal to our primary concerns, so, rather than introduce an 'activities' category into our taxonomy, we shall take up another comment of LR's and treat 'wipe' verbs as verbs of surface-contact. They also note a subspecies of 'wipe' verbs, the denominal 'instrument' verbs such as *vacuum, mop, hoover, brush*.

We may bring these observations together with the analysis of the previous sections to draw up the taxonomy in Fig. 9.17.

Alternations `unspec, cause, erg` and `refl` are dealt with above. The node `DITRANSITIVE` relates to a permissive reading in which any verb which subcategorises for three items is 'ditransitive', irrespective of the syntactic form of the items subcategorised for. The node follows the same pattern as `TRANSITIVE`:

```
DITRANSITIVE: <> == TRANSITIVE
    <syn subcat re re fi sem binding> == UNI:<third>
    <sem args re re fi binding> == UNI:<third>
    <sem pred> == ("<word>" /3)
    <syn subcat re re> == COMP:<>
    <sem args re re> == ARG:<>.
```

[8] It will still be necessary to state a `sem pred` equation at non-lexical nodes such as `CAUSE`, as this is where rewritten causatives inherit their semantics.



Unbroken lines: default inheritance.

Broken lines: alternations, label in italics,
arrows point from children to parents.

Figure 9.17: Verb taxonomy

9.7.1 'Clear' verbs

The alternations from CLEAR are motivated in section 9.3.2. First we address the "clear the table of dishes" and "clear the table" forms.

```
Clear:<>   == CLEAR
    <word> == clear.

CLEAR:<>  == DITRANSITIVE
      <sem args re fi semfeats location> == yes
      <syn subcat re re fi syn> == PP:<>
      <syn subcat re re fi syn pform> == of
      <sem args re re fi semfeats> == PATIENT:<>
      <alt no-of> == C-OF-S:<>
      <alt no-of sem> ==  "<sem>"
      <alt no-of sem args re re fi binding>  == ex-q.

PP:<maj> == p.
```

The base form of the verb is the "clear the table of dishes" one, and correspondingly, the predicate has three arguments of which the second is the location and the third, the locatum. In our minimal account of semantic features as specifying role-types, we detail a location by adding

```
<location> = yes
```

to the set of semantic features. The third complement is a PP rather than the default NP. PPs have p as their major category.[9]

The alternation to the simple transitive is labelled **no-of**, as it is obtained from the base form through dropping the of-complement. The only change to the semantics is that the third argument is existentially quantified.

To state the equivalence between "clear X of Y" and "clear Y from X" we use a 'remove' alternation and bind the second argument to the third subcat list member and *vice versa*.

```
CLEAR:<alt remove> == REMOVE:<>
    <alt remove sem args re fi> == "<sem args re re fi>"
    <alt remove sem args re re fi> == "<sem args re fi>".

REMOVE:<> == PUT
      <syn subcat re re fi syn pform> == from.

PUT:<> == DITRANSITIVE
      <syn subcat re re fi> == PP
      <sem args re re fi semfeats location> == yes.
```

[9] This account, like GPSG (Gazdar et al., 1985) and HPSG, does not distinguish NP and PP semantics. The route for inheriting syntax is quite separate from the route for inheriting semantics, and the two only come together where the AVMs are built and re-entrancies interpreted. There is no straightforward way within the theory to state generalisations concerning the relations between the syntax and semantics of subcategorised-for items. This may be a shortcoming of the fragment, although it has not produced difficulties so far, and indeed has clarified the theory-building process. An approach which did not use unification variables and in which the order of the subcat list was, by default, directly the order of semantic arguments, is an alternative which has been investigated but proved difficult to implement, presented problems in relation to the re-entrancy variables wanted in HPSG lexical entries, and did not appear to provide advantages over the current version.

We need to swap the **of** for **from**. This is the **pform** for all 'remove' verbs, whether 'remove' is a base or extended form, so can be stated at **REMOVE**.

Almost nothing need be said regarding ergative *clear*. Paths for 'clear' verbs starting **alt no-of alt erg** will inherit first from **CLEAR**, then from **C-OF-S**, and then, as the **alt no-of** is stripped leaving **alt erg** as the leading subpath, from **PHYS-PROC**. Only one addition to the theory is needed. The third semantic argument should inherit its existential quantification from the line:

```
CLEAR: <alt no-of sem args re re fi binding>  == ex-q.
```

Currently, it does not, as **Clear/erg** queries will have leading subpath is still **alt no-of alt erg**. We remedy this with:

```
CLEAR: <alt no-of alt erg sem args re re fi binding>  ==
                        <alt no-of sem args re re fi binding>.
```

9.7.2 'Wipe' verbs

LR consider the 'wipe' verbs to be verbs of surface contact with non-basic **remove** senses. As with causatives there is a new element introduced into the meanings of the extended sense, here, the 'remove' component. Thus we cannot directly express the meaning of the 'remove' sense of a 'wipe' verb using only a single predicate, but will need to build an expression for the semantics which uses both WIPE and REMOVE.

```
Wipe:<> == WIPE
    <word> == wipe.

WIPE:<> == SURFACE-CONTACT
     <alt remove> == REMOVE:<>
     <alt remove sem> == AND:<sem>
     <alt remove sem args fi> == "<sem>"
     <alt remove sem args fi args re fi binding> == UNI:<third>
     <alt remove sem args re fi> == REMOVE:<sem>.

SURFACE-CONTACT:<> == TRANSITIVE.

REMOVE:<sem pred> == remove /3.

AND:<> == TRANSITIVE
    <sem pred> == and /2.
```

If I wipe the crumbs off the table, the syntax is that of a verb of removal, and the meaning is that I both wipe the table and remove the crumbs from the table. The default for paths prefixed **alt remove** ensures that the syntax, including ditransitivity and PFORM specifications, is inherited from **REMOVE**. For the semantics, the primary predicate is **and/2**, which has two arguments, both of which have predicate-argument structure. The first has the same predicate-argument structure as the base form of the verb with the location as second argument. The second has the predicate-argument structure of a 'remove' verb, with the locatum as the second argument. The lambda calculus equivalent is:

$$\begin{aligned} &\lambda z \lambda y \lambda x \qquad (WIPE/REMOVE(x,y,z)) = \\ &\lambda z \lambda y \lambda x \quad (AND/2(WIPE/2(x,z), REMOVE/3(x,y,z))). \end{aligned}$$

`AND` inherits from `TRANSITIVE` as a formal convenience. Like predicates relating to base-form transitive verbs, `and/2` takes two arguments in its argument list. It would be redundant to restate the mechanism for terminating a two-item list.

Supporting evidence for this analysis comes from the way in which 'wipe' verbs - but not the 'clear' or 'remove' verbs - can be used as verbs of putting as well as verbs of removing, as in "Wipe the grease on the axle" and "Brush the egg on the pastries". The alternation seems very similar to the 'remove' alternation. A 'put' sense for 'wipe' verbs is very easily added to the lexicon, re-using some of the structure built for 'remove' senses:

```
WIPE:
    <alt put> == <alt remove>
    <alt put syn> == PUT:<syn>
    <alt put sem args re fi> == PUT:<sem>.

PUT:<sem pred> == put /3
    <syn subcat re re fi syn pform> == on.
```

As discussed above in relation to `from`, `on` stands in for the set of prepositions which is a more suitable value for `pform`.

Little needs saying about the 'instrument' subclass of 'wipe' verbs:

```
INSTRUMENT:<> == WIPE.

Brush:<> == INSTRUMENT          %also: rake shear shovel vacuum
    <word> == brush.
```

A generalisation we miss here is that all these verbs are derived from nouns, in all but one case with no morphological change. Derivational morphology has received a DATR treatment (Evans, 1992) and in a larger fragment these verbal entries could be inherited from nominal ones.

9.7.3 The conative alternation

As part of their evidence that 'wipe' verbs are verbs of surface contact, LR point out that some of them participate in the conative alternation, as between "cut the bread" and "cut at the bread", and "rub the lamp" and "rub at the lamp". They say:

> This construction is attested with verbs whose meaning includes notions of both meaning and contact ... The verbs found in this construction are drawn from several semantic classes; they include verbs of impact by contact, such as *hit* or *kick*, and verbs of contact and effect, such as *cut* or *hack*. Simple verbs of contact and simple verbs of motion are not found in this construction (**Terry touched at the cat, *Nina moved at the table*). The conative construction is also not found with verbs of change of state, including the *clear* verbs ... as they lack the appropriate meaning components. (p 135)

Our analysis here will be a very simple one. All verbs undergoing the conative alternation will have an `alt conative` prefix stripped and will then inherit from `CONATIVE`. There will be a non-primitive predicate for the conative form. Regarding the "slight change in meaning" (p 135) associated with the conative, we introduce a one-argument predicate. Whatever the meaning-shift is, we represent it by `conative /1` which takes as its argument the basic verb which is to have its meaning shifted. In lambda calculus the meaning-relation may be described as:

$$\lambda y \lambda x \text{ RUB/2/CONATIVE/2}(x, y) = \lambda y \lambda x \text{ CONATIVE/1}(\text{ RUB/2 } (x, y)).$$

The DATR is:

```
SURFACE-CONTACT:<alt conative>    == CONATIVE.

CONATIVE:<>                       == TRANSITIVE
     <sem pred>                   == (conative /1)
     <sem args fi>                == "<sem>"
     <sem args re>                == NIL:<>
     <syn subcat re fi syn>       == PP:<>
     <syn subcat re fi syn pform> == at.
```

The last two lines express the syntactic difference. A conative form still has two complements so inherits from **TRANSITIVE**, but we override the default that complements are NPs by stating that the second complement is a PP with pform *at*.

