

Word Grammar, Unification, and The Syntax of Italian Clitics

Massimo Salvatore Volino

PhD

University of Edinburgh

1990



To My Mother

Contents

Declaration	5
Abstract	7
1 General Introduction	8
1.1 Goals Of Thesis	8
1.2 Linguistic Traditions	9
1.3 Current Research	11
1.4 Introduction To The General Framework	12
1.4.1 The Appeal Of Word Grammar	12
1.4.2 The Homogeneity Hypothesis	13
1.4.3 Grammars Without Components	17
1.4.4 Grammars Without Levels	18
1.5 Brief Overview	19
2 Introduction To Word Grammar (1984 Framework)	21
2.1 Preliminaries	21
2.1.1 The ISA Hierarchy and Inheritance in Word Grammar	23
2.1.2 The Co-Word Relation and Dependency	26
2.1.3 Summary	33
2.2 Problems With The Approach	34
2.2.1 A Grammar Fragment	35
2.2.2 The Proliferation Of Entities	38

2.2.3	The Proliferation of Relations	40
2.2.4	Problems With The Formation of Propositions	41
2.2.5	Further Problems	42
3	Word Unification Grammar	46
3.1	Over-Diffuseness	49
3.2	The Isolation of Primitives	50
3.2.1	Relations in Word Unification Grammar	50
3.2.2	Entities in Word Unification Grammar	52
3.3	Propositions in Word Unification Grammar	56
3.3.1	Taxonomic Propositions	56
3.3.2	Subcategorisation Propositions	61
3.4	Inference in Word Unification Grammar	63
3.4.1	The Operating Principle For ISA	63
3.4.2	Inheritance in Word Unification Grammar	64
3.4.3	Summary	67
3.5	A Formalisation Of Dependency	68
3.6	Definition of Grammar	70
3.6.1	A Word Unification Grammar Fragment of English	75
4	Clitics	85
4.1	Descriptive Generalisations About Clitics	86
4.1.1	Bound Clitics	86
4.1.2	Simple Clitics	86
4.1.3	Special Clitics	87
4.1.4	Differences In The Order Of Clitics Relative To The Host	88
4.1.5	The Historical Perspective	89
4.1.6	The Classification Of Clitics	89
4.2	The Pronominal System of Italian	90
4.2.1	The Strong Pronominal Forms	91

4.2.2	The Weak or Clitic Pronominal Forms	92
4.2.3	The Form Of The Clitics In Clusters	93
4.3	The Syntax of Italian Clitics	93
4.3.1	The Non-Exceptional Behaviour Of Clitics	94
4.3.2	Clitic Climbing	96
4.3.3	Object Agreement	99
4.3.4	Change Of Auxiliary	101
4.3.5	Interaction Of Clitic Climbing, Change Of Auxiliary And Object Agreement	102
4.4	The Relative Ordering Of Clitics	105
4.4.1	Summary and Benchmarks	106
5	The Syntax of Clitics In Word Unification Grammar	107
5.1	Previous Analyses	107
5.2	Clitics As Words	111
5.3	Problems Inherent In The Framework	114
5.4	The Analysis And Possible Solutions	115
5.4.1	A Visitor Account	115
5.4.2	A Non-Visitor Account	118
5.5	Bicephality or Double Headedness	119
5.6	Bicephality in Word Unification Grammar	121
5.6.1	The Non-Exceptional Behaviour Of Clitics	123
5.6.2	Clitic Climbing	125
5.6.3	Double Dependencies	126
5.6.4	The Relative Ordering Of Clitics	128
5.6.5	Clitic Climbing Revisited	133
5.6.6	Object Agreement	138
5.7	Advantages and Disadvantages of The Approach	139

6	Word Unification Grammar and Categorical Grammar	143
6.1	Categorical Grammar	144
6.2	Comparison	146
6.2.1	Lexicalism	146
6.2.2	Constituency	147
6.2.3	Classification	148
6.2.4	Heads	148
6.2.5	Further Differences	154
6.3	A Categorical-Like Word Grammar	155
6.3.1	Clitic Placement	155
6.3.2	The Non-Exceptional Behaviour of Clitics	157
6.3.3	Clitic Climbing	159
6.3.4	Object Agreement	161
6.4	Conclusions	163
	Implementation	165
	Bibliography	175

Declaration

I declare that this thesis has been composed by myself and that the research reported herein has been conducted by myself unless otherwise indicated.

Massimo S. Volino

Edinburgh, 1st July 1990

Acknowledgments

There are many people who I would like to thank for the various ways in which they have helped me arrive at this thesis. Foremost amongst them is my mother, my brothers and sisters, Paul, Sandra, Marcella, Marco, and Danny, and the rest of my family, upwards, downwards and sideways.

There are also my friends some within the field and some that share other interests. All have helped and supervised me in their own way, and thanks are due accordingly. These include, Guy Barry, Nick Braisby, Stephana Broadbent, Biff and Toots, Mike The Bike, Jo Calder, Francesco Cara, Humph Couchman, Mark Dance, Paul Dance, Roger Dodge, John Eglin, Martin Evans, Mark Hepple, Dick Hudson, Jane Hunter, Keith Johnston, Ewan Klein (official), Sandy MacNeil, Gill Martin, Donald Mitchell, Greg Mitchell, Clem Moore, Jeremy Olesen, Irene Orr, Janie Park, Jim Park Jnr, Jim and Buddy Park, Grandma Park, Mike Reape, Jim and Jan Rule, Udi and Ayellet Rahat, Steve praise His Majesty the Lion of Judah, Antonio Sanfillipo, Jerry Seligman, Callum Smith, Jeff Stewart, Peter Swanson, Isobel Vance, Mark Wall, Pete Whitelock (official), Scott, Tracy and the wee yun, Deidre Wilson, Dave Woodhouse.

Of course a special debt is owed to Dick Hudson not only for the theory of Word Grammar but for giving me endless amounts of his free time, tolerating my phone calls out of the blue, commenting on practically all drafts of everything I've ever written much of which rightly ended up in the dustbin. His ability as a linguist is only surpassed by his ability as a teacher and I would like to take this opportunity to thank him.

Abstract

Modern linguistics is currently replete with competing theories, all with differing goals and intentions. This is not an altogether desirable situation. The aim of this thesis is to develop one of these theories, namely Word Grammar (Hudson 1984a), with this in mind. After an exposition of the theory, which should leave the reader with a clearer idea of the workings of the theory of Word Grammar, I will be concerned to put the intuitions behind the theory on as formal a footing as possible. This will involve the development of yet another formalism. However, this formalisation will involve the use of standard techniques. Extensions to the grammar, where necessary, shall be made with devices now current in the field such as Unification. In such a way, I hope to bring Word Grammar more into line with other formalisms, thus aiding a convergence rather than a divergence of theories. As part of the test for this new formalism it will then be applied to the problem of clitic placement in Italian.

Chapter 1

General Introduction

1.1 Goals Of Thesis

The overall aim of this thesis is to investigate and develop the theory of Word Grammar (Hudson, 1983a - b, 1984a - b, 1985a - h, 1986a - g, 1987a - d, 1988a - d, 1989, 1991) and the central notions of dependency on which it is based. One major hindrance in this exercise is a lack of clarity and explicitness with some of the fundamental concepts of the theory. As a result, a significant part of the investigation will involve making the theory more precise and explicit. The advantages of doing this are summed up neatly by Chomsky,

“The search for rigorous formulation in Linguistics has a much more serious motivation than mere concern for logical niceties ... Precisely constructed models for linguistic structure can play an important role, both negative and positive, in the process of discovery itself. By pushing a precise but inadequate formulation to an unacceptable conclusion, we can often expose the exact source of this inadequacy and, consequently, gain a deeper understanding of the linguistic data. More positively, a formalized theory may automatically provide solutions for many problems other than those for which it was explicitly designed.”

(Chomsky, 1957, preface)

In addition to the points made by Chomsky, it is only when formalised that a theory becomes independent of its author and enters into a more public domain. This thesis, however, will not be concerned solely with formalisation. In order to take up the point made by Chomsky above of “pushing a precise but inadequate formulation” of a theory, I will not only test the formalisation developed against a body of data but will also make

a comparison between that formalisation and a related framework, namely Categorical Grammar. Briefly then, this thesis will attempt to achieve its overall aim mentioned in the opening paragraph, by attaining the following sub-goals:

- (1) a formalising Word Grammar
- b formalising dependency
- c testing the formalisation against a specific body of data
- d comparing Word Grammar with a related framework
- e implementing the formalism

As Word Grammar is a little known theory, it will be useful to give some background details by way of introduction, before considering the goals set out in (1). This will occupy most of the rest of this chapter.

1.2 Linguistic Traditions

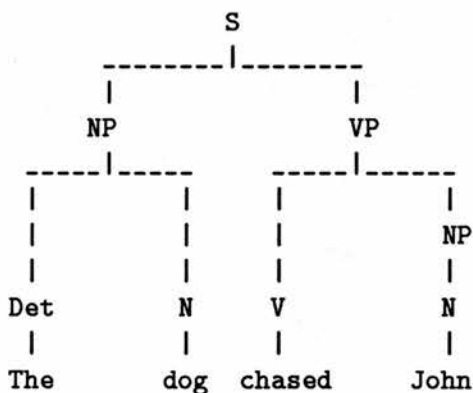
Probably still the most influential tradition in modern linguistics, certainly since the 1930's, is the *constituency* approach, popularised by Bloomfield, (1933) and continued through Harris (1951) to Chomsky (1957). In the constituency tradition, the study of syntax has gradually replaced phonetics which had previously been the dominant branch of linguistics. This trend continued to the situation under Chomsky (1965) where syntax is seen as the single most important aspect in the study of language, prior to, and autonomous of, all other aspects of the study of language. As a consequence, most recent research, especially work done in the Chomskyan tradition, has been directed towards a greater understanding of syntax.

In general terms, the basis for the analysis of syntactic structure in the constituency tradition is the relationship of parts to a whole. What this means is that in building up the structure of a sentence under analysis, we go from nouns to noun phrases, from adjectives to adjective phrases, from prepositions to preposition phrases, from verbs to verb phrases and so on. Further, in the analysis of simplex sentences, the constituency tradition would take the subject of a sentence to be a noun phrase.