Not all 'wipe' verbs participate readily in the conative alternation. The issue is as addressed in the previous chapter. We have set up a fragment to indicate which verbs *may* participate in which alternations, but not which *do*. There were two mechanisms presented in the last chapter: a negative one, of building an **altlist** stating which alternations may apply to a word (and by implication which may not), and a positive one of stating where a usage-type was attested. Both mechanisms are also applicable to the fragment of the verb lexicon.

The quotation from LR claims the conative alternation may be associated with any verbs with both 'contact' and 'motion' meaning components. So a conative sense is to be added to the **altlists** of just those verbs. Meaning components of verbs are distinct from the semantic features we have considered so far. The latter, such as 'volition' and 'causally-affected', apply or fail to apply to entities, prototypically physical objects such as people and chairs, which are standardly linguistically realised as nominals. In the verbal lexicon, they relate to the arguments of verbs. By contrast, 'contact' and 'motion' apply or fail to apply directly to verbs rather than to their arguments. So, in addition to **semfeats**, we introduce a new feature, **vsemfeats** for components of verbal meaning simpliciter. So:

```
Rub:<sem vsemfeats contact> = yes.
Rub:<sem vsemfeats motion> = yes.
Move:<sem vsemfeats contact> = no.
Move:<sem vsemfeats motion> = yes.
Touch:<sem vsemfeats contact> = yes.
Touch:<sem vsemfeats motion> = no.
```

conative should be added to the **altlist** for any transitive verb if and only if the answer to both these queries is **yes**. The mechanism for doing this in DATR, which involves more sophisticated DATR programming techniques than we have used so far, is presented in Appendix C.

It is not at all clear what we might include in the inventory of **vsemfeats**. The analysis has freely used an inventory of the semantic features that verbs expect of their arguments.[10] But where LR say they consider it their project to uncover "syntactically relevant components of verb meaning" (p 123) they are concerned with aspects of meaning which relate directly to the verb rather than to its arguments, and it is these 'verb meaning components' or 'verbal semantic features' that the attribute **vsemfeats** is provided for. LR mention several candidate **vsemfeats**, 'contact' and 'motion' among them, but their analysis falls far short of providing a motivated inventory of primitive features. This is of course scarcely surprising. A large number of critiques of Katz & Fodor (1963)'s decompositional semantics point to the theoretical and practical

[10] Dowty talks of entailments that a verb licenses about its arguments rather than semantic features: the distinction is not important here.

snares in the enterprise when aimed at nominal semantic features (e.g. Bolinger (1965), Fillmore (1978)) Verbal ones are still less charted territory. Works such as the Vendler verb classification, Talmy (1985) on 'Lexicalisation Patterns', and LR itself, provide some hints of what the essential `vsemfeats`, and the structure of the domain, might be, but the endeavour is in its infancy.

9.8 Shortcomings of LR's approach

The style of analysis in LR presents a range of alternations relating distinct forms of verbs, and treats the various possible forms as either present or absent in the lexicon. In the tradition of post-Chomsky theoretical linguistics, it is concerned with rules determining what is and is not in the language.

The argument developed in Chapters 1 and 4 and given empirical support in Chapters 5 and 6 shows that the 'existence' of a word sense is a function of its frequency and predictability: that identifying and individuating word senses is often difficult: that identifying and individuating alternations is still more difficult: and what is or is not part of the language is an open question. LR make minor concessions to the blurred edges of what is part of the language —"the verb *wipe*, unlike *clear*, only marginally, if at all, allows the locatum argument to be expressed in an *of* phrase"(p 128)— but the sentence used to illustrate the possibility:

> ∗Kay wiped the counter of fingerprints.

is starred and the analysis proceeds on the basis that 'wipe' verbs simply do not have this option. The formalisation also proceeds on that basis. The methodology of starred sentences carries the implication that there are some ways a word will never be used. But this is territory where intuitions are often weak and unstable. A shift from introspective to corpus evidence is imperative. The corpus will only ever provide frequency statistics: it will give likelihoods, not yes/no answers. There are significant difficulties involved in integrating statistical and rule-based methods (as discussed in section 7.7) and this is an area where much more work must be done.

9.8.1 Optional arguments and adjuncts

One area where the 'yes/no' methodology is unsatisfactory is in distinguishing optional arguments and adjuncts. "Taking the candy from the baby" and "removing the candy from the baby" not only describe the same activity but seem to do so using parallel linguistic structures. However according to LR and the formalism presented above, *remove* subcategorises for a prepositional phrase with *from* or similar. *Take* is not discussed, but the usual analysis would be that the prepositional phrase is an adjunct. The one is a lexical matter, described at the lexical entry for *remove*. In this case, *from* acts simply as a case-marker and the role ascribed to the NP following it is part of the meaning of the verb. With *take*, on the other hand, the lexical entry tells us nothing about the PP and the role the PP plays is determined by the meaning of *from*. This is counter-intuitive. *From* seems to make a very similar contribution to both sentences.

The distinction between complements and adjuncts has occupied many authors, and the goal here is not to contribute to the debate, but just to note the desideratum that complement-like adjuncts and adjunct-like complements are treated in comparable ways, so that the two 'baby' sentences can receive equivalent analyses. LR do not discuss adjuncts, or more generally, the interactions between the lexicon they describe and syntactic processes such as the potential for VPs to take adjuncts. A fuller account of the structure of the verb lexicon would include a discussion of those interactions, and a complete theory would demonstrate how the lexical meaning of *from* contributed to the meanings of both sentences, perhaps to varying degrees.

9.8.2 Syntax, semantics and circularity

LR's enterprise runs a risk of circularity which the authors do not mention. Their goal is to establish which 'components of meaning' have linguistic significance. The risk is that an alternation not only provides evidence of linguistic significance, but also proves criterial for whether a given 'component of meaning' is present. For example, *shut* is presented as belonging to the same class as *thicken* (p 134): they are both change-of-state verbs. But what is the evidence that they share components of meaning? The clearest evidence is syntactic: they are both related to adjectives and undergo a characteristic alternation. LR do not discuss how 'components of meaning' are to be identified independently of syntactic, and in particular alternation-based, criteria. So their claim to be unearthing correlations between syntactic and semantic domains is a shallow one: to be strengthened, the philosophically tortuous question of non-syntactic criteria for identifying 'components of meaning' must be addressed.

This does not undermine their, or our, project. The lexicon is a complex entity, its structure is a legitimate and important object of enquiry in its own right. Also, the patterns of subcategorisation for a verb are central to what its meaning is (Fisher, Gleitman, & Gleitman, 1991). The study should not be seen as an investigation of generalisations connecting alternations and meaning: rather, it is simply a study of (aspects of) meaning.

LR claim that:

> The *clear* verbs and the *wipe* verbs might appear to be verbs of removal even in the location-as-object frame, but this impression results from real-world knowledge associated with the events these verbs denote and is not part of their linguistic classification. The analysis of the meaning of the *clear* and *wipe* verbs ... explains why these verbs behave differently from verbs of removal. (p 137)

Their appeal to the distinction between linguistic and real-world knowledge falls foul of the arguments in section 8.2. The distinction is arbitrary and the same representation scheme should be used for both. LR push the fuzzy boundaries between usage-types out of the 'linguistic' domain, in order that linguistics can be presented as a territory where rule-based accounts apply. The claim that the account explains rather than describes is premised on the circularity described above. The view taken in this thesis is simply that the *clear* verbs and the *wipe* verbs sometimes behave as verbs of removal and denote 'removal' events. The most elegant account of this behaviour is that there is an alternation from each of these classes to the class where verbs of removal have their base sense. The representation then expresses concisely both the possibility that *clear* and *wipe* verbs will pattern as *remove*, and the possibility that they will not, and sits comfortably within the overall structure of the verb lexicon.

9.9 Summary

This chapter has shown how verb alternations can be formalised in a way that captures and exploits generalisations about the behaviour of verb classes. The linguistic data was taken from two articles: Atkins et al. (1986) and Levin & Rappoport Hovav (1991). The alternations investigated were ones where the core meaning of the verb was not affected by the alternation but the numbers and semantic roles of complements were. As in the previous chapter, the formalism used was DATR and the basic hierarchy was motivated both by syntax, and the need to express generalisations about syntactic behaviour, and classes identified according to meaning. A further constraint was that the lexical entries should be of a kind specified by a grammar formalism, so they could be used for parsing or generation and would be part of an integrated formal account of the language. The grammar formalism used was HPSG.

First, HPSG-style verbal lexical entries, and the mappings between them corresponding to alternations, were described. But at this stage, the generalisations were not captured. So then these entries were translated into DATR, and arranged into a taxonomy so an alternation only needed expressing once, at a non-terminal node from which the verbs to which it applied would inherit. The theory was developed to succinctly represent lexical entries for seven classes of base verbs and eight alternations applying to or between one or more of the classes. Information about syntax, semantics, and patterns of polysemy was concisely expressed in a manner both theoretically and computationally appealing.

Finally some constraints on the approach were discussed. The articles and the formal theory assumed a simple distinction between what was, and what was not, part of a language. An important area for the development of the approach is to establish ways in which statistical information, regarding more and less likely uses of words in a language, can be incorporated.

Chapter 10

Conclusion

What is polysemy? How is it that language-users can effortlessly comprehend and generate novel uses of words? How might natural language processing computer systems deal with multiple meanings, or novel meanings? These were the questions with which we set the thesis going. Have they been answered?