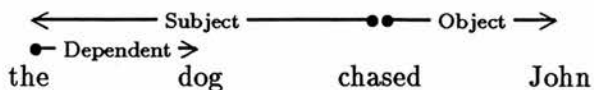
But there is another tradition. This is the *dependency* tradition (Tesnière, 1959). Though it receded in popularity after the Bloomfieldians, it is now arousing greater and

greater interest and slowly re-establishing itself. The basis for the analysis of syntactic structure in this tradition is the relationship that a single word has to another single word. That is to say that in the dependency tradition we are dealing with the relationship of one part to another. So, in contradistinction to the constituency tradition, dependency analyses relate prepositions to nouns, nouns to verbs and so on. But this part:part relationship is not symmetrical; one part or word will be the *head* word and as such will determine many properties of the other word as will become apparent below. The remaining word is called the *dependent* (of the head word) and it may also exert an influence back upon the head, albeit to a lesser degree. In the analysis of simplex sentences the dependency tradition would take the subject of a verb to be a noun and not a noun phrase. The verb would be the head of the noun and the noun would be the dependent of the verb. These two traditions would result in the following analyses for the sentence, *The dog chased John*:

(2) A Constituency Analysis



(3) A Dependency Analysis



The diagrams in (2) and (3)¹ contain much implicit information that is relevant to how the respective structures were arrived at, but which need not concern us here. What the diagrams clearly illustrate is the points of difference noted above; that in constituency we

¹I would like to thank Mark Hepple for writing the \LaTeX macro to draw dependency diagrams

are concerned with the part:whole relation as in the case of the V and NP that form the VP, whereas in dependency we are more concerned with the part:part relation as in the relationship of the verb to the object noun. Also clearly illustrated is the fact that in (2) the subject of the verb is the NP whereas in (3) it is a single word and not a phrase.

1.3 Current Research

These two approaches seem quite different on the face of it. But the two traditions are not completely incompatible. In fact, one of the most promising developments in linguistics over the past twenty years has been the way in which both traditions have converged in many respects. The X-bar theory of constituency structure which was introduced by Jackendoff (1977) is a case in point. Central to X-bar theory is the idea that each phrase or construction has an element which is crucial to that phrase. The introduction of the X-bar notation was intended to capture this close connection between a phrase and one of its elements. Thus the central element of a noun phrase is a noun, the central element of a preposition phrase is a preposition, the central element of a verb phrase is a verb and so on. It should not be surprising that the nodes that are picked out correspond more or less systematically with the heads of the head-dependent relationships of dependency.²

Another example of this mutual convergence of traditions is the tendency across theories towards greater and greater lexicalisation of grammatical information. This can be seen in the theory of Government and Binding (Chomsky, 1981) where the role of the lexicon has become increasingly important and in *Categorial Grammar(s)* (Adjukiewicz 1935), where much grammatical information is to be found in the lexical category. In spite of this mutual convergence, constituency has remained the dominant tradition in most current research.

Despite the dominance of constituency, dependency flourishes in many areas around the world. These areas include Starosta's Lexicase group in Hawaii (Starosta, 1982, 1988) and Sgall's group in Prague (Sgall, Hajicova, Panevova, 1986, Sgall and Hajicova, 1986).

²cf. also the use of the 'H' notation in GPSG (Gazdar et al. 1985)

Other groups include the group headed by Kunze (Kunze, 1987) in East Germany, Hellwig's Eurotra group in Heidelberg (Hellwig, 1986) and Valkonen's group in Finland, (Valkonen et al. 1987). In this country there is the work on dependency based theories of phonology by Anderson at Edinburgh, (Anderson, 1977, 1979, 1984) and, of course, the theory under discussion, Hudson's Word Grammar (op. cit.) in London.

The convergence noted above is continued by researchers in both traditions with theorists from both traditions borrowing liberally from each other. For example, dependency theorists have found it problematic to do without constituency, especially when it comes to accounting for phenomena such as co-ordination and gapping. By the same token, constituency theorists, to varying degrees, have found it necessary to incorporate the dependency notion of head. It is within this general setting of convergence that this thesis is written. I will now turn to a general discussion of the theoretical background to Word Grammar.

1.4 Introduction To The General Framework

The first problem encountered by anyone working within the Word Grammar framework is that the term, *Word Grammar*, is used for the more general theoretical claims as well as for the linguistic (sub-)part with which the theory mainly deals. It is important, even if merely for the sake of clarity, to keep the two distinct. I will use the term *General Theory*, in connection with the general cognitive/epistemic structure, and simply *Word Grammar* (as opposed to a Word Grammar grammar) for the linguistic specific sub-part of that declarative knowledge. I will begin the discussion by looking first at what might be called the *appeal* of Word Grammar.

1.4.1 The Appeal Of Word Grammar

Set against the backdrop of the dominance until the early 1980's of Chomskyan linguistics, amongst the most striking and appealing features of Word Grammar are the following:

- (4) a homogeneity of linguistic and non-linguistic knowledge
- b absence of components
- c (lexical) surface-based syntax and semantics
- d hierarchically organised lexicon
- e dependency syntax

The reason these features are somewhat striking is that when they were proposed, they went against many of the mainstays of the Chomskyan approach. They are appealing because in going against the Chomskyan tradition, these features alleviated many of the uneasy feelings about that approach (some of which will be discussed below), that were already beginning to appear by the end of the seventies. Though many of the attributes in (4) may be found separately elsewhere in current research into linguistics, this particular set, as we shall see, uniquely defines Word Grammar. I will discuss each of these features in turn below. In doing so, I will attempt not only to justify their inclusion in the theory on independent grounds, as far as possible, but additionally to give some flavour of the interesting ways in which they interact. In many ways (4a) may be seen as the psychological “adhesive” which motivates and justifies much of the approach, and at the same time underpins many of the other features in (4). As such, it is to this that I will turn to first.

1.4.2 The Homogeneity Hypothesis

Probably the most distinctive claim of Word Grammar is that knowledge of language is essentially indistinguishable from other kinds of declarative knowledge. This I will call the *Homogeneity Hypothesis*.³ This term will be used in contrast to the *Innateness Hypothesis* of Chomsky (but the two are not necessarily mutually exclusive).

The Homogeneity Hypothesis does not necessarily imply that knowledge of language is itself indistinguishable from other bodies of knowledge, nor that there is no place in the theory for modular approaches to the study of the mind. Quite the contrary. The only claim being made is that the way linguistic knowledge is stored, used, and updated, and the mechanisms that allow such operations to take place are not significantly different from the way other declarative kinds of knowledge (or information) are stored, used, and

³Thanks to Mark Hepple for this term, though its use is strictly my own

updated. Of course, the Homogeneity Hypothesis cannot be proven one way or the other at this time, but what I will do is argue for it along the lines given in (5) below:

- (5) a There are certain methodological and common-sense reasons why an approach based on this hypothesis is worth considering.
- b Given that the Homogeneity Hypothesis can be motivated, certain desirable consequences (that are actually in line with most current research) about the nature and organisation of the ensuing grammar follow naturally and to a certain extent automatically. These are principally the features (4b) to (4d).
- c When coupled with the features in (4) above, theories incorporating the Homogeneity Hypothesis may provide insightful and interesting analyses of natural language phenomena and thus by implication help to advance the study of the structure and organisation of non-linguistic knowledge.

The points made in (5) are related to the goals of this thesis as set out in (1). As such, justification for the Homogeneity Hypothesis will run beyond the end of this chapter.

Let us now turn to (5a) above. We can see the Homogeneity Hypothesis as towards one end of the nature/nurture debate; the one that places the emphasis on the nurture side of the coin. I take it that the relevant questions in the nature/nurture debate include at least the following:

- (6) a What aspects of language should be handled by general learning devices?
- b What aspects of language should be handled by innate devices specific to language?
- c What aspects of language should be handled by learning devices specific to language?
- d What aspects of language should be handled by more general innate devices?

Though, obviously, I will not attempt to answer these questions here, the point is what possible light might be thrown onto these problems by any given theoretical stance? The general approach adopted by Hudson is apparent from the following:

“If language is a mental phenomenon then we must inquire into its relations to other mental phenomena. In particular how similar is the structure of language to the structure of other types of knowledge? ... There must of course be some difference ... otherwise the two would not be distinct, but we need not assume any differences beyond those which lead us to distinguish them in the first place.”

(Hudson, 1984a, pp. 35 to 36)

But we can go further than this common-sense approach. Given that even those most strongly opposed to explanations of the mind in terms of innate devices still admit that some degree of mental structuring appears to be innate, the relevant question seems to be *how much* is innate. From this viewpoint it seems methodologically sounder (irrespective of the facts of the matter about what is or is not innate) to take the Homogeneity Hypothesis as an initial hypothesis until that approach fails. There are several reasons for this; firstly, to proceed from the premise that there is an innate language faculty in order to demonstrate that same premise, or even discover the precise nature of such a faculty, seems unsound. Secondly adopting the Homogeneity Hypothesis is motivated even further by the fact that despite much research within frameworks that adopt some form of the Innateness Hypothesis, there has been as yet no clear characterisation of what is or what is not admissible as an innate device. The result of not having such a characterisation is that potentially anything is permissible as an innate mechanism. This implies that there is no sense in which we could ever fail to find some (abstract) device which accounted for some body of data. Further, it is not the case that either by the success or the failure of the Innateness Hypothesis that we necessarily learn, with any kind of certainty, any more about the nature of language or, by implication, the nature of mind as these questions are characterised in (6) above.

In contrast to the Innateness Hypothesis, an advantage of adopting the Homogeneity Hypothesis would be that if that approach did then prove false, because an innate device was required to account for some linguistic phenomenon, then it would do so only because all attempts to account for that data by non-specific, non-innate means would have failed. At this stage we will at least be more confident of having found a truly innate mechanism. As a result we would then have a clearer idea about answers to (6a) and (6b) and hopefully some insights into (6c) and (6d).

There are further reasons for not adopting the Innateness Hypothesis too readily. The first of these come from the genetic epistemology of Piaget. It could be reasonably argued that some form of the Homogeneity Hypothesis is present in the genetic epistemology of

Piaget. In justifying his approach, Piaget makes the following point:

“...recourse to innate factors never resolves problems; it merely passes them onto biology”

(Piaget, 1964, p. 117)

If we agree with Piaget that the Innateness Hypothesis is not so much an explanation but rather a burden-passer, and that additionally what we are interested in is explanation of language in psychological terms, it seems methodologically sounder to limit the number of innate devices that we postulate and therefore to keep as much of the burden of explanation in the realm of psychology as possible. From this it seems reasonable to conclude that the now standard view of linguistics as a branch of psychology, first popularised by Chomsky, seems more compatible with the Homogeneity Hypothesis, which attempts to keep the burden of explanation within psychology, than with the Innateness Hypothesis which defers the burden to the realms of biology. Eventually, I would hope that some amalgamation of ideas from both the Word Grammar and the genetic epistemology approaches would be possible but that must be left as the subject of a different investigation as it is beyond the scope of the present one.