10.1 What is Polysemy?

The blind men of Hindustan attempt to comprehend the elephant:

> The first approached the elephant,
> And happening to fall
> Against his broad and sturdy side,
> At once began to bawl,
> "Bless me, it seems the elephant
> Is very like a wall."
>
> The second, feeling of his tusk,
> Cried, "Ho! What have we here
> So very round and smooth and sharp?
> To me 'tis mighty clear
> This wonder of an elephant
> Is very like a spear."
>
> The third approached the animal,
> And happening to take
> The squirming trunk within his hands,
> Then boldly up and spake;
> "I see," quoth he, "the elephant
> Is very like a snake."
>
> The fourth stretched out his eager hand
> And felt about the knee,
> "What most this mighty beast is like
> Is mighty plain," quoth he;
> " 'Tis clear enough the elephant
> Is very like a tree."

> The fifth who chanced to touch the ear
> Said, "Even the blindest man
> Can tell what this resembles most;
> Deny the fact who can,
> This marvel of an elephant
> Is very like a fan."
>
> The sixth no sooner had begun
> About the beast to grope
> Than, seizing on the swinging tail
> That fell within his scope,
> "I see," cried he, "the elephant
> Is very like a rope."
>
> — John Godfrey Saxe

Polysemy, like the elephant, may at first encounter seem like a variety of things: like homonymy, ellipsis, metaphor; like syntactic variation, collocation, pragmatic reasoning. The question, "What does it mean to say a word has many meanings?", may be addressed from any of these angles. This thesis has aimed to show the beast in its entirety.

The empirical studies looked directly at polysemy as found in a dictionary, and the relations between that and the ways words were used in a corpus. An uncritical consideration of a dictionary might give rise to the following thought. It lists words, thereby providing a key to a set of entities in the language. Likewise, it lists word senses, so must also be providing a key to another set of entities. The dictionary studies set this misconception to rest. Different words are different because they have different spellings and sounds. There is no comparable fact of the matter for determining what makes a word sense different.

The two studies exposed a range of phenomena involving words having a variety of uses, and a corresponding range of lexicographical devices. We identified the SFIP criterion: a usage-type merits listing in a dictionary when it is Sufficiently Frequent and Insufficiently Predictable, so that it is neither too obscure nor so obvious as to be a waste of space. The two parts interrelate. The more common a pattern, the stronger the prediction that further words follow it. For paper lexicography, what is sufficiently frequent and insufficiently predictable depends on the size and target audience of the dictionary. For inheritance-based computational lexicons, 'predictability' falls under the more general treatment of the inheritance structure of the lexicon.

Polsemy was not found to form any kind of 'natural kind'. It described a crossroads. In the direction of less predictability lay homonymy. In the direction of greater predictability lay alternations. Some alternations, such as that between *mink*, *beaver* or *fox* referring to an animal or a coat, were located in the borderland between polysemy and metonymy. Others, such as the relation between transitive and unspecified-object *cook, drive* and *write* fell between polysemy and a syntax-based account of optional arguments. In other directions lay collocations and analogy: the expression *nervous wreck* could be described in the dictionary under a distinct sense of *wreck*, or with an entry to itself, or might not receive a treatment at all on the basis that it was an analogy that the dictionary reader could make sense of without recourse to the dictionary. *Nervous wreck* fell in the territory between polysemy, collocations and analogy.

10.2 Novel uses

> "When *I* use a word," Humpty Dumpty said, in rather a scornful tone, "it means just what I choose it to mean, neither more nor less."
>
> "The question is" said Alice, "whether you can make a word mean so many different things".

"The question is" said Humpty Dumpty, "which is to be master – that's all".

– Lewis Carroll; Alice's Adventures in Wonderland

Was Humpty-Dumpty right? Can a word mean whatever we choose it to mean? Certainly, we can sometimes use words in novel ways, to mean something other than what they usually mean. Possibly, in sufficiently extreme circumstances, almost any word can be used to mean almost anything. But a word only succeeds in meaning what the speaker intends it to mean if the meaning is apparent to other users of the language. The hearer must be able to work out what the speaker meant in using the word. It must be predictable that the word could be used to mean that in that setting. There lies the link between novel usages and polysemy. Entirely novel usages are of nil frequency (by definition) but are completely predictable. (That is not to say that poets' choices of words are predictable: only that it is predictable that that is the meaning that the word would take on in that setting.) As frequency increases and predictability decreases, so a word is understood less on the basis of context and more directly on the basis of its lexical entry, and we slide gently towards polysemy.

There are two kinds of novel use, corresponding to two kinds of ways in which the meaning might be determined. It might be determined by linguistic context, or by non-linguistic context. (These are of course two ends to a scale, rather than distinct phenomena.) Original metaphors and metonyms are at the end of the scale where linguistic context plays only a supporting role. A use of a verb with a set of complements it has never been used with before lies at the other extreme. Linguistic and non-linguistic aspects frequently cannot be disentangled. Where two words are similar in meaning, it is usual for there to be both a (non-linguistic) similarity to their denotations and a (linguistic) similarity to their distributions. Novel uses very often depend on words being used in ways which are unfamiliar for that word, but relatively familiar for other words similar to it in meaning: the novel use is then interpretable by virtue of the linguistic and non-linguistic features it shares with its neighbours in semantic space.

While Humpty Dumpty may be right, this does not invalidate the lexicographer's attempt to specify the meanings for a word. The lexicographer concerns him or herself with the usage-types meeting the SFIP criterion. Not only is this the information required to interpret words as most commonly used; it also contains — implicitly in a paper dictionary, explicitly in an inheritance-based computational lexicon — the patterns to extrapolate from to interpret a range of less common or novel uses.

10.3 NLP and Polysemy

Wilks's problem provided a point of entry for the study of polysemy. It presented a well-defined goal: to enable NLP systems to deal appropriately with natural language inputs, where many of the words in the input have more than one sense listed in the system's lexicon. Chapter 2 chronicled the history of assaults on the problem. These generally assumed the Bank Model, which was found lacking in Chapter 6. Building on the 'crossroads' nature of polysemy, Chapter 7 proposed that NLP needed a variety of mechanisms for representing polysemous senses and interpreting non-standard usage-types, and that in each case, the mechanism was also required for some other phenomenon; homonymy, collocations, analogy, or metonymy.

Where polysemy was regular, and could be described in terms of an alternation relating to a number of words, it was redundant to express the facts about the alternate form at each lexical entry where it applied. There was a generalisation: this only needed stating once, provided that there were explicit mechanisms for stating that words inherited it. Chapters 8 and 9 showed how the lexicon could be organised to do this in a manner well-suited to NLP applications, introducing and formalising some facts about the polysemous behaviour of two classes of words. These formal

theories are of course small fragments, but may serve as prototypes for larger-scale projects in formal lexicography.

10.4 Further work

The thesis suggests several avenues for further research.

10.4.1 Evaluation of disambiguation systems

Chapter 2 reviewed a number of NLP systems for lexical disambiguation. Some such as McRoy's appeared most impressive. However, as McRoy herself notes, it is no straightforward matter to describe how successful such a system is, or to compare it with other systems. If the Bank Model applied universally, so every word had a set number of meanings and people all agreed which meaning applied in each citation, then, once a 'test corpus' had had every polysemous word disambiguated, it would be possible to say what proportion of the time a given system arrived at an answer which matched the test corpus. But none of these conditions hold. Firstly, dictionaries and people tend to disagree about how many senses a word has, or what they are. The question is clearly also dependent on domains: do we consider only 'general vocabulary' or also specialised domains: if so, how many, and how do we identify them or know which one we are in? Secondly, people face various dilemmas when they try to slot usages into dictionary senses. This makes the provision of a test corpus problematic. The best answer a person or NLP system could give will sometimes be 'halfway between senses *a* and *b*', or 'including elements of *a* and *b*', or 'an extended sense of *a*'. Thirdly, a system that makes many near misses will, for some purposes be better, and for others worse, than one that does not commit itself when the evidence for selecting a sense is equivocal. The various dimensions of the problem are all touched on in this thesis, but to weld them into an evaluation system will be a substantial further piece of research.

10.4.2 Using semantic fields to identify alternations (and *vice versa*)

Semantic fields and alternations are, to a degree, interdefined. We expect words in the same semantic field to undergo the same alternations, and the fact that they undergo the same alternations counts as evidence that they are in the same semantic field. This is a matter deserving more attention. An empirical study might take the following form. First, choose a semantic field or set of near-synonyms. Then consider the senses given for each of the words in a number of dictionaries. Then determine the extent to which the same alternations apply to all the words by looking at corpus evidence, and seeing where the frequency distributions between different senses was equivalent for all the words in the group. Such a study would take further our understanding of the relationship between frequency factors and the inheritance structure of the lexicon. It would provide clues to the ways in which the meanings of novel word-uses were predictable from the distribution of that word and its near neighbours in a corpus.

10.4.3 Frequencies and inheritance structure

There is a theoretical corollary to the empirical exercise described above. As this thesis has argued, it is appropriate to represent the lexicon in an inheritance-based formalism, yet also, frequency data play an essential role in determining the structure of the lexicon. Ways for the formalisms to operate with frequency data must be devised. This is both important for a theoretically adequate account of the lexicon, and practical. Word-sense frequencies have already started playing a role in lexical disambiguation programs, but the statistics are hard to come by. We might expect words participating in alternations to, by default, share the same relative frequencies between

senses. Thus if 80% of occurrences of both *melt* and *freeze* are intransitive and 20% causative, we might expect the same to hold for *evaporate* and other physical-process verbs where, owing to lower absolute frequencies, we have no direct evidence. The 80:20 ratio could then be stated at the higher node in the inheritance hierarchy. There is a host of difficult questions regarding how such relative frequencies might be calculated, where they might be inherited, and what they would be good for. In particular, when should a word be assumed to follow the pattern for its semantic field and when must it to be treated as a special case? If those questions can be answered, then scant information on relative frequencies of different senses could be aggregated and used to contribute to our understanding of the behaviour of whole classes of words. Such reasoning is likely to play a major role in our understanding of the lexicon in the future.