Related arguments that suggest adopting the Homogeneity Hypothesis come from assumptions about evolution. Chomsky has written extensively on the modularity of mind (Chomsky, 1968, 1975, 1980, 1986) comparing the language faculty with a physical organ both in terms of the evolution of the species and individually in terms of a pre-determined maturation. However, if we investigate this claim carefully and look at the physical apparatus of speech, for example, we find that all the organs of speech have a primary (or at least previous) function which has nothing to do with speech at all. These include lips, teeth, tongue, palate, and so on. When seen in terms of evolution, we seem to be forced to acknowledge that, at some stage at least, the primary or prior function of these organs must have been dominant. It seems likely then that the speech function was developed on top of the primary function.

As language or speech seems to be an evolutionarily late development physically, it seems natural to assume that the mental devices that contributed to this physical process

developed at least in tandem with the physical developments and hence later as well. It also seems natural to assume that these mental devices, like their physical counterparts, developed on top of, or at least made use of, existing structures and devices to a significant degree. This again seems to suggest that if there is a language organ it should in many significant respects share features with other modules. A similar point is made by Hudson:

“If we assume that our Being had designed the general conceptual workings of the human mind before starting on language (which is surely a reasonable assumption) it is hard to see any reason why the Being should develop a completely new set of structures for dealing with linguistic expressions. On the contrary, it would be most natural if the Being exploited the existing structures as far as was possible, and only introduced the minimum number of new types of structure.”

(Hudson, 1985g, p. 286)

This in general terms, then, can be described as the motivation for adopting the Homogeneity Hypothesis. We will now see how the view that there is essentially no difference between linguistic and non-linguistic knowledge leads more or less directly to many of the other distinctive features of Word Grammar mentioned in (4).⁴

1.4.3 Grammars Without Components

Perhaps one of the more obvious consequences of the Homogeneity Hypothesis is how this view naturally leads us to a grammar in which it is harder to make many of the distinctions which form the boundaries of the components of grammar. This in effect leads down the path towards a componentless grammar. This situation arises partly due to the fact that it would be inconsistent for Word Grammar or any other theory to claim, on the one hand, that there is no difference between linguistic and non-linguistic knowledge, and, on the other, that within the domain of linguistic knowledge there are major differences between phonology, morphology, syntax, and so on. It seems, then, that Word Grammar is committed to as few components of grammar as possible.

However, one might also want a grammar with as few components as possible for independent reasons. The fact that many of the distinctions between components of gram-

⁴In the following subsections I will adopt the Chomskyan terminology for components and levels, as it is against the Chomskyan backdrop that the discussion revolves.

mar are notoriously difficult to draw, despite enormous amounts of resources put into the project, is a case in point. Further, in many areas of the study of language, we seem obliged to straddle (at least) two components. The syntax/morphology boundary is a case in point.

This radical approach to grammar also goes against the Chomskyan preconception that components should be prevented from interacting. In Word Grammar, there is no autonomous syntax (or autonomous anything else for that matter). The autonomous syntax approach was yet another attitude which grew out of the Innateness Hypothesis and was a source of uneasiness to many linguists both within and outside that framework. In fact, in this respect, Word Grammar is in line with most current trends which also advocate that syntax and semantics run interactively in parallel.

It is important to remember that a componentless grammar, or any of the other features in (4) is not necessarily required by approaches that adopt the Homogeneity Hypothesis. In fact it will prove more convenient in the discussion to adopt the terminology of components. The point is that questions relating to which component of grammar we are dealing with will be of no interest and have little or no relevance. The practical effect of a grammar without components is that we effectively have a lexico-grammatical theory; that is to say a theory where much grammatical information is stored in the lexicon, again in line with current trends.

1.4.4 Grammars Without Levels

A second, less obvious, consequence of approaches based on the Homogeneity Hypothesis is a grammar where there is little or no place for analyses that rely on underlying representations and invisible multiple abstract derivations (or the use of levels). This is a less obvious consequence because the reasoning behind it is a little more subtle.

As is generally agreed, underlying representations and the abstract derivations they necessitate are by definition not accessible, in the sense of being present in the data, to a child. Instead, they constitute (in some form) part of the biological endowment. All that is accessible to a child are the surface representations. So if the underlying representations

are not accessible and therefore can not in any real sense be “acquired”, their existence must be facilitated by some appeal to the Innateness Hypothesis. Again, that is not to say that the Homogeneity Hypothesis precludes such analyses, merely that any such statements that are made have to be justified on the basis of the direct evidence of the data, and in accordance with other general cognitive abilities for making abstractions in other domains. Only in this way will we have kept faith with the Homogeneity Hypothesis. The point is that if we were to adopt the use of levels or underlying representations and still maintain the Homogeneity Hypothesis we would be obliged to motivate the use of levels for other non-linguistic domains as well. Where this can not be motivated, the result will be the anchoring of analyses firmly to the surface data.

It seems, then, that if we wish to maintain consistently the Homogeneity Hypothesis we are forced to find and demonstrate other methods whereby not only acquisition but analysis of the complex phenomena acquired could take place. What this means practically is finding some way in which a child could possibly acquire a language based on the data alone (probably in accordance with some general learning mechanisms) and a grammar that is essentially surface-based.

In summary then, the Homogeneity Hypothesis seems to lead naturally to a surface-based monostratal lexico-grammatical theory, a position that is in line with much current research into grammatical formalisms in both linguistics and computational linguistics.

1.5 Brief Overview

The rest of this thesis is organised along the following lines:

Chapter 2 will present a more detailed description of the Word Grammar framework and will attempt to explain some of the fundamental notions. Some problems with the Word Grammar framework will then be discussed.

In Chapter 3, I will introduce some modifications to the framework. In doing so I will attempt to explore and, where necessary, extend Word Grammar as a theoretical framework for the analysis and understanding of natural language phenomena. The modified

framework will be called *Word Unification Grammar*. The move to Word Unification Grammar also facilitates comparison with other linguistic formalisms and in particular with the family of Categorical and Unification-based Grammars. It also has the advantage of facilitating the incorporation of many current ideas about linguistic formalisms into the framework. This chapter will see the completion of goals (1a) and (1b).

After the exposition of Word Unification Grammar, Chapters 4 and 5 will attempt to evaluate Word Grammar's usefulness as a theory for linguistic description. The approach taken here will be to examine a body of problematic data in Chapter 4 and to try and develop an account for it in Chapter 5. The data chosen will provide a significant test for some central claims of the theory. A second advantage of this approach is that it facilitates comparison with other mainstream grammatical frameworks at the level of linguistic coverage and formalisation of linguistic intuitions. This will see the completion of the goal in (1c).

In Chapter 6, a comparison will be made with Word (Unification) Grammar and Categorical Grammar. One form this will take will be the development of a categorical grammar that is similar to a Word Unification Grammar. This will see the completion of goal (1d). Chapter 6 will also contain the problems and conclusions.

The appendix will contain an implementation of Word Unification Grammar where we will see most clearly the benefits of the modifications made to Word Grammar in Chapter 3, which facilitated an assessment of Word (Unification) Grammar.

Let us now take a closer look at Word Grammar.

Chapter 2

Introduction To Word Grammar (1984 Framework)

2.1 Preliminaries

We have now seen some of the motivation for the general theoretical setting of Word Grammar, and some of the general advantages of such an approach. Before we proceed with the introduction to Word Grammar proper, a few opening comments would be in order. Detailed discussion of the theory may be found in Hudson (op. cit.). I will be concerned with expounding only that syntactic part which is essential and relevant to an understanding of the phenomenon under discussion in Chapter 4. Following the exposition there will be a discussion of some of the problems with the theory. As should be apparent from Chapter 1, this will be done more in the spirit of constructive criticism than destructive diatribe.

The goal of Hudson's enterprise, as was made clear in the opening chapter, is to develop a general account of knowledge and knowledge representation, within which linguistic knowledge will form a special case. In section 1.4 this position was motivated and the general consequences of such an approach were discussed. The aim of *this* chapter is to see how the intuitions behind the Homogeneity Hypothesis may be incorporated into a grammar: Hudson begins with the following statement about the nature and organisation of Word Grammar ;

“A language is a network of entities related by propositions”

(Hudson, 1984a, p. 1)

This statement, though simple, still gives us four concepts to investigate; namely *network*, *proposition*, *relation*, and *entity*. We will now look at each of these concepts in turn. The view of language as a network is a claim of such generality that if taken at face value it seems to say nothing contentful at all. It would be difficult to see how it could not be applied to almost any theory of language. What makes Word Grammar distinctive, and gives the above statement content is the nature of this network and the use it is put to. Briefly, the Word Grammar network derives its nature from the use of (declarative) propositions to express (grammatical) facts. The propositions express two place relations between entities. I will look at the entities first and then move on to the relations.

The entities that we will be concerned with (details of which may be found in Hudson, 1984a, pp. 4 - 7), are all linguistic and all make reference to some notion of *word*. Exactly what ontology would be required by the General Theory is a moot point, but one that is outside the scope of this thesis. The set of linguistic entities, which is small and fixed, includes the following, as listed in Hudson (1984a);

- (1) a individual (independent) words
- b parts of words
- c strings of words
- d the meanings of words
- e the situations in which the words are uttered

As we will see later, under one interpretation at least these do not exhaust the linguistic possibilities. I will not, however, take the exhaustive listing of all the possible entities as problematic, especially when we bear in mind that the work carried out on non-linguistic entities is minimal. Instead I will be satisfied with identifying all and only those entities that are required to account for the phenomena in Chapter 4.

This brings us to the set of relations, which is also small and fixed. All relations in Word Grammar are two place relations. As with the entities, more detail may be found in Hudson (1984a, pp. 7 - 12). The set of relations includes the following;

- (2) a composition
- b model-instance/ISA
- c co-entity/co-word
- d referent
- e (utterance) event

A proposition is formed by making two entities enter into one of the binary relationships (We will see precisely how later). In the theory as presented in Hudson (1984a) there are no type restrictions on which entities can form propositions with which relations. As we will see below, it will prove necessary to impose such restrictions if we are to restrict our grammar to making only meaningful statements. None of the relations involved in Word Grammar is specifically linguistic, but obviously some of the entities, such as those given in (1), are.

The above, then, gives a brief description of what is meant by the opening quotation. In the next section I will take a closer look at some of the relations mentioned in (2) above. Given that this thesis will concentrate mainly on syntactic relations, I will be concerned for the most part with the model-instance and co-word relations which will occupy the next two subsections. Associated with these two relations are certain operating principles which will also be discussed.

2.1.1 The ISA Hierarchy and Inheritance in Word Grammar

The predominant relation that governs the organisation of the network in Word Grammar and hence plays a major role in all parts of the grammar is the *model-instance* relation. This is more commonly known as the *ISA* relation and in the interest of clarity it is this term that I will employ. Though the ISA relation has received many interpretations, the basic intuition behind it and the sense in which it is similar to the model-instance relation, is as a sort of subset/subsumption relation. This relation, especially when it is used in conjunction with defaults, is familiar to researchers in Knowledge Representation in the form of semantic networks (Quillian, 1968, Brachman, 1979, 1983, 1985, Etherington and Reiter, 1983, Minsky, 1975, Touretzky, 1986, Woods, 1975) and in the form of higher order (non-monotonic, auto-epistemic and so on) logics (Doyle, 1978, Hayes, 1977a - b,

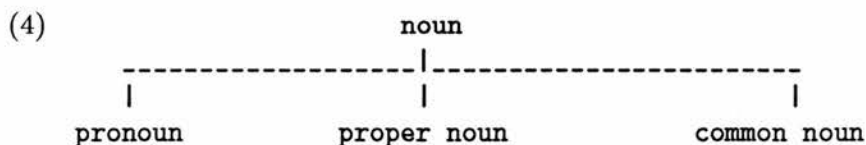
McDermott and Doyle, 1980, Moore, 1983, Reiter, 1980 and Turner, 1984).

Such ideas have been applied to both problem solving and natural language understanding in Knowledge Representation. They are also common in Linguistics in the form of the Elsewhere Condition and incorporated into specific linguistic frameworks such as Generalised Phrase Structure Grammar (Gazdar et al. 1985) and Head-Driven Phrase Structure Grammar (Pollard, 1986) and receive discussion in Flickinger (1987) Flickinger, Pollard and Wasow (1985), as well as Word Grammar (Hudson, *op. cit.*).

The use of a relation which has common use in both the areas of Knowledge Representation and Linguistics has obvious advantages for frameworks that adopt the Homogeneity Hypothesis. To get a clearer idea of the ISA relation we need to look at some trivial examples of propositions containing it. Let us look at a simple, non-controversial taxonomy of nouns such as the following;

- (3) a pronoun ISA noun
- b proper noun ISA noun
- c common noun ISA noun

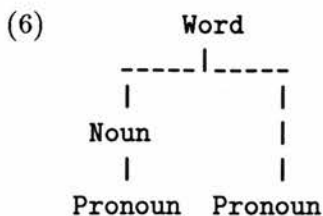
These same facts may be represented diagrammatically and both forms of representation will be used throughout. The propositions in (3) are represented as follows, where the lines represent ISA relations or links:



Returning to the example given in (3) we now need to see how the network is structured with a trivial extension in the range of facts we wish to encode, such as those in (5) below;

- (5) a noun ISA word
- b pronoun ISA word

The facts in (3a) and (5) could be expressed diagrammatically as follows;



nodes in the hierarchy such as the following (where *Relation* stands for one of the relations in (2) (except ISA) and A, B, X, and Y are entities):

- (9) a A Relation X
- b B Relation Y
- c B ISA A

Briefly, the first use of the taxonomy is to attach properties associated with entities at the most general node and for these properties to filter down the ISA hierarchy accumulatively. This will be called *simple* inheritance. This is the case where given 9c) we would want *B* to inherit the property *Relation X*, as well as the property *Relation Y*. The second use of this hierarchy, and a major motivational force behind using the ISA relation, is to state *exceptional* (or more specific) properties of entities. This is the case where in addition to (9) it is known that *X* and *Y* are incompatible in some relevant way. In these cases we would obviously have a conflict about what to inherit. What is required in such cases is some decision procedure. This is provided by the taxonomic ordering imposed by the hierarchy and the attachment of the properties at the relevant positions. In other words the more specific properties, that is those attached lower down the hierarchy, override the more general properties where there is a conflict. This general statement is called the *Selective Inheritance Principle* and receives the following statement from Hudson;

(10) **Selective Inheritance Principle**

“If I is an instance of M, then any proposition which applies to M must apply to I as well (with ‘I’ substituted for ‘M’) provided this proposition does not contradict any proposition which is already known to apply to I”

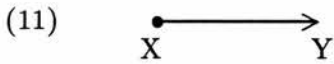
(Hudson, 1984a, p. 18)

The differences between these two uses of the hierarchy will become more apparent below. This completes the brief description of the taxonomic hierarchy and its use in conjunction with the inheritance of (exceptional) properties.

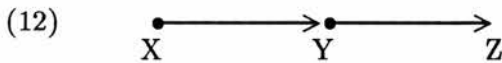
2.1.2 The Co-Word Relation and Dependency

The central relation governing word order in Word Grammar is the *CO-WORD* relation. The *CO-WORD* relation is based on the notion of dependency. Though the general ideas

of dependency arguably go all the way back to the earliest grammarians (Pāṇini c. fourth century B.C.) it was not until the appearance of Tesnière’s seminal work, *Eléments de Syntaxe Structurale*, in 1959 that we find the first comprehensive rendition of the theory as it is known today. It is primarily his work that most modern research on dependency theories goes back to. The basic idea of dependency is that the (syntactic) relations between words can be expressed by the use of directed links between them such as the following;



In this case we call X the *head* and Y the *dependent* and we say that *Y depends on X* or equivalently that *X is the head of Y*. Like many of the primitives of linguistic theory, the notion of head has resisted universal definition and formalisation. All that it appears that we can do is suggest broad tendencies for identifying them. I will look a little closer at the notion of head in Chapter 5 but for a more detailed investigation, the interested reader is referred to Zwicky (1985) and (the reply by) Hudson (1987a). The dependency relation is traditionally thought of as a transitive, asymmetric¹ relation between one word and another. So, in (12) below;



we say that if Y (directly) depends on X, and Z (directly) depends on Y, then Z (indirectly) depends on X. It is useful to distinguish these types of dependency. The term *dependent* will be used exclusively for the cases like (11) of direct dependency. The term *subordinate* will be used as the more general description to cover cases of indirect dependency such as the relation of Z to X in (12) as well as cases of direct dependency. We will investigate these relations more fully in the next chapter.

Following Tesnière, more formal works appeared on dependency theory. These included the works of Gaifman (1965), Hays (1964) and Robinson (1970). Robinson has the

¹Robinson (1970) states that the dependency relation is “transitive, irreflexive and antisymmetric” which is equivalent

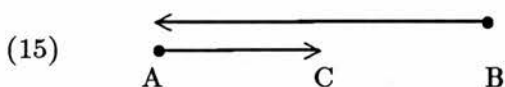
following, “informally presented” axioms of dependency grammars;

- (13) a “one and only one element may be independent;
 b all others depend directly on some element;
 c no element depends directly on more than one other; and
 d if A depends directly on B and some element C intervenes between them (in linear order of string) then C directly depends on A or on B or on some other intervening element.”

(Robinson 1970 p. 260)

The statement in (13d) has come to be known as the *Adjacency Principle*. As we shall see in the next chapter this statement is neither sufficient nor precise enough for our purposes.

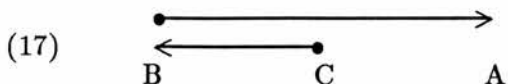
The Adjacency Principle in (13d) allows the following structures, and their mirror images;



Of course (13d) also allows for any further intervening element D, between A and B. Before we can go on, we should look at an example of a violation of the Adjacency Principle. Such a case would be the structure in (16) below which would be the structure assigned to something like *big the cat*. As this structure violates the Adjacency Principle it is correctly predicted to be ungrammatical, though the test of this will be left as an exercise for the interested reader.



Similarly the mirror image of (16) namely (17) below is also excluded;



This, then, is the statement of standard or classical dependency theory. Hudson, however, has found it necessary to depart from the standard. I will not be interested here in

reviewing the arguments for (or against) all the departures from the classical definition made by Hudson but will simply adopt the modifications and introduce them as they arise. Hudson's modifications of (13c) in particular, allowing a word to have more than one head, will be required in my own analysis later, where they will be discussed. Let us now move on to look at some of the implications of a grammar in which the central notion is word.

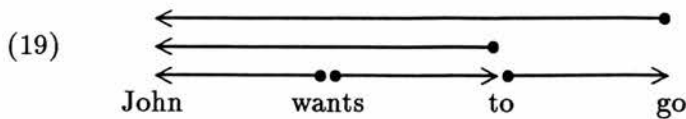
As nothing larger than words is referred to in the grammar, Word Grammar departs from approaches where the notion of constituent or phrase is central. This does not mean that there is no place for some notion of constituency or that constituency cannot be captured in dependency terms. Quite the contrary; we can, in fact, give the following rule for converting dependency analyses into "constituents";

- (18) "... a constituent can be defined as some word plus all the words depending on it, either directly or indirectly"

(Hudson, 1984, p. 92)

So there is a fairly direct relationship. That is not to say that the two are identical. Firstly, constituency structure alone lacks the information about the asymmetrical relation between a head and a modifier which is available in dependency structure. Indeed, the fact that heads are important as is the asymmetrical relation they have with their dependents has also been realised by the proponents of constituency theories. This is evident in X-bar theory and in the ID rules of Generalised Phrase Structure Grammar and Head-Driven Phrase Structure Grammar.

Further, we note that most phrase structure constituency has a very rigid notion of structure. Word Grammar not only licenses many of the constituents of phrase structure grammars but also licenses overlapping and discontinuous constituents. It is generally agreed that these are required for such diverse phenomena as Verb Raising in Dutch (den Besten and Edmondson, 1981) and Complex Fronting in German (Uszkoreit, 1987) but they are also assigned in Word Grammar to such structures as Control in English (Hudson 1984a, 1985a, 1986g). For example, the dependency structure assigned to the sentence *John wants to go* would be that in (19);



Control structures are in fact handled by what is called *modifier sharing* in Word Grammar, i.e. the subject in (19) is shared by the words *wants*, *to* and *go*. The (more flexible) structures given in (19) may be arrived at only by violation of (13c) which bans such things as modifier sharing by stating that elements must have only one head. The structure assigned to (19) under the classical definition of dependency would necessarily exclude the dependencies from *to* and *go* to *John* in order to satisfy (13c). However, it is due to problems associated with assigning adequate structures to examples such as (19) that (13c) has been violated by Word Grammar (Hudson 1984b, 1986g, 1987b). Remembering the rule given in (18) we can derive the following constituents from (19);

- (20) a John
 b John wants to go
 c John to go
 d John go

From (20) we can see that the constituents (20c) and (20d) are discontinuous. The VP constituents in (21) however, do not follow from the translation schema.