10.5 To conclude: summary of principal contributions

The thesis, then, has contributed to our understanding of polysemy in a number of ways. To finish, we restate four principal conclusions and contributions.
The thesis:

- shows the Bank Model to be fatally flawed;
- presents the SFIP criterion, describing when it is appropriate for a paper dictionary to list a usage-type, and describes the relation of polysemy to the four phenomena it falls between: homonymy, collocations, alternations and analogy;
- presents formal theories of two fragments of the lexicon, concisely describing alternations and related aspects of lexical structure.

Appendix A

Words examined in matching study

Showing: words examined; parts of speech; number of senses for each part of speech (excluding senses for collocations and including, for nouns, any senses specific to the plural); and in the last column, '1' if the word had only one sense for each part of speech, 'N' if every usage could be classified as one and only one sense, and 'Y' if it could not.

Word & forms	POS	No. of senses	Result
absence/s	n	2	N
accompany/s/ed/ing	v	3	Y
additional	a	1	1
alternative	n, a	2, 3	Y
anxious	a	3	Y
application/s	n	6	Y
apply/s/ed/ing	v, a[1]	5, 1	Y
appointment/s	n	3	N
arrive/s/ed/ing	v	5	Y
article/s	n	5	Y
aspect/s	n	4	Y
benefit/s/ed/ing	v, n	2, 4	Y
boat/s	n	3	Y
border/s/ed/ing	v, n	3, 2	Y
brain/s	n	3	Y
bread	n	3	Y

Word & forms	POS	No. of senses	Result
busy	a	4	Y
camp/s	n, v	3, 1	Y
capable	a	3	N
capacity/s	n	3	N
chapel/s	n	5	Y
clean	v, a, adv	5, 2, 1	Y
clothes	n	1	1
coffee/s	n	2	Y
colour/s	n	7	Y
coloured	a, n, v	3, 2, 3	Y
competition/s	n	3	Y
composition/s	n	6	Y
conception/s	n	3	Y
connection/s	n	6	Y
continent/s	n, a	1, 1	1
cool	v, a	1, 6	Y

[1] adjectival form is 'applied'

Word & forms	POS	No. of senses	Result
credit/s	n, v	7, 2	Y
criminal/s	a, n	3, 1	Y
critical	a	3	Y
curious	a	2	Y
custom/s	n	6	N
dangerous	a	1	1
decide/s/ed/ing	v, a	3, 2	Y
delightful	a	1	1
derived	a, v	0[2], 3	Y
design/s	v, n	2, 6	Y
destroy/s/ed/ing	v	2	N
detail/s	n	2	Y
distinction/s	n	3	Y
divine	a, v	2, 2	N
don/s/ed/ing	n, v	1, 1	1
dust/s/ed/ing	n, v	5, 2	Y
eleven/s	n	2	N
embassy/s	n	1	1
emphasis/es	n	1	1
energy/s	n	3	Y
ensure/s/ed/ing	v	1	1
enter/s/ed/ing	v	7	Y
entrance/s	n	3	Y
escape/s/ed/ing	n, v	2, 3	Y
establishment/s	n	4	Y
evil/s	n, a	1, 2	N

Word & forms	POS	No. of senses	Result
exchange/s	n, v	4, 1	Y
exciting	a	1	1
exercise/s/ed/ing	n, v	5, 3	Y
expensive	a	1	1
explain/s/ed/ing	v	2	Y
factory/s	n	1	1
farming	n	1	1
fashion/s	n, v	3, 1	Y
favourite/s	a, n	1, 3	Y
federal	a	2	Y
feed/s/ed/ing	n, v	4, 5	Y
flights	n	7	Y
football/s	n	4	Y
formal	a	4	Y
frame/s	n, v	6, 3	Y
friendship/s	n	2	Y
gallery/s	n	4	Y
gas/s/ed/ing	n, v	7, 2	Y
generation/s	n	4	Y
gift/s	n	3	Y
guest/s	n	4	Y
herring/s	n	1	1
hit/s/ing	n, v	5, 5	Y
ideal/s	n, a	2, 3	Y
image/s	n	6	Y

[2] no dictionary entry for adjectival form

Appendix B

Matching study: examples of misfits

The following are examples of citations which cannot be satisfactorily classified as one and only one of the dictionary senses of the word. Words have been selected to demonstrate a variety of kinds of cases.

arrive *v* [I] **1** to reach a place at the end of a journey: *We arrived home safely.* | *What time does the plane arrive in New York?* —compare DEPART (1) **2** to come to a place, esp. by arrangement: *Shall we start now, or shall we wait for the others to arrive?* **3** to be brought or delivered to a place: *Has the post arrived yet?* | *I'm still waiting for those books I ordered to arrive.* **4** to happen as expected or arranged; come: *At last the great day arrived.* | *Her baby arrived* (=was born) *yesterday.* **5** to win success: *They felt they had really arrived when they made their first record.*

1	barge. When the American symphony orchestra	arrive in Marlow on July 15 they will use an 18-foot
2	climax in the island orgy. Here, the guests	arrive in ghost-like yachts, the wildly flapping
3	3,000 men, who must take about a fortnight to	arrive. If the UN forces were thick enough on the
4	are things wrong with the film, but the print	arrived from the cutting room only a few hours before
5	him of the Vienna outcome. Last night Mr Rusk	arrived in London in time to join the Buckingham

Comments: The distinction between s1 and s2 seems very slight, and there was rarely any evidence in the citations that provided grounds for making the classification. In case 4, s3 is equally applicable.

chapel *n* **1** [C] a place, such as a small church, a room in a hospital, prison, etc. (but not a PARISH church), used for Christian worship **2** [C] a room or area in a church with its own ALTAR, used esp. for private prayer and small religious services **3** [C] **a** (esp. in England and Wales) a place of Christian worship used by NONCONFORMISTS (=those who do not belong to the established church or the Roman Catholic church) **b** (in Scotland) a Roman Catholic church **4** [U] the religious services held in such places: *He goes to chapel every Sunday night.* | *I'll meet you after chapel.* **5** [C+*sing./pl. v*] a branch of a union in jobs such as printing and JOURNALISM: *The chapel has/have voted to go back to work.* | *a chapel meeting*

1	has shown this to be the remains of the	chapel of the shrine of our Lady, visited and described
2	the levelling of the sloping site by the	chapel builders had destroyed much of the original
3	had later been added to the east end of the	chapel. A great porch at the west end, of still later
4	from home. Excavation work identifies shrine	chapel. Walsingham evidence. Excavation of the small
5	the chapel, it is known to have stood above the	chapel floor. As this floor had been almost completely

Comments: Although the distinction between s1, s2 and s3 was quite clear, there was rarely evidence in the citation for determining which applied. In some but probably not all cases, more context would resolve the issue.

colour *BrE* ‖ **color** *AmE n* **1** [U] the quality in objects which allows the eyes to see the difference between (for example) a red flower and a blue flower when both are the same size and shape: *The book has illustrations in colour.* | *These insects can change colour.* | *a colour television* **2** [C] red, blue, green, black, brown, yellow, white, etc.: *"What colour is this paint?" "It's red."* | *"What colour did you paint the door?" "I painted it red."* **3** [S;U] the general appearance of a person's skin, esp. as this shows the state of their health: *He lost colour* (=became pale) *during his illness.* | *The fever gave her a* **high colour**. (=a lot of colour) | *The cold wind brought colour to her cheeks.* (=made them red) **4** [C] the colour of a person's skin showing which race they belong to: *people of all colours* (=black, brown, white, etc.)—see also COLOURED **5** [U] details or behaviour of a place, thing or person, that interest the mind or eye and excite the imagination; character: *She loved the life, noise and colour of the market.* | *The lecturer told a few jokes and anecdotes to* **add colour to** *his talk.* —see also LOCAL COLOUR **6 give/lend colour to** to make (something, esp. something unusual) appear likely or true: *Her wet hair lent colour to her claim that she had fallen into the lake.* **7 off colour** *infml* not in good health: *You look a little off colour today.* **8 see the colour of someone's money** *infml* to have clear proof that someone has enough money to pay: *I don't trust him to pay us—I want to see the colour of his money first.* —see also COLOURS, OFF COLOUR, PRIMARY COLOUR

colours *BrE* ‖ **colors** *AmE n*[P] **1** a special sign, cap, BADGE etc., worn as a sign of one's club, school, team, etc.: *He won his colours* (=was chosen for the team) *for football this year.* **2** the official flag of a country, ship, part of the army etc.: *the regimental colours* **3 one's true colours** one's real (esp. unpleasant) character, esp. when seen for the first time: *I liked him at first, but now he has shown his true colours.* | *I've seen him in his true colours.* —see also FLYING COLOURS, **sail under false colours**.

1	peering in through the window, brought back	colour which (although it was divorced from any
2	only a small amount of grey. As hair loses its	colour pigment, alterations take place in its structure
3	comparative freedom of expression in the use of	colour and in sheer ingenuity in the use of all kinds
4	curved or angular, rightly or wrongly lacking in	colour - and behind the description is the implication
5	: 'Wicki's blacks and greys are not only the	colours of the lost and the forgotten, but they are
6	and excitingly applied and inventively combined	colours play some of their very sunniest compositions.