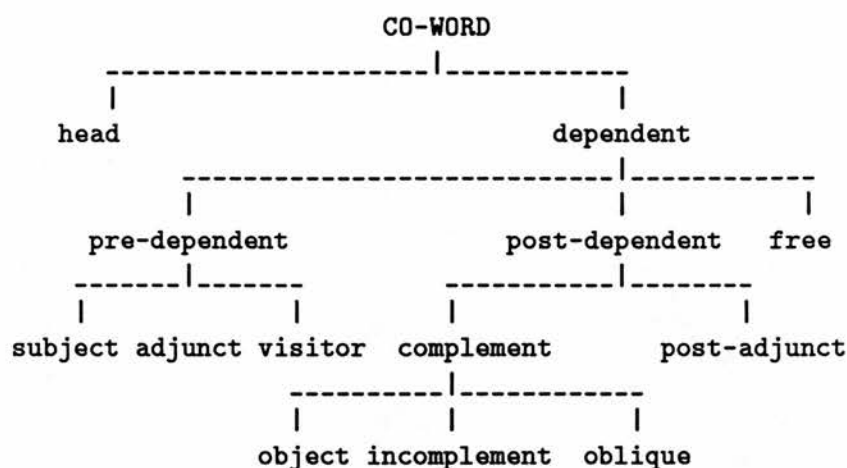
- (21) a wants to go
 b to go

As the debate about the existence or otherwise of VPs continues in constituency frameworks, and will not infringe on the current discussion anyway, I will again leave this as an open question.

Another major distinctive claim of Word Grammar that follows from the use of the CO-WORD relation, is that dependency links form the basis of *grammatical relations*. This contrasts with many other frameworks where grammatical relations are derived. Detailed discussion for the non-derivedness of grammatical relations can be found in Hudson (1984a, 1985a, 1986b, 1986g, 1988a, 1988c, 1989). Essentially the discussion centres around the claim that it is because the grammar must frequently deal with more than one dependent per word, as in the example in (19) that these grammatical relations categories arise, i.e.

as a means of distinguishing dependents. As the dependency links are basic to Word Grammar so must grammatical relations be. As grammatical relations in Word Grammar are seen as subdivisions of dependency, they form an ISA hierarchy of their own. The CO-WORD relation then can be divided up into at least the following, for English;

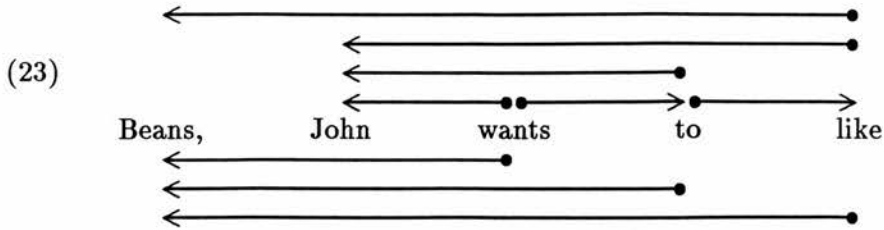
(22) **The Relational Hierarchy**



This hierarchy may also be thought of as the Subcategorisation Hierarchy, in the sense that if, for example, an entity, *X* enters into the *subject* relation with another entity, *Y*, then this is because one of them subcategorises for the other. We will return to this point in Chapter 3. Most of the above have their standard meanings. However, a few need explaining. The grammatical relation of *incomplement* which is an abbreviation of *incomplete complement* is invoked in cases of Control, predicatives and such like as in (19). The idea behind this is that an entity (usually a non-finite verb) may be seen as an incomplete complement of the matrix verb in the sense that it needs a subject (or more generally an element) but must share that element with its head. It is incomplete in the sense that, as it is assumed that all verbs subcategorise for a subject, its subcategorisation frame is not satisfied independently of its head. The grammatical relation of *visitor*, on the other hand is invoked to capture things like unbounded movement, extraction, and so on. The idea here is of a “moved” word, which in looking for its head visits various possible² candidates before finding the correct one. An example of this would be the following where the visitor links have been

²What counts as a possible visiting site is given by the grammar.

drawn below the sentence;



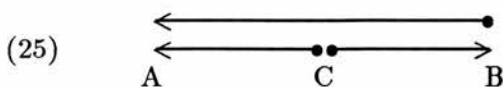
In (23) above the visitor links below the line represent the various possibilities tried by the displaced element *Beans* in attempting to find a head that it can be a dependent of, eventually succeeding with *like*.

The basic licensing principle governing word order in dependency is the Adjacency Principle. In violating (13c) Word Grammar has had to extend this notion. This has been necessary for examples like (19) and (23) where the incompletion (dependency) link between *to* and *John* and some of the visitor links violate the Adjacency Principle by the presence of *want*. The dependency link between *to* and *John* violates the Adjacency Principle because although *to* is the head of *John* they are separated by an element, *wants* which is a dependent of neither, but in fact the head of both. The general intuition behind the Adjacency Principle is to sanction those dependency lines not sanctioned by strict adjacency. As we have extended the possible number of dependency links by violating (13c) we now need to extend the Adjacency Principle to sanction these extra links explicitly. In order to cover the cases in (19) and (23), (13d) needs to be extended to include the case where a head and a dependent are separated by a third word that is the head of both of the other words. (13d) may be revised as follows;

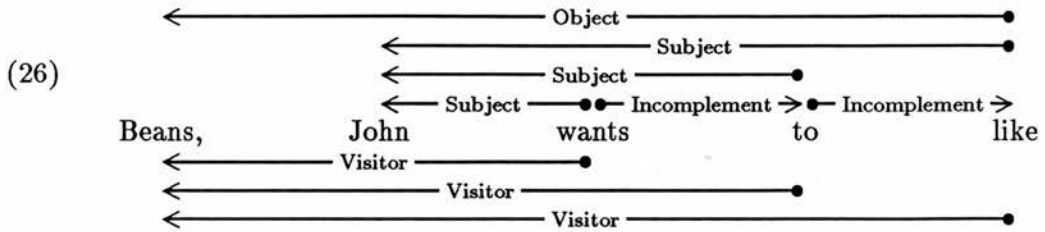
(24) **The Adjacency Principle**

if A depends directly on B and some element C intervenes between them then either C directly depends on A or on B or on some other intervening element, or A and B (directly) depend on C or some other intervening element C

The statement in (24) will allow general structures, such as the following in addition to (15) and (16) above.



Examples (19) and (23) now no longer violate the Adjacency Principle as it is given in (24). Having looked at such examples we are now in a better position to see what is meant by the statement that a head frequently has to deal with more than one dependent and that the grammatical hierarchy provides a useful means of distinguishing dependents. The use of the visitor link not only creates further discontinuous constituents but also adds to the number of potential dependents that a head may have. Using our Relational Hierarchy given in (22), (23) would now look like the following;



The claim is that though *wants* may have three dependents, they can all be distinguished by the different grammatical relation that it has towards each of them, namely *Subject*, *Visitor* and *Incomplement*. A further implication of the approach that has grammatical relations founded on dependency relations, is that any structure which in other frameworks is considered to have undergone some kind of “movement”, in Word Grammar involves a change in grammatical relations. That this is the case can be seen by the addition of of the *Visitor* link from *Beans* to *wants* as well as the *Object* link.

The grammar, then, rests centrally on the two separate but interacting hierarchies; the Classificatory Hierarchy given in (8) above and the Relational Hierarchy given in (22) above. In the analysis of sentences we can say that on this approach the total structure of a word-string in Word Grammar can be expressed as the sum of the properties inherited by the words in the string, in the form of propositions and the connections made in the course of finding the correct heads and dependents, i.e. in satisfying the requirements of the inherited propositions.

2.1.3 Summary

Word Grammar is a general theory of knowledge/cognitive structure. It sees all declarative knowledge as a, partially hierarchically, ordered network of entities related by an unordered

list of propositions. Hence, it could be described as a sort of heterarchy. The main relation in the network is the ISA relation. A sub-part of this heterarchy is knowledge about linguistic entities. Put another way, one of the distinctive features of Word Grammar, is that it presents the principles that govern language structure as special cases of the general principles that govern other cognitive structures.

In Word Grammar, the total structure of a word-string can be expressed as the sum of the structures inherited by the words in the string, in the form of propositions. The main relation in the word-string is the CO-WORD relation, which is seen as a co-occurrence explicitly sanctioned by the grammar. Dependency relations, which are instances of the CO-WORD relation, are basic. Dependency is a relation between two words, one of which is the head, the other dependent. (The dependent of a dependent of an entity is called a subordinate of that entity). There is an asymmetry between head and dependent; it is the head that provides the link between the dependent and the rest of the sentence, and it is the dependent that relies on the head for many of its properties. Grammatical relations are basic and not derived, and form a separate but interacting part of the heterarchy. Moreover it is because the grammar must frequently deal with more than one dependent per word that these categories of grammatical relations arise. It is because dependency is basic to Word Grammar that the grammatical relations are too.

2.2 Problems With The Approach

In this section I will concentrate on some of the negative aspects of Word Grammar. This I see as a necessary step in the development of any theory. The impression might be given that the state of the theory is much worse than it actually is. There are several reasons why this impression would be false. Firstly, Word Grammar is constantly being updated (Hudson 1991, forthcoming) and refined. Secondly, I felt it necessary in the following to emphasize certain deficiencies in Word Grammar to make it clear where the problems lay, in my opinion and in order to make it more obvious how my own modifications were necessitated and how my own work differed though developed from Hudson's.

There are several problems associated with Word Grammar, some belonging to the formalism some belonging to the grammar and some belonging to the concepts on which the grammar is based. It is difficult to assess the linguistic claims without having a formalism that is clear enough to allow such an assessment. It is thus the formalism that the discussion in this section will centre on. From this perspective we will be in a better position to assess the Word Grammar framework. With this in mind we now turn to a description of a grammar fragment.