Comments: Case 1 would possibly be resolved (between s1, s3 and s5) by more context. In case 2 there is nothing to choose between s1 and s2, unless the fact that it appears in a noun-noun compound is taken as grounds for matching it with sense 1, since one of s1's examples is a noun-noun compound. 3 might, but might not, refer to s1 colour. It could refer to s5 in a novel or to the 'music' sense, which applies to case 6 and which can be seen as a metaphorical extension of any of s1, s2, s3 and s5. Case 5 makes explicit the fact that the usage refers both to s2 and also to s1 of colours(pl), and probably also to s3.

competition *n* **1** [C] a test of strength, skill, ability, etc.: *to go in for/enter a competition* | *a crossword competition [+to-v] a competition to find a designer for the new airport building* **2** [U (**with, between, for**)] the act of competing; the struggle between several people or groups to win something or gain an advantage; RIVALRY: *There was intense/keen/fierce competition between the journalists to get the story.* | *He was* **in competition with** *some world-class runners, so he did well to win the race.* | *The two products/companies are in direct competition.* (=are/produce similar products at similar prices) | *They believe that competition in business benefits the consumer.* **3** [U] the (other) competitors: *Anyone wanting to enter the computer business faces tough competition.* | *It's important in business to keep a careful watch on the competition.* | *They had to keep their prices low because of foreign competition.*

1	of the new building should be put up to open	competition - and a building might emerge at last
2	houses. For better or for worse this would bring	competition to the licensed trade. He said : 'that
3	protection from teenage drinkers as well as more	competition. Nearly 400 wives and relatives of licensees
4	while the Americans enjoyed the fiercest	competition in their trials to select their defender.

Comments: As often with noun-noun compounds, case 1 is simply indeterminate between s1 and s2. Cases 2, 3 and 4 introduce subtle ambiguities that might well not be resolved by more context. In cases 2 and 3 the ambiguity is between s2 (more competitive) and s3 (more competitors). In case 4 an s4

reading is also available, in which 'competition' might denote the Major League and the 'Americans', the spectators.

critical *adj* **1** of or being a moment of great danger, difficulty, or uncertainty, when a sudden change to a better or worse condition is likely; of or being a CRISIS: *a critical stage in his illness/in the negotiations | a matter of critical importance | We arrived at the critical moment. | The next two weeks will be critical (for the company).* **2** providing a careful judgement of the good and bad qualities of something: *a critical analysis/assessment of the government's record | critical writings | Her new book received* **critical acclaim.** (=was praised by the critics) **3** [(**of**)] finding fault; judging severely: *Why are you so critical of everything I wear?*

1	more competitive spirit in industry, a far more	critical attitude towards costs, whatever their origin,
2	few nervous children are diffident about facing a	critical audience of their own fellows. To overcome
3	other communities. As the Eskimo artists are self-	critical, and their work is being fostered with knowledge
4	still lodge a humble protest as well as deliver a	critical ukase, and the phrase 'it's a bit much

Comments: 'Critical' (and 'self-critical') participate in a common pattern relating a 'neutral' sense of the word to a 'far end of the scale' sense. Unlike 'tall' and 'heavy', but like 'distinction' (see below), usages of 'critical' can carry aspects of both senses, with it often being hard to determine the balance of the two aspects.

design *n* **1** [C] a drawing or pattern showing how something is to be made: *Have you seen the latest designs for the new library?* **2** [U] also **designing**— the art of making such drawings or patterns: *a course in dress design* **3** [U] the arrangement of the parts in any man-made product, such as a machine or work of art, as this influences the product's practical usefulness, artistic quality, etc.: *The success of this car shows the importance of good design in helping to sell the product.* **4** [C] a decorative pattern, esp. one that is not repeated: *a carpet with a floral design in the centre* —see picture **5** [C] a plan in the mind; SCHEME **6 by design** as a result of purposeful planning; intentionally: *She arrived just as we were leaving, but whether this was by accident or by design I'm not sure.*

1	necessary to explore the many possibilities of	design in this ancient craft. In addition to chapters
2	are later diversified by that very characteristic	design of a rectangle surmounted by a semi-circle
3	years old. If the new proposals are accepted, the	design of the new building should be put up to open
4	of the leg with toe. For instance, the	design at (F), fig 1, is formed entirely on the
5	and river and on the right hand side of the	design appears one of the inscribed rock tablets

Comments: Case 1 has aspects of both non-count senses (s2 and s3); case 2, of two count ones (s1 and s4). Case 3 lies between s1, s2 and s3. Cases 4 and 5 might be s1 or s4, or the all-embracing s3.

distinction *n* **1** [C;U (**between**)] the fact of being different; clear difference: *I can't see any distinction between these two cases. | It's important to* **draw a distinction** *between the policies of the leaders and the views of their supporters.* **2** [S;U] the quality of being unusually good; excellence: *a writer of real distinction* **3** [C] a special mark of honour, fame or excellence: *These are the highest distinctions that have ever been given by our government. | She got a distinction in her chemistry exam. | This country* **enjoys the dubious distinction** *of having the highest rate of inflation in the world.*

1	before the war, shares with Rilke and Kafka the	distinction of having origins which seem to escape
2	edge to a collar; these add a charming, feminine	distinction that nothing else can give. Although
3	separate bream species. It can not even claim the	distinction of being a bream 'variety' or '

Comments: See 'critical' above.

dust *n* **1** [U] dry powder made of extremely small grains of waste matter, esp. of the kind that settles on indoor surfaces: *There was a layer of dust on the books before I cleaned them. | atmospheric dust* **2** [U] finely powdered earth: *The car raised a cloud of dust as it went down the earth road. | the heat and dust of India | The rain soon settled/laid the dust.* (=stopped it from rising, by making the ground wet)

3 [U] fine powder made of small pieces of the stated substance: *gold dust* | *coal dust* **4** [U] *lit* the earthly remains of bodies once alive: *the dust of our ancestors* **5** [S] an act of dusting: *I gave the living room a quick dust.* **6 kick up/raise a dust (about)** *infml* to argue and shout (about) **7 when the dust has settled** *infml* when the confusion is over —see also DUSTY, **bite the dust**

1	in the pits. In South Wales 346 died from	dust in 1959 alone. Miners who had worked during
2	who had worked during the long period of '	dust-approved conditions,' including younger men
3	there will be resistance.' 346 in a year.	Dust kills many more people than gas, bad roofs,
4	this combination. The body returns to the	dust, the life returns to God, and the spirit disappears.
5	that his seed should be like the stars, the	dust and the sand that can not be numbered, we know
6	holes in the drill body must be kept free from	dust, screws should be checked for tightness regularly,
7	mild steel and other softish metals splinters and	dust are a sign that more pressure is required, so

Comments: Similar to 'chapel' above, except that in this case, some dust is equally dust/1, dust/2 and dust/4, which is unlike 'chapel' since (almost) every individual chapel is either a chapel/1, chapel/2 or chapel/3. Note that, firstly, although case 1 refers to coal dust, it would be inappropriate to classify it as sense 4 since that requires that the substance (i. e. coal) is stated, as in the example noun-noun compounds. Secondly, case 4 alludes strongly to s4 but the dust that the body returns to (in this citation) is not the dust that the body turns into; the sense 4 figure of speech is implied but not employed.

image *n* **1** [C (**of**)] a picture formed in the mind: *She had a clear image of how she would look in twenty years time.* **2** [C] a picture formed of an object in front of a mirror or LENS, such as the picture formed on the film inside a camera or one's REFLECTION in a mirror **3** [C] the general opinion about a person, organization, etc., that has been formed or intentionally created in people's minds: *The government will have to improve its image if it wants to win the next general election.* | *The company tries to project an image of being innovative and progressive.* **4** [(*the*) S (**of**)] a copy: *He's the (very) image of his father.* **5** [*the*+S+**of**] a phrase giving an idea of something in a poetical form, esp. a METAPHOR or SIMILE **6** *old use* likeness; form: *According to the bible, man was made in the image of God.* —*see also* MIRROR IMAGE, SPITTING IMAGE

1	his personality is impressed on every delphic	image. How it is that Celtic mystery and individual
2	always alive and always changing; but the visual	image is in keeping with the spoken word. We accept
3	the Hollywood 'senator' had a noble looking	image - as public relations prose sometimes puts
4	of the people. Television and the political	image shows what was actually happening to the minds
5	Gyorgy Kepes notes that we respond to the	images of the artist because their forms and harmonies
6	work he could, with a cheerful heart, compose	images while listening to the songs of the Faubourg.

Comments: see main text. All cases vary on one dimension from 'physical' (s2, s4 and s6) to 'conceptual' (s1, s3 and s5) and on another from 'public' (s2 and s3) to 'private' (s1). Note the 'quoted' use of 'image' ('as public relations prose puts ...') in case 3 which cannot but defy classification.

Appendix C

Altlists and the conative in DATR

As discussed in section 9.7.3, LR claim the conative alternation is associated with verbs with both 'contact' and 'motion' meaning components. So a conative sense is to be added to the `altlists` of just those verbs. Using the hierarchy built in Chapter 9, we add `Rub`, `Move` and `Touch`, the salient `vsemfeats` specifications, and `altlist` values capturing alternations already discussed, to the theory as follows:

```
Rub:<>                          == WIPE.

WIPE: <sem vsemfeats motion>    == yes
      <altlist>                 == remove put SURFACE-CONTACT.

SURFACE-CONTACT:
      <sem vsemfeats contact>   == yes.

Touch:<>                        == SURFACE-CONTACT
      <sem vsemfeats motion>    == no.

Move: <>                        == TRANSITIVE
      <sem vsemfeats contact>   == no
      <sem vsemfeats motion>    == yes
      <altlist>                 == refl TRANSITIVE.
```

(`Move` undergoes the reflexive alternation, "I moved" meaning "I moved myself".) From this we want to derive that `Rub` but not `Move` or `Touch` has `conative` on its `altlist`.

As in the DATR for gathering collocations in Chapter 8, we gather alternations in the `altlist` by adding members to a sequence as we inherit up the hierarchy. `Rub`, `Move` and `Touch` all inherit, directly or indirectly, from `TRANSITIVE`, so the rule for adding `conative` to the `altlist` is stated there. Any other alternations —unconditional or, like `conative`, conditional— which apply to all transitives will also be stated there, and after gathering alternations from `TRANSITIVE`, an `altlist` query will proceed to see whether there are any more to be gathered from still higher up the hierarchy, from `VERB`. `VERB` is to be found at the end of the `TRANSITIVE:<altlist>` line below (and several lines of the proof) because, after determining whether `conative` is to be added to the `altlist`, DATR will look to see whether there are any more additions to be made at `VERB`.

We add `conative` if and only if, when we go back from `TRANSITIVE` to the base node and ask the two queries, we get `yes` twice. The truth-table aspect of the problem is identified at the `AND` node with the path prefix `truth-table`.

```
TRANSITIVE: <altlist>  == CONATIVE:<if> VERB.

CONATIVE: <if>         == < AND:<truth-table
                                "<sem vsemfeats motion>"
                                "<sem vsemfeats contact>" > >
          <yes>        == conative.

AND:<truth-table yes yes>   == yes
    <truth-table>           == no.
```

The `altlist` for `Rub` inherits from the `altlist` for `TRANSITIVE`. The `altlist` for `TRANSITIVE` has the value for the query `CONATIVE:<if>` added to its front and then inherits from `VERB`. When a node, path or node-path pair occurs nested inside angle brackets on the right hand side of a DATR equation, we first evaluate the innermost query as if the surrounding angle brackets were not present, and then take the value that that query evaluates to and treat it as an element in the remaining structure. `CONATIVE:<if>` exploits the mechanism twice over. First, we evaluate the two `vsemfeats` queries at the global context node which will be the basic node for the verb. For `Rub`, both will give `yes`. These then become the second and third path elements following `AND:<truth-table>`, so the node-path pair at `AND` evaluates to `yes`. Back at `CONATIVE`, this is treated as a path which is inherited locally to give the value `conative`, which is placed in the `altlist` for `Rub`. The main steps in the derivation are shown below.

```
Rub:<altlist>
    TRANSITIVE:<altlist>
    CONATIVE:<if>  VERB
    CONATIVE: < AND:<truth-table
                        "<sem vsemfeats motion>"
                        "<sem vsemfeats contact>"  > > VERB
    CONATIVE: < AND:<truth-table
                      Rub: <sem vsemfeats motion>
                      Rub: <sem vsemfeats contact> > > VERB
    CONATIVE: < AND:<truth-table
                      yes  yes                     > > VERB
    CONATIVE: <yes>  VERB
    conative VERB
```

Whenever DATR does not find a value, a query fails. For the `altlist` mechanism to work, we must make it possible for a negative result to be returned without the query failing. Now the query `Move:<altlist>` should return the one-item sequence, `refl`. But if the part of the query determining whether `conative` is to be added to the `altlist` fails, then the whole query fails and no value is defined for `Move:<altlist>`. We put the matter right by returning the empty sequence as a negative result to an `altlist` query. We add the following equations:

```
VERB:<altlist>      == .
CONATIVE:<no>       == .
```

From a DATR perspective, we have defined a value for the node-path pair, albeit a null value. So now values are defined, the queries are defined, and we have:

```
Rub:<altlist>       = remove put conative.
Touch:<altlist>     = .
Move:<altlist>      = refl.
```

When introducing the `altlist` in section 8.6 we noted that it was a second order feature, conveying a different kind of information to that in the remainder of the theory. Now, there is also a technical difference. In the main theory, a negative result causes a query to fail, but in the `altlist` part of the theory, a negative result is represented as the empty sequence.

References

Aarts, J. & Meijs, W. (1990). *Theory and Practice in Corpus Linguistics.* Rodopi, Amsterdam.

Ahlswede, T. E. & Evens, M. (1988). Parsing versus text processing in the analysis of dictionary definitions. In *ACL Proceedings, 26th Annual Meeting*, pp. 217–224 New York.

Ahlswede, T. E. (1985). A linguistic string grammar for adjective definitions in w7. In Williams, S. (Ed.), *Humans and Machines: 4th Delaware Symposium on Language Studies*, pp. 101–128. Ablex, Norwood.

Amsler, R. A. (1980). *The Structure of the Merriam-Webster Pocket Dictionary.* Ph.D. thesis, University of Texas at Austin.

Apresjan, J. D. (1974). Regular polysemy. *Linguistics*, *142*, 5–32.

Atkins, B. T. S., Kegl, J., & Levin, B. (1986). Explicit and implicit information in dictionaries. In *Advances in Lexicography: Proc. Second Ann. Conf. of the UW Centre for the New OED*, pp. 45–65 Waterloo, Canada.

Atkins, B. T. S., Kegl, J., & Levin, B. (1988). Anatomy of a verb entry: from linguistic theory to lexicographic practice. *International Journal of Lexicography*, *1*(2), 84–126.

Atkins, B. T. S. (1987). Semantic ID tags: corpus evidence for dictionary senses. In *Proc. Third Ann. Conf. of the UW Centre for the New OED*, pp. 17–36 Waterloo, Canada.

Atkins, B. T. S. (1990). Presentation given to SALT workshop on corpus resources, Wadham College, Oxford, 3-4 Jan.

Ayto, J. R. (1983). On specifying meaning. In Hartmann, R. R. K. (Ed.), *Lexicography: Principles and Practice*, pp. 89–98. Academic Press, London.

Barwise, J. & Perry, J. (1983). *Situations and Attitudes.* MIT Press, Cambridge, Mass.

Beckwith, R., Fellbaum, C., Gross, D., & Miller, G. (1991). WORDNET: a lexical database organized on psycholinguistic principles. In Zernik, U. (Ed.), *Lexical Acquisition: Exploiting On-Line Resources to Build a Lexicon*, pp. 211–232. Lawrence Erlbaum, Hillsdale, New Jersey.

Benson, M., Benson, E., & Ilson, R. (1985). *The BBI Combinatory Dictionary of English.* John Benjamins, Amsterdam.

Berlin, B. (1978). Ethnobiological classification. In Rosch, E. & Lloyd, B. B. (Eds.), *Cognition and Categorization*, pp. 9–26. Lawrence Erlbaum, New Jersey.

Best, J. B. (1992). *Cognitive Psychology* (Third edition). West Publishing Company, St. Paul, Minnesota.

Bobrow, D. G. & Winograd, T. (1977). An overview of KRL, a knowledge representation language. *Cognitive Science*, *1*(1), 275–304.

Boguraev, B. K. & Briscoe, E. J. (Eds.). (1989). *Computational Lexicography for Natural Language Processing.* Longman, Harlow.

Boguraev, B. K. & Levin, B. (1990). Models for lexical knowledge bases. In *Electronic Text Research: Proc. Sixth Ann. Conf. of the UW Centre for the New OED*, pp. 65–78 Waterloo, Canada.

Boguraev, B. K. & Pustejovsky, J. (1990). Lexical ambiguity and the role of knowledge representation in lexicon design. In *COLING 90*, Vol. 2, pp. 36–41 Helsinki.

Boguraev, B. K. (1979). *Automatic Resolution of Word-Sense Ambiguity.* Ph.D. thesis, Cambridge University.

Bolinger, D. (1965). The atomization of meaning. *Language*, *41*, 555–573.

Briscoe, E. J., Copestake, A. A., & Boguraev, B. K. (1990). Enjoy the paper: lexical semantics via lexicology. In *COLING 90*, Vol. 2, pp. 42–47 Helsinki.

Byrd, R. J., Calzolari, N., Chodorow, M. S., Klavans, J. L., Neff, M. S., & Rizk, O. A. (1987). Tools and methods for computational lexicology. *Computational Linguistics*, *13*, 219–240.

Byrd, R. J. (1989). Discovering relationships among word senses. In *Dictionaries in the Electronic Age: Proc. Fifth Ann. Conf. of the UW Centre for the New OED*, pp. 67–80 Oxford.

Cahill, L. J. & Evans, R. (1990). An application of DATR: the TIC lexicon. In *Proc. ECAI-90*, pp. 120–125.

Cahill, L. J. (1990). Syllable-based morphology. In *COLING 90*, Vol. 3, pp. 48–53 Helsinki.

Calzolari, N. & Picchi, E. (1988). Acquisition of semantic information from an on-line dictionary. In *COLING 88*, pp. 87–92 Budapest.

Calzolari, N. & Zampolli, A. (1988). From monolingual to bilingual automated lexicons: is there a continuum?. *Lexicographica*, *4*, 130–144.

Calzolari, N. (1989). Lexical databases and textual corpora: perspectives of integration for a lexical knowledge base. In Zernik, U. (Ed.), *First International Workshop on Lexical Acquisition* IJCAI, Detroit.

Carpenter, B. (1990). Typed feature-structure constraints.. Lecture Notes, Leuven LLI Summer School.

Chierchia, G. & McConnell-Ginet, S. (1990). *Meaning and Grammar. An Introduction to Semantics.* MIT Press, Cambridge, Mass.

Church, K. & Hanks, P. (1989). Word association norms, mutual information and lexicography. In *ACL Proceedings, 27th Annual Meeting*, pp. 76–83 Vancouver.

Church, K., Gale, W., Hanks, P., & Hindle, D. (1991). Using statistics in lexical analysis. In Zernik, U. (Ed.), *Lexical Acquisition: Exploiting On-Line Resources to Build a Lexicon*, pp. 115–164. Lawrence Erlbaum, Hillsdale, New Jersey.

Church, K. W. (1989). A stochastic parts program and noun phrase parser for unrestricted text. In *ACL Proceedings, 27th Annual Meeting*, pp. 136–143 Vancouver.

Copestake, A. A. & Briscoe, E. J. (1991). Lexical operations in a unification-based framework. In Pustejovsky, J. & Bergler, S. (Eds.), *Lexical semantics and knowledge representation: ACL SIGLEX Workshop* Berkeley, California.