2.2.1 A Grammar Fragment

The following grammar is one taken from Hudson, distributed in 1987 under the heading *A Sample Of Word Grammar*³. It must be remembered that as this grammar is taken from a 1987 version we should expect certain modifications to the grammar presented in Hudson (1984) to have appeared in order to increase linguistic coverage and as in the usual course of a development. It is the 1987 version of the grammar that seems most pertinent to be discussed, for not only is it one of the more recent, but it also incorporates many of the specifics that any larger grammar must come to terms with. At this point, I will not be concerned to discuss particular syntactic structures, as the discussion will centre on the Word Grammar formalism rather than analyses. The numbering is as it appears in that paper. The grammar is divided into the following sections;

- (27) A Word Classes
- B Inflections
- C Morphology
- D Vocabulary
- E Spelling
- F Grammatical Relations and General Valencies
- G Lexical Syntactic Valencies
- H Semantic Relations
- I General Semantics

It will not be necessary to give a complete listing here and I will be concerned to give enough of the grammar to give a flavour of the framework.

³note that by 1987 Hudson has also changed the name of the model-instance relation to ISA

A. Word Classes

- A1 noun isa word
- A2 verb isa word
- A3 ad-word isa word
- A4 common noun isa noun
- A5 pronoun isa noun
- A6 auxiliary-verb isa verb
- A7 adjective isa adword

B. Inflections

- B1 (inflection of word) isa word
- B2 (lexeme of (inflection of word)) isa word
- B3 word isa (lexeme of word)

- B11 noun has (a number)
- B12 number is (singular or plural)
- B13 (number of noun) is singular
- B14 (number of (inflection of noun)) is plural

- B21 verb has (a finiteness)
- B22 finiteness is (finite or non-finite)
- B23 (finiteness of verb) is finite
- B24 (finiteness of (verb (which has (a head)))) is non-finite
- B25 (verb ((finiteness of which) is finite)) has (a tensedness)
- B26 tensedness is (imperative or tensed)

- B28 tense is (past or present)

C. Morphology

- C1 (spelling of (inflection of common-noun)) is
(a ((spelling of (lexeme of common-noun)), <s>))
- C2 (spelling of (inflection of (verb ((tense of which) is past)))) is
(a ((spelling of (lexeme of word)), <ed>))

D. Vocabulary

- D1 w-boy isa common-noun

E. Spelling

- E1 (spelling of w-boy) is (a <boy>)

F. Grammatical Relations and General Valencies

- F1 head isa companion
- F2 dependent isa companion
- F3 predependent isa dependent
- F4 postdependent isa dependent
- F5 subject isa predependent

- F8 object isa postdependent
- F22 object is (a noun)
- F24 that-object is (a (tensed of verb))
- F27 (subject of (incomplement of verb)) is (subject of verb)
- F28 (verb (tensedness of which) is tensed) has (a subject)

- F32 (head of (common-noun ((number of which) is singular) is (a pronoun)
- F35 auxiliary verb has (ano incomplement)
- F36 (subject of (tensed of auxiliary-verb)) (precedes or follows)
(tensed of auxiliary-verb)
- F40 word has (a head)
- F41 (verb ((finiteness of which) is finite)) has (ano head)
- F42 word has (mano pre-adjunct)
- F43 (head of (dependent of word)) is word
- F44 word has (no complement)

G. Lexical Syntactic Valencies

- G1 w-expect has ((a to-complement) and (ano direct-object))
- G2 w-expect/1 has (a direct object)
- G3 w-see has (a direct-object)

- G7 w-the has (a complement)
- G8 w-a has (a complement)

- G11 (subject of (incomplement of w-to)) is (subject of (head of w-to))

There are many problems associated with the formulation of the grammar in this section.

Those that are most serious can be grouped along the lines given in (28);

- (28) a proliferation of entities
- b proliferation of relations
- c problems with propositions
- d problems associated with the grammar

I will discuss these problems in the order in which they are listed in the next few sections.

Solutions, where they exist, will be presented in the next chapter.

2.2.2 The Proliferation Of Entities

It should be fairly obvious from the grammar fragment presented above that there are more types of entities here than were allowed for in the original formulation, where it was stated that there were only five types of entities. As was mentioned above, though, the actual numbers of entities is not considered problematic provided that we are aware of what they are and that they receive a listing somewhere.

But there still remain problems with the entities in the grammar given in section 2.2.1. In the initial description of Word Grammar it appeared as though the entities were simple entities. But in the grammar fragment we find much more complex entities. These include the following;

- (29) a *a, an, ano* (1 or 0), *mano* (0 or greater), *no, of, which, or, and*,
- b Entities such as
 - B2 (lexeme of (inflection of word)) ...
 - F28 ... (a subject)
 - B12 ... (singular or plural)
 - B24 (finiteness of (verb (which has (a head))))...
 - B25 (verb ((finiteness of which) is finite))...
 - F32 (head of (common-noun ((number of which) is singular)))...
 - F24 ... (a (tensed of verb))
 - B11 ... (a number)
 - B12 number...
 - B13 (number of noun)...
 - C2 (spelling of (inflection of (verb ((tense of which) is past))))...
 - C2 ... (a ((spelling of (lexeme of word)), <ed>))

There is no *prima facie* reason for not including such entities as we find in (29). The problems arise when we try to formalise these entities and can no longer rely on their intuitive interpretations. But before we look at the entities as totalities let us look at their sub-parts. For reasons that will become apparent, I will call the elements in (29a) *operators*. The operator *which* seems to have two uses in the above and so it is reasonable to assume that there are two possible interpretations. These two meanings are apparent in (B24) and (B25) where we have (*which has (a head)*) and (*finiteness of which*). Again we see the same problem with *of*. In (B2) we see (*lexeme of (inflection of word)*) whereas in (F32) we find (*number of which*). Again we have two uses and (at least) two possible interpretations.

This brings us to the other operators, *a*, *an*, *ano*, and *mano*. In (F28) we see the entity (*a subject*). It is not clear from the grammar how this entity differs from the entity *subject*. It is also not clear how to interpret *a* in (F28), (F24) and (B11) consistently. This is partly due to the fact that the interpretation of *a* changes slightly depending on which relation it is involved with (F24 vs B11 vs F28) and partly due to the fact that its interpretation changes relative to the entity it is combined with (*(a subject)* versus *(a number)*). The basic point is that there seems to be no clear syntax for the representational language, and in fact none is given. As a result, there can be no interpretation of this formalism beyond the intuitive. As we will repeatedly find, there is no syntax because the elements of the language are not used uniformly, making it difficult to write such rules.

Let us now look at the entities as totalities. Even when seen in its context in (F24) it is difficult to assign any interpretation at all to *is (a (tensed of verb))*. The examples in (B11) to (B13) all contain the element *number*. But in (B11) it is quantified, in (B12) it stands alone and in (B13) it appears modified by the words *of noun*. This ability to be, apparently randomly, modified by a following bracketed phrase is problematic for a consistent interpretation of the subparts of these complex entities, as well as for efficient implementation.

In (B12) we see the possibility that a complex entity may be formed by the disjunction (or even the conjunction) of other (complex) entities. This is not problematic as far as the formalism itself is concerned, though it was not allowed in the original formulation. However, for purposes of implementation when this option is used (possibly recursively) in conjunction with an operator like *ano*, or *mano* the computational cost in following every possible path would be enormous, even in simple cases.

There is a further development of the formalism in (C2). In that example we see the possibility that an entity may take the form of a list of entities separated by commas. All these devices, while adding greatly to the flexibility of the formalism, come at the expense of having no clear interpretation for the entities, no clear restrictions on what can or can not be an entity, and no clear rules or restrictions for forming other entities not given in

the fragment. It seems that what has gone under the guise of formalism is really some restricted form of English with brackets, incorporating many of its ambiguities.

2.2.3 The Proliferation of Relations

As can be imagined, the problems associated with the entities have repercussions for the relations, and many of the same basic problems carry over. Examples of these include the following;

- (30) a B1 (inflection of word) isa word
- b B2 (lexeme of (inflection of word)) is word
- c B3 word isa (lexeme of word)
- d F40 word has (a head)
- e B11 noun has (a number)

Before I discuss these examples I would like to make a further point about the operators in (29a) above. I have included them in that section as they only seem to be used in the construction of complex entities. However it is not clear that they are entities as they seem to function like relations. This would then be the first case of the proliferation of relations. But let us move on to the relations proper.

In (30) above we have a distinction made between *ISA* and *IS*. Again the precise nature of this distinction is quite elusive and apparently ad hoc especially when we consider the content of the statements made in (27) (though see Hudson 1987d). Further, in (B25) and (F32) we see simple entities such as *verb* in (*verb ((finiteness of which) is finite)*) and *common-noun* in (*common-noun ((number of which) is plural)*) verging on the relational as suggested by the bracketing. Additionally, we see in those same examples relations embedded within the complex entities. Again, there seems to be no clear syntax.

The relation *HAS* in (30) seems to incorporate the ambiguities of the non-auxiliary *have* in English. For example, having a head as in (30d) does not seem to be the same as having a number as in 30e). The former is a subcategorisation statement whereas the latter is more like an intrinsic property of the element in question. Again, finding consistent interpretations proves elusive.

2.2.4 Problems With The Formation of Propositions

Given the previous two sections it should be obvious that there would be problems with constructing propositions. If there is no clear syntax for constructing either entities or relations then there is little hope for getting an overall syntax for propositions. This is further complicated because when these ideas were introduced there were simple entities and simple relations and there was at least the hope that these could be simply formalised. It was also stated that there was a small, fixed number of entities and relations. At least as far as the entities are concerned this no longer seems the case.

To get a better picture of the problems with the formalism mentioned above, I will now attempt a (generous) reconstruction of what a syntax of the grammar given might look like. In order to do this we will need to distinguish, relations, entities and operators, all of which may be used in forming propositions. The operators and relations will be listed below. Any element given in the grammar fragment above that is neither a relation nor an operator will be taken to be an entity. A possible syntax for the grammar fragment is presented below;

- (31)
- | | | | |
|---|-------------|----|---|
| a | Proposition | := | Entity Relation Entity |
| b | Relation | := | has is isa |
| c | Operator | := | a an ano mano no |
| d | Entity | := | word noun verb etc
inflection lexeme etc
number singular plural
finiteness finite etc
head dependent etc |
| e | | := | (Operator Entity) |
| f | | := | (Entity of Entity) |
| g | | := | (Entity, < Entity >) |
| h | | := | (Entity and Entity) |
| i | | := | (Entity or Entity) |
| j | | := | (Entity (Proposition)) or if <i>which is not an entity</i> ,
i) (Entity (which Relation Entity))
ii) (Entity ((Entity of which) Relation Entity)) |

If in the above, the element *which* is taken as an entity then the statement of the expansion of *Entity* in (31j) is simple. If *which* is not an entity but a fixed form given by the expansions along the lines of *and*, (31h) or *of*, (31f) then we need the two more complex expansions given in (31j) namely (i) and (ii). If *which* is an entity, our formalism will be permitted

to make non-sensical statements such as (*which or which*) from (33i). Clearly this is undesirable. If *which* is a fixed form given by the expansions of *Entity*, then it appears that we must have one statement for *is* and *isa*, namely (31j) (ii) and one for *has*, namely (31j) (i). This seems over-stipulative.

The problem just mentioned with making *which* an entity is in fact, a specific case of a more general problem that applies across all the expansions of *Entity*: namely that the syntax given is not restrictive enough in the statements that it allows. In attempting to give a syntax for the grammar fragment presented above we have abstracted away from the particulars of the propositions given in the fragment and in doing so we have ended up with statements that overgeneralise. Further, knowing which expansions to use, when to use them and how to interpret the propositions formed requires meta-knowledge. Though the syntax given in (31) correctly allows entities such as (*a head*) (by (31e)) which are required by the grammar fragment, it also allows entities such as (*a (a head)*) or (*mano (a head)*) amongst other non-sensical statements.

The overgeneration of propositions caused by the lack of an explicit syntax for the representational language is above and beyond the overgeneration of propositions that will occur due to the lack of restrictions on what category of element may be related to what other category in proposition forming. In addition it is not clear whether the statements given in (31) exhaust the possibilities. From the above many of the inadequacies present in the grammar fragment and mentioned in the previous subsections, should now be apparent, as should the ensuing difficulties in getting consistent interpretations for the entities involved.

2.2.5 Further Problems

So far we have dealt with the more technical deficiencies of the formalism. I would now like to move on to some problems associated with the fundamental concepts.

The attentive reader will have noticed a discrepancy creeping into the framework even before the grammar fragment in section 2.2.1 was presented. If, as claimed in the exposition of Word Grammar, propositions are used to relate entities, we are presented with a problem

for entities like *head* and *dependent*. We recall that such entities are subdivisions of CO-WORD and enter into the ISA relation. Thus we find statements in the grammar like the following, encoding the diagram given in (22) above;

- (32) a subject isa pre-dependent
- b object isa complement

We also recall however that CO-WORD is a relation not an entity. It would now seem, then, that propositions can state relations between relations as well as relations between entities. Though this is not problematic *per se*, such extensions if required must be explicitly sanctioned in the syntax of the formalism, and highlights further the gap between theory and practice. One way would be to give things like *head* dual status as both entities and relations. A second would be to make them either one or the other. Possible approaches will be discussed in the next chapter. The next set of problems are associated with the ISA hierarchy. In the grammar fragment above we find statements such as the following;

- (33) a G1 w-expect has (a direct object)
- b G3 w-see has (a direct object)
- c B24 (finiteness of (verb (which has (a head)))) is non-finite

Now, if the purpose of adopting a hierarchical frame-like representation was that grammatical statements could be made efficiently at the highest most general node, as was mentioned in section 2.1.1, then statements such as those in (33a) and (33b) seem, initially at least, to defeat much of the purpose of this approach. This could be more efficiently stated if there were some notion of transitive verb to refer to in the grammar. If (33c) means that any verb which has a head is non-finite, then clearly this is too general a statement to cover tensed embedded clauses.

A more serious problem concerns the Selective Inheritance Principle. All the problems mentioned so far combine to make it difficult to stipulate a clear, consistent and coherent statement of the Selective Inheritance Principle (see (10) above), especially when we consider the complexity and diversity of the entities contained in the fragment. It is for this reason that the Selective Inheritance Principle has gone through many formulations, both formal and less formal, and has proven along with the Adjacency Principle to be one

of the less stable elements of the Word Grammar enterprise. A further hindrance in the specification of the Selective Inheritance Principle comes from the diffuse structure and organisation of the grammar. Hudson sees this as an advantage;

“The network structure is compatible with a more diffuse organisation than this, and I have argued that such diffuseness is a good thing because it seems to correspond to the reality of language data.”

(Hudson, 1984a, p. 4)

This problem is harder to illustrate as it involves the grammar as a whole. The point is that given such diffuseness it is hard to describe a general rule that will make all and only the necessary inferences. Consider the following propositions taken from section 2.2.1;

- (34) a B1 (inflection of word) isa word
b B2 (lexeme of (inflection of word)) isa word
c B3 word isa (lexeme of word)

On some reasonably general version of the Selective Inheritance Principle, the propositions in (34) taken as whole seem to verge on the circular. The Selective Inheritance Principle as given in (10) above (repeated below for convenience) was stated for a grammar with much simpler entities.

(35) **Selective Inheritance Principle**

“If I is an instance of M, then any proposition which applies to M must apply to I as well (with ‘I’ substituted for ‘M’) provided this proposition does not contradict any proposition which is already known to apply to I”

(Hudson, 1984, p. 18)

If it turns out that there are cases where the Selective Inheritance Principle needs to make reference to the internal structure of any of the entities given in section 2.2.1, then it will be difficult to prevent them from doing so in general. Such a facility is in fact required by the grammar in the course of derivations. This has disastrous consequences as in (34) it would be possible to replace recursively any occurrence of the entity *word* by the entity (*lexeme of word*) by (34c). As the replacement of *word* in this case contains as a sub-element the entity *word*, this process could be repeated indefinitely. Even if this could be prevented, on the formulation of the Selective Inheritance Principle as given above, we can still derive propositions from (34) such as those in (38) below;

(36) (lexeme of (inflection of word)) isa (lexeme of word)

whose interpretation is open to question.

There is a final problem with the Selective Inheritance Principle as given above. There is no statement given of what to do in cases of multiple inheritance, i.e. what to do in cases where an entity has two models and the properties inherited conflict. If the multiple inheritance problems are avoided then some statement of how this is achieved is required.

Briefly, the final set of problems concerns the notions behind dependency. Firstly, Robinson's formulation of dependency was informal, and as such imprecise. Further, given that Hudson violates Robinson's formulation of dependency, we need some reformulation of the principles under which Word Grammar operates, but more importantly of the Adjacency Principle, neither of which is forthcoming. These, then, are some of the major problems as I see them with the Word Grammar formalism. In the next chapter I will try to develop a less problematic formalism.

Chapter 3

Word Unification Grammar

In this chapter I would like to discuss some of my own modifications to the Word Grammar framework. In doing so, I shall be concerned primarily with the problems raised in Chapter 2. In attempting to formalise Word Grammar as much as possible, the grammar has necessarily become less complete, and its coverage more restricted. I have called the framework incorporating my modifications *Word Unification Grammar* because the formalism developed crucially uses notions of Unification.

There are two relevant strands of work in Unification connected with natural language processing. The first concerns logic programming, Robinson (1965) and the second the development of feature structures Kay (1979, 1985a) Shieber (1986). The idea of programming in logic developed from the work of Robinson and the first language to allow one to do so was PROLOG. Detailed discussion of this language may be found in Clocksin and Mellish (1981) and Sterling and Shapiro, (1986). PROLOG's use of Unification is restricted to what is called *Term Unification* because of its use of structures resembling first order terms.

As Term Unification was too limited, structures of a more general type were developed for natural language processing. Kay (op. cit.), introduced the notion of feature structures which were a formalisation of the linguistic notion of features. Feature structures can be used to represent many diverse kinds of information and hence seem suited to the current approach. This kind of Unification defined over feature structures is called Graph

Unification. Detailed discussion of the notions involved may be found in Shieber (1986). As Knight (1989) notes, feature structures are similar to first order terms but have several of the conditions lifted. These are that in feature structures:

- (1) a substructures are identified by name and not inferred from argument position
- b the propositions do not require a fixed arity to unify as they do in term unification (as in PROLOG) and as such may be used to represent partial information
- c the distinction between function and argument is removed and
- d variables and co-reference are treated differently in that in term unification, variables are simultaneously place holders, for future instantiation and may also enforce constraints among different parts of a term

The above, then, are the ways in which Term Unification and Graph Unification differ. In feature structures, Unification provides a kind of constraint checking mechanism for merging information from different sources. The type of Unification adopted in Word Unification Grammar is, in principle, Graph Unification because of the flexibility stated above. However given that the Word Unification Grammar formalism was required to be implemented in PROLOG from the outset, and further, that PROLOG only uses Term Unification, the kind of Unification that will be adopted in the exposition is Term Unification. One benefit of this is that the relationship between the grammar presented and its implementation in the appendix will be more apparent. One repercussion of using Term Unification stems from the fact that terms unlike graphs must be of a fixed arity. Where entities require terms of differing arity, the term containing the largest arity will be used and the term with the smaller arity will contain the constant *nil* for those parts of the larger term that do not involve it. This will become clearer below.

Unification was introduced into the Word Grammar formalism for several reasons. Firstly, as has already been mentioned above, one test of the adequacy of not only the formalisation presented, but also of the framework is that it be implementable. The reason for this is that not only does implementation enforce a great deal of formalisation and explicitness on a framework and hence help to clarify it, but implementation may also provide significant tests of the rules of the grammar. The constant revisions required in

going from theory to implementation and back again has turned out to be a most fruitful one.

Given that PROLOG was specifically developed for applications to natural language processing, and that the data structures used in PROLOG appeared compatible with Word Grammar propositions, it seemed a natural choice. But there are more far reaching reasons for formalising Word Grammar using Unification, connected with Kay's work in Functional Unification Grammar.

Unification has been shown by Kay to be a powerful and useful tool to any grammar writer. Further, in recent years, many linguistic formalisms have either incorporated or been based around Unification. These include, Head-Driven Phrase Structure Grammar (Pollard, 1985, Pollard and Sag, 1987) Categorical Unification Grammar (Uzkoreit 1986) and Unification Categorical Grammar (Zeevat et al. 1987). Further still, PATR-II (Shieber 1986) was developed as a general formalism in which Unification-based grammars could be written in a wide variety of ways. Unification thus brings Word Grammar more in line with other frameworks with a more widespread appeal, and also facilitates comparison, both formally and in terms of linguistic coverage.

Before we proceed with a discussion of my modifications, it will be helpful to have a concise statement of the problems associated with the Word Grammar framework discussed in the previous chapter. My approaches to these problems will be discussed in turn in the next few sections. Overcoming these problems successfully may be viewed simultaneously as fulfillment of the sub-goal (1a) set out in Chapter 1 and as a condition of adequacy against which my own formalism should be judged. These may be summarised as follows:

- (2) a Too diffuse a grammar
No attention given to the network as a hierarchical classification device
- b The isolation of primitives
No explicit syntax for the formation of (complex) entities
No fixed (i.e. an indefinite) number of entities
- c No explicit syntax for the formation of propositions
Indefinite number of propositions
No type restrictions on entities and relations
- d No directed inference mechanism(s)
No transitive ISA statement
No clear or usable statement of the Selective Inheritance Principle
No precise definition of the Adjacency Principle

The problems in (2) above will now be discussed in turn and my solutions given.