Copestake, A. A. (1991). Defaults in the LRL. In Briscoe, E. J., Copestake, A. A., & de Paiva, V. (Eds.), *Proc. ACQUILEX workshop on default inheritance in the lexicon* Tech. report 238, University of Cambridge Computer Laboratory.

Cowie, A. P. & Mackin, R. (1975). *Oxford Dictionary of Current Idiomatic English. Volume 1: Verbs with Prepositions and Particles.* OUP, Oxford.

Cowie, A. P. et al. (1982). *Oxford Dictionary of Current Idiomatic English. Volume 2.* OUP, Oxford.

Cruse, D. A. (1986). *Lexical Semantics.* CUP, Cambridge, England.

Dahlgren, K. (1988). *Naive Semantics for Natural Language Understanding.* Kluwer, Boston.

de Marcken, C. G. (1990). Parsing the LOB corpus. In *ACL Proceedings, 28th Annual Meeting*, pp. 243–251 Pittsburgh.

de Saussure, F. (1915). *Course in General Linguistics.* Published in 1959 by McGraw-Hill, New York.

DeRose, S. J. (1988). Grammatical category disambiguation by statistical optimization. *Computational Linguistics*, *14*(1), 31–39.

Dowty, D. R. (1979). *Word Meaning in Montague Semantics.* Reidel, Dordrecht.

Dowty, D. R. (1991). Thematic proto-roles and argument selection. *Language*, *67*(3), 547–619.

Evans, R. & Gazdar, G. (1989a). Inference in DATR. In *ACL Proceedings, 4th European Conference*, pp. 1–9 Manchester.

Evans, R. & Gazdar, G. (1989b). The semantics of DATR. In Cohn, A. G. (Ed.), *Proc. Seventh Conference of the AISB*, pp. 79–87 Falmer, Sussex.

Evans, R. & Gazdar, G. (1990). The DATR papers. Tech. rep. CSRP 139, School of Cognitive and Computing Sciences, University of Sussex, Falmer, Sussex.

Evans, R. (1992). Derivational morphology in DATR. In Coates, R. & Cahill, L. J. (Eds.), *Sussex Papers in general and computational linguistics presented to the LAGB*, CSRP 239. School of Cognitive and Computing Sciences, University of Sussex, Falmer, Sussex.

Evens, M. W. (Ed.). (1988). *Relational Models of the Lexicon.* CUP, Cambridge, England.

Fass, D. (1991). Met⋆: a method for discriminating metonymy and metaphor by computer. *Computational Linguistics*, *17*(1), 49–90.

Fellbaum, C. (1990). English verbs as a semantic net. *International Journal of Lexicography*, *3*(4), 278–301.

Fillmore, C. J. (1968). The case for case. In Bach, E. & Harms, R. (Eds.), *Universals in Linguistic Theory.* Holt, Rinehart and Winston, New York.

Fillmore, C. J. (1978). The organization of semantic information in the lexicon. In *Papers for the Parasession on the lexicon.* Chicago Linguistics Society.

Fillmore, C. J. (1989). Two dictionaries. *International Journal of Lexicography*, *2*(1), 57–83.

Firth, J. R. (1957). *Papers in Linguistics 1934–1951.* OUP, London.

Fisher, C., Gleitman, H., & Gleitman, L. R. (1991). On the semantic content of subcategorization frames. *Cognitive Psychology*, *23*(3), 331–392.

Francis, W. N. (1964). *A standard sample of present-day American English for use with digital computers. Report to the U.S. Office of Education on Cooperative Research, Project No. E-007.* Brown University, Providence.

Garside, R., Leech, G. N., & Sampson, G. (Eds.). (1987). *The Computational Analysis of English.* Longman, Harlow.

Garside, R. (1987). The claws word-tagging system. In Garside, R., Leech, G. N., & Sampson, G. (Eds.), *The Computational Analysis of English.* Longman, Harlow.

Gazdar, G., Klein, E., Pullum, G. K., & Sag, I. (1985). *Generalized Phrase Structure Grammar.* Blackwell, Oxford.

Gazdar, G. (1987). *The grammatical form of words: inflectional morphology and the lexicon.* ESRC RC/00/24/2003.

Geeraerts, D. (1990). The lexicographical treatment of prototypical polysemy. In Tsohatzidis, S. L. (Ed.), *Meanings and Prototypes: Studies in Linguistic Classification*, pp. 195–210. Routledge, London.

Gibbon, D. (1990). Prosodic association by template inheritance. In *Proc. Workshop on Inheritance in Natural Language Processing*, pp. 65–81 Tilburg. ITK.

Gibbs, R. (1983). Do people always process the literal meanings of indirect requests?. *Journal of Experimental Psychology: Learning, Memory and Cognition, 9*(3), 524–533.

Gibbs, R. (1984). Literal meaning and psychological theory. *Cognitive Science, 8*(3), 275–304.

Gibbs, R. (1986). Skating on thin ice: literal meaning and understanding idioms in conversation. *Discourse Processes, 9*, 17–30.

Givon, T. (1967a). Some noun-to-noun derivational affixes. Tech. rep. SP-2893, System Development Corporation, Santa Monica, California.

Givon, T. (1967b). Transformations of ellipsis, sense development and rules of lexical derivation. Tech. rep. SP-2896, System Development Corporation, Santa Monica, California.

Guthrie, L., Slator, B. M., Wilks, Y., & Bruce, R. (1990). Is there content in empty heads?. In *COLING 90*, Vol. 3, pp. 138–143 Helsinki.

Guthrie, J. A., Guthrie, L., Wilks, Y., & Aidinejad, H. (1991). Subject-dependent co-occurrence and word sense disambiguation. In *Proc. 29th Annual Meeting of the Association for Computational Linguistics*, pp. 00–00 Berkeley.

Halliday, M. A. K. (1966). Lexis as a linguistic level. In Bazell, C. E., Catford, J. C., & Halliday, M. A. K. (Eds.), *In Memory of J. R. Firth*, pp. 148–162. Longman, Harlow.

Hearst, M. A. (1991). Noun homograph disambiguation using local context in large text corpora. In *Using Corpora: Proc. Seventh Ann. Conf. of the UW Centre for the New OED*, pp. 1–22 Waterloo, Canada.

Hindle, D. (1989). Acquiring disambiguation rules from text. In *ACL Proceedings, 27th Annual Meeting*, pp. 118–125 Vancouver.

Hindle, D. (1990). Noun classification from predicate-argument structures. In *ACL Proceedings, 28th Annual Meeting*, pp. 268–275 Pittsburgh.

Hirst, G. (1986). Why dictionaries should list case structures. In *Advances in Lexicography: Proc. Second Ann. Conf. of the UW Centre for the New OED* Waterloo, Canada.

Hirst, G. (1987). *Semantic Interpretation and the Resolution of Ambiguity.* CUP, Cambridge, England.

Hofland, K. & Johansson, S. (Eds.). (1982). *Word Frequencies in British and American English.* The Norwegian Computing Centre for the Humanities, Bergen, Norway.

Hopcroft, J. E. & Ullman, J. D. (1979). *Introduction to Automata Theory, Languages and Computation.* Addison-Wesley, London.

Ilson, R. & Mel'čuk, I. (1989). English *bake* revisited. *International Journal of Lexicography*, *2*(4), 325–346.

Jakobson, R. & Halle, M. (1956). *Fundamentals of Language.* Mouton.

Jenkins, E. A. (1990). Enhancements to the Sussex Prolog DATR implementation. In Evans, R. & Gazdar, G. (Eds.), *The* DATR *Papers*, CSRP 139. School of Cognitive and Computing Sciences, University of Sussex, Falmer, Sussex.

Jensen, K. & Binot, J.-L. (1987). Disambiguating prepositional phrase attachment by using on-line dictionary definitions. *Computational Linguistics*, *13*, 251–260.

Johnson-Laird, P. N. (1983). *Mental Models.* CUP, Cambridge, England.

Katz, J. J. & Fodor, J. A. (1963). The structure of a semantic theory. *Language*, *39*, 170–210.

Kilbury, J., Naerger, P., & Renz, I. (1991). DATR as a lexical component for PATR. In *European ACL, Fifth Conference* Berlin.

Klavans, J. L., Chodorow, M. S., & Wacholder, N. (1990). From dictionary to knowledge base via taxonomy. In *Electronic Text Research: Proc. Sixth Ann. Conf. of the UW Centre for the New OED*, pp. 110–132 Waterloo, Canada.

Klavans, J. (1988). COMPLEX: a computational lexicon for natural language systems. In *COLING 88* Budapest.

Krishnamurthy, R. (1987). The process of compilation. In Sinclair, J. M. (Ed.), *Looking Up: An Account of the COBUILD Project in Lexical Computing*, chap. 3. Collins, London.

Lakoff, G. & Johnson, M. (1980). *Metaphors We Live By.* University of Chicago Press.

Lakoff, G. (1987). *Women, Fire and Dangerous Things.* University of Chicago Press.

Leech, G. (1981). *Semantics.* CUP, Cambridge, England.

Lesk, M. E. (1986). Automatic sense disambiguation using machine readable dictionaries: how to tell a pine cone from an ice cream cone. In *Proc. 1986 SIGDOC Conference* Toronto, Canada.

Levin, B. & Rappoport Hovav, M. (1991). Wiping the slate clean: a lexical semantic exploration. *Cognition*, *41*, 123–151.

Levin, B. (1991). Building a lexicon: the contribution of linguistics. *International Journal of Lexicography*, *4*(3), 205–226.

Levinson, S. C. (1983). *Pragmatics.* CUP, Cambridge, England.

Lyons, J. (1977). *Semantics*. CUP, Cambridge, England.

MacWhinney, B. (1989). Competition and lexical categorization. In Corrigan, R., Eckman, F., & Noonan, M. (Eds.), *Linguistic Categorization*, pp. 195–241. John Benjamins, Amsterdam.