3.1 Over-Diffuseness

As was mentioned in the previous chapter, the problem of the over-diffuseness of grammar is one that is hard to illustrate as it is an implicit property of a grammar as a whole. Further at the current state of knowledge it is difficult to know what advantages (or disadvantages) there are to such diffuseness. For example, let us return to propositions (B11) to (B14) of the fragment presented in (20) of section 2.2.1 of the previous chapter, repeated here for convenience:

- (3) a B11 noun has (a number)
- b B12 number is (singular or plural)
- c B13 (number of noun) is singular
- d B14 (number of (inflection of noun)) is plural

It is difficult to know what advantages the diffuseness of (3) brings. A formalism that incorporated a similar amount of information but did so in a more condensed form, such as feature structures, would seem to have several advantages over (3) especially when taken over a reasonable sized fragment. It would not only be more convenient when it came to grasping what was being encoded by the grammar, but such a formalism would also be easier to verify and easier to compute with. What we must bear in mind is that in making such changes we remain as faithful as possible to the framework.

One of the most significant changes made to the Word Grammar formalism, then, was the introduction of entities that resembled traditional lexical entries. We will see

examples of these below. As should be apparent from section 2.2.1 there is nothing in the Word Grammar formalism that could reasonably be referred to as such. One of the major motivations for the introduction of lexical entries was the pressure caused from requiring efficiently implementable grammatical statements. One of the most effective and efficient ways of implementing Word Grammar proved to be the condensing of much of the grammatical information into lexical entries. This resulted in a reduction in the overall diffuseness.

The lexical entries introduced make significant use of features. This was a natural step to take given the above discussion on Unification. The use of features was explicitly rejected in the 1984 framework:

“I have made no mention of syntactic features so far in my explanation of word grammar, because we do not need them as part of the theory.”

(Hudson 1984a, p. 1)

This position was weakened slightly by the appearance of the fragment in section 2.2.1 where we find evidence of entities which at least resemble features. The features of Word Unification Grammar have, where possible, been hierarchically arranged. This seems not only a natural development given that we want to take advantage of the net formalism as a hierarchical classification device, but also one of the only way to make the changes required in going from Word Grammar to Word Unification Grammar yet still keeping faith with the underlying intuitions. We will investigate the lexical entries further below, but first we need to look closer at the relations and entities.

3.2 The Isolation of Primitives

The isolation of primitives is relevant for the formation of propositions. Propositions as we now know are formed from entities and relations. This means that we must answer the questions, what are the primitive relations and what are the primitive entities?

3.2.1 Relations in Word Unification Grammar

The relations that we find in the fragment in 2.2.1 include the following:

- (4) a is
- b isa
- c has

As mentioned in the previous chapter the precise distinction between (4a) and (4b) is unclear. Further, the relation in (4c) is used both for subcategorisation requirements (G2) and for feature statements (B11). Both these positions are unsatisfactory. The relations that will be required in Word Unification Grammar are the following:

- (5) a ISA
- b NEEDS

The ISA relation in Word Unification Grammar is used solely to form the hierarchies of the grammar and its meaning is simply the subset relationship. As was mentioned in the previous chapter, there are two uses of the taxonomies formed by the ISA relation. The way these are incorporated into the Word Unification Grammar formalism will be discussed when we come to look more closely at ISA and inheritance below. The NEEDS relation will be used for making subcategorisation statements and as such factors out one part of the ambiguities of the Word Grammar HAS relation. That part of the Word Grammar HAS relation that was used to make propositions related to featural properties of entities will now be incorporated into (hierarchically organised) feature structures that form lexical entries. How this is achieved will be described below.

This brings us to the CO-WORD relation of Word Grammar , which specifies the dependency relation between two words. In section 2.2.5 we noted that there were certain discrepancies specific to the CO-WORD relationship given that propositions are used to express relations between entities. The discrepancy was that the CO-WORD relations formed a separate ISA hierarchy. What this meant was that propositions could now be used to express relations between relations between entities. Again in section 2.2.5 some solutions to this problem were discussed. The possible approaches we could take to this problem include the following:



- (6) a assigning elements like *subject* dual status as both entities and relations
- b making CO-WORD relational and excluding it from entering into the ISA relation.
- c making CO-WORD an entity
- d changing nothing and allowing for these discrepancies
- e doing away with them altogether

Given that we want our formalism, initially at least, to be comparatively simple and clear, we should be wary of adopting either (6a) or (6d) unnecessarily. Adopting (6b) seems counter productive given that the main reasons for including the CO-WORD relation was as a means of taxonomically distinguishing dependents. This leaves (6c) and (6e). Given the reasonably convincing justification for having a taxonomy of dependents discussed above and elsewhere, (6e) seems too radical a step to take initially. As such the approach taken below was to make what was thought of as a relation in Word Grammar, a hierarchically organised set of (featural) entities in Word Unification Grammar. We will see precisely how this is encoded in the section on subcategorisation.

This then leaves us with only two relational primitives in Word Unification Grammar, namely ISA and NEEDS. Before we can see exactly how these are used we need to look at what the primitives are out of which we may form simple and complex entities which may enter into one of our primitive relations. It is only at this point that we can see how they interact with the relations.

3.2.2 Entities in Word Unification Grammar

One of the problems associated with the entities (both simple and complex) in Word Grammar was the lack of an explicit syntax. As such far from there being a small, fixed number of them it now appeared as if there were an indefinite number of entities. This was an unacceptable state of affairs. Word Unification Grammar allows two types of complex entity. These may be schematised as follows:

- (7) a [Root, Grammatical-Features, Agreement-Features],
which may be abbreviated to [Root, GF, Agr]
- b i) [Optionality, Grammatical-Relation, [Root, GF, Agr]],
which may be abbreviated to [Op, GR, [Root, GF, Agr]]
- ii) [Optionality, head, [Root, GF, Agr]],
which may be abbreviated to [Op, head, [Root, GF, Agr]]

Some explanation of (7) is in order. The complex entity, *Grammatical-Relation* in (7b) i) is the Word Unification Grammar equivalent of the Word Grammar, CO-WORD except that in Word Unification Grammar it will be realised as a feature. Further, *Grammatical-Relation* is restricted to that sub-part of the CO-WORD hierarchy given in (22) of the previous chapter that refers to, or is below, *dependent* i.e it does not include the entity *head*. Heads are dealt with by (7b) ii). From this we can see that (7b) i) and ii) are really specific cases of the more general [Op, CO-WORD, [Root, GF, AGR]]. The reason for making this distinction is that requiring a head is seen as different from requiring a dependent, though of course in any given sentence there is a large degree of overlap between the head requirements and the dependent requirements.

The list [Root, GF, AGR] will be dropped in (7b) ii) for ease of exposition, reducing (7b) ii) to [Op, head]. This is possible due to the fact that it is assumed that dependents do not specify features of their heads and as such the list will be redundant. In addition, given that if X is a dependent of Y then Y is the head of X, there would be a reasonable amount of overlap in the information contained were the general structure maintained. Theoretically though, (7b) ii) will not be distinguished from (7b) i).

Here, as below, it must be remembered that for presentation purposes we are dealing with term-like structures in the Unification sense. Finally, I will not be concerned to develop a morphological component here and will just list inflected forms. It is assumed that the variable, *Root* will be instantiated by the relevant morphological propositions when they come to be incorporated. The root form of a word is not to be confused with the root of a sentence, i.e. that element that does not have a head.

In order to get a clearer idea of what the complex entities in (7) mean we have to unpack them. Ideally these complex entities in (7) would be made up from simple entities.

For convenience and ease of exposition however, subsets of these simple entities have been grouped together, as in the variable *Agr*, for example. We also want to restrict our formalism not only in the sense of giving it a precise syntax but also in the sense of restricting it to making only meaningful propositions. This has been achieved by restricting the constant that any variable may be instantiated to. This is what is meant below by the term *typed variable*, an aspect that was missing in the Word Grammar formalism. The schemata in (7) then may be unpacked as follows:¹

Variables and the Entities they Range Over

- (8) *Root* is a typed variable that ranges over the constants which are the root forms of words.
- (9) *Agr* is a variable over the three place list of the following form:
 - [*Person*, *Number*, *Case*], where
 - Person* ∈ (1, 2, 3)
 - Number* ∈ (singular, plural)
 - Case* ∈ (nominative, accusative, dative)
- (10) *Op* ∈ (optional, obligatory)
- (11) *GR* ∈ (dependent, pre-dependent, post-dependent, free-adjunct, subject, pre-adjunct, complement, post-adjunct, object, oblique, by-complement, that-complement, tensed-complement, incomplement, indirect-object, to-complement, inf-complement, ing-complement, en-complement)
- (12) the feature *nil* which may stand in the place of any of the constants to signify not applicable.

¹I will adopt the PROLOG convention of using an initial uppercase letter to denote a variable over entities and lower case to denote a constant entity. As all the variables will keep the same name throughout anyway, there should be no confusion

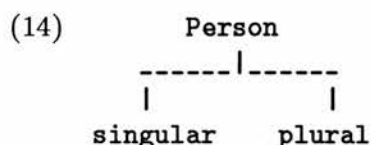
- (13) *GF* is a variable which ranges over two possible types of three place lists. These have the following forms:
- a For Verbs, *GF* takes the following form:
 [*Type, Subcategorisation, Finiteness*], which may be abbreviated to, [*Type, Subcat, Fin*], where
- | | | |
|---------------|---|--|
| <i>Type</i> | ∈ | (normal, auxiliary, modal) |
| <i>Subcat</i> | ∈ | (intransitive, transitive, ditransitive, incomplement, that-complement, tensed-complement, to-complement, infinitive-complement, ing-complement) |
| <i>Fin</i> | ∈ | (finiteness, finite, non-finite, tensed, present, past, infinitival, participle, infinitive, gerund, active, passive) |
- b For Nouns, *GF* takes the following form:
 [*Type, Definiteness, Countability*], which may be abbreviated to, [*Type, Def, Count*], where
- | | | |
|--------------|---|---|
| <i>Type</i> | ∈ | (noun, common-noun, non-common-noun, proper, determiner, pronoun) |
| <i>Def</i> | ∈ | (definite, indefinite) |
| <i>Count