Malakhovski, L. V. (1987). Homonyms in english dictionaries. In Burchfield, R. W. (Ed.), *Studies in Lexicography*, pp. 36–51. OUP, Oxford.

Marcus, M. P., Santorini, B., & Magerman, D. (1990). First steps towards an annotated database of English. Tech. rep., Dept Computer Information and Science, University of Pennsylvania, Philadelphia.

Marcus, M. (1980). *A Theory of Syntactic Recognition for Natural Language*. MIT Press, Cambridge, Mass.

Marr, D. (1982). *Vision*. W. H. Freeman, San Francisco.

Martin, W., Heymans, R., & Platteau, F. (1988). DILEMMA, an automatic lemmatizer. *Antwerp Papers in Linguistics*, *56*, 5–62.

Martin, J. H. (1990). *A Computational Model of Metaphor Interpretation*. Academic Press, Boston.

McRoy, S. W. (1992). Using multiple knowledge sources for word sense discrimination. *Computational Linguistics*, *18*(1).

Medin, D. L. (1989). Concepts and conceptual structure. *American Psychologist*, *44*, 1469–1481.

Meijs, W. (Ed.). (1987). *Corpus Linguistics and Beyond*. Rodopi, Amsterdam.

Mel'čuk, I. & Polguère, A. (1987). A formal lexicon in meaning-text theory. *Computational Linguistics*, *13*, 261–275.

Moore, R. C. (1989). Unification-based semantic interpretation. In *ACL Proceedings, 27th Annual Meeting*, pp. 33–41 Vancouver.

Nakamura, J.-i. & Nagao, M. (1988). Extraction of semantic information from an ordinary English dictionary and its evaluation. In *COLING 88*, pp. 459–464 Budapest.

Neff, M. S. & Boguraev, B. K. (1989). Dictionaries, dictionary grammars and dictionary entry parsing. In *ACL Proceedings, 27th Annual Meeting*, pp. 91–101 Vancouver.

Nunberg, G. & Zaenen, A. (1991). Systematic polysemy in lexicology and lexicography. Tech. rep. UCB/CSD91/618, Linguistics in lexicography subject group, Stanford University, California.

Nunberg, G. (1978). *The Pragmatics of Reference*. University of Indiana Linguistics Club, Bloomington, Indiana.

Ostler, N. & Atkins, B. T. S. (1991). Predictable meaning shift: some linguistic properties of lexical implication rules. In Pustejovsky, J. & Bergler, S. (Eds.), *Lexical semantics and knowledge representation: ACL SIGLEX Workshop* Berkeley, California.

Palmer, F. R. (1976). *Semantics*. CUP, Cambridge, England.

Pinker, S. (1989). *Learnability and cognition: the acquisition of argument structure*. MIT Press, Cambridge, Mass.

Pollard, C. & Sag, I. A. (1987). *An Information-Based Approach to Syntax and Semantics. Volume 1: Fundamentals.* CSLI Lecture Notes, No. 13. Chicago University Press, Chicago.

Pollard, C. & Sag, I. A. (1992). *An Information-Based Approach to Syntax and Semantics. Volume 2: Agreement, Binding and Control.* CSLI Lecture Notes. Chicago University Press, Chicago.

Pustejovsky, J. (1989). Current issues in computational lexical semantics. In *ACL Proceedings, 4th European Conference*, pp. xvii–xxiii Manchester.

Pustejovsky, J. (1991a). The generative lexicon. *Computational Linguistics*, *17*(4), 409–441.

Pustejovsky, J. (1991b). The syntax of event structure. *Cognition*, *41*, 47–81.

Quine, W. v. O. (1960). *Word and Object.* MIT, Cambridge, Mass.

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1972). *A Grammar of Contemporary English.* Longman, London.

Ravin, Y. (1990). Disambiguating and interpreting verb definitions. In *ACL Proceedings, 28th Annual Meeting*, pp. 260–267 Pittsburgh.

Robins, R. H. (1987). Polysemy and the lexicographer. In Burchfield, R. W. (Ed.), *Studies in Lexicography*, pp. 52–75. OUP, Oxford.

Rosch, E. & Lloyd, B. B. (Eds.). (1978). *Cognition and Categorization.* Lawrence Erlbaum, New Jersey.

Ruhl, C. (1989). *On Monosemy: A study in linguistic semantics.* SUNY Press, Albany, New York.

Rumelhart, D. E. & McClelland, J. L. (Eds.). (1986). *Parallel Distributed Processing: Explorations in the Microstructure of cognition.* MIT Press, Cambridge, Mass.

Russell, G., Ballim, A., Carroll, J., & Armstrong-Warwick, S. (1991). A practical approach to multiple default inheritance for unification-based lexicons. In Briscoe, E. J., Copestake, A. A., & de Paiva, V. (Eds.), *Proc. ACQUILEX workshop on default inheritance in the lexicon* Tech. report 238, University of Cambridge Computer Laboratory.

Shieber, S. (1986). *An Introduction to Unification-Based Approaches to Grammar.* CSLI Lecture Notes 4. CSLI, Stanford.

Sinclair, J. M. (Ed.). (1987). *Looking Up: An Account of the COBUILD Project in Lexical Computing.* Collins, London.

Slator, B. M. & Wilks, Y. A. (1987). Towards semantic structures from dictionary entries. Tech. rep. MCCS-87-96, Computing Research Laboratory, New Mexico State University, New Mexico.

Smajda, F. A. & McKeown, K. R. (1990). Automatically extracting and representing collocations for language generation. In *ACL Proceedings, 28th Annual Meeting*, pp. 252–259 Pittsburgh.

Small, S. L. (1980). *Word Expert Parsing: A Theory of Distributed Word-Based Natural Language Understanding.* Ph.D. thesis, Department of Computer Science, University of Maryland, Maryland.

Sparck Jones, K. (1986). *Synonymy and semantic classification.* Edinburgh University Press, Edinburgh.

Stock, P. F. (1983). Polysemy. In *Proc. Exeter Lexicography Conference*, pp. 131–140.

Talmy, L. (1985). Lexicalisation patterns. *Language Typology and Syntactic Description*, *3*, 57–149.

Taraban, R., McDonald, J. L., & MacWhinney, B. (1989). Category learning in a connectionist model: learning to decline the german definite article. In Corrigan, R., Eckman, F., & Noonan, M. (Eds.), *Linguistic Categorization*, pp. 163–193. John Benjamins, Amsterdam.

ten Hacken, P. (1990). Reading distinction in machine translation. In *COLING 90* Helsinki.

Turner, L., Grover, C., & Briscoe, E. J. (1989). The syntactic regularity of english noun phrases. In *ACL Proceedings, 4th European Conference*, pp. 256–263 Manchester.

Vendler, Z. (1967). *Linguistics in Philosophy*. Cornell University Press, Ithaca, New York.

Veronis, J. & Ide, N. M. (1990). Word sense disambiguation with very large neural networks extracted from machine readable dictionaries. In *COLING 90*, Vol. 2, pp. 389–394 Helsinki.

Vossen, P., Meijs, W., & den Broeder, M. (1989). Meaning and structure in dictionary definitions. In Boguraev, B. K. & Briscoe, E. J. (Eds.), *Computational Lexicography for Natural Language Processing*, pp. 171–192. Longman, Harlow.

Walker, D. E. & Amsler, R. A. (1986). The use of machine-readable dictionaries in sublanguage analysis. In Grishman, R. & Kittredge, R. (Eds.), *Analysing Language in Restricted Domains*. Lawrence Erlbaum, Hillsdale, New Jersey.

Way, E. C. (1991). *Knowledge Representation and Metaphor*. Kluwer, Dordrecht.

Weinreich, U. (1980). *On Semantics*. University of Pennsylvania Press.

Whittemore, G., Ferrara, K., & Brunner, H. (1990). Empirical study of predictive powers of simple attachment schemes for post-modifier prepositional phrases. In *ACL Proceedings, 28th Annual Meeting*, pp. 23–30 Pittsburgh.

Wierzbicka, A. (1987). Boys will be boys: 'radical semantics' vs. 'radical pragmatics'. *Language*, *63*(1), 95–114.

Wiggins, D. (1971). On sentence sense, word-sense and difference of word-sense. towards a philosophical theory of dictionaries. In Steinberg, D. D. & Jakobitz, L. A. (Eds.), *Semantics: An Interdisciplinary Reader*. CUP, Cambridge, England.

Wilensky, R., Mayfield, J., Chin, D., Luria, M., Martin, J., & Wu, D. (1988). The Berkeley UNIX consultant project. *Computational Linguistics*, *14*(4).

Wilensky, R. (1989). Primal content and actual content: an antidote to literal meaning. *Journal of Pragmatics*, *13*, 163–186.

Wilensky, R. (1990). Extending the lexicon by exploiting subregularities. In *COLING 90*, Vol. 2, pp. 407–412 Helsinki.

Wilks, Y., Fass, D., Guo, C.-M., McDonald, J., Plate, T., & Slator, B. (1989). A tractable machine dictionary as a resource for computational semantics. In Boguraev, B. K. & Briscoe, E. J. (Eds.), *Computational Lexicography for Natural Language Processing*, pp. 193–238. Longman, Harlow.

Wilks, Y. A. (1975). A preferential pattern-seeking semantics for natural language inference. *Artificial Intelligence*, *6*, 53–74.

Wittgenstein, L. (1953). *Philosophical Investigations*. Macmillan, London.

Zernik, U. & Dyer, M. G. (1987). The self-extending phrasal lexicon. *Computational Linguistics*, *13*, 308–327.

Zgusta, L. (1971). *Manual of Lexicography*. Mouton, The Hague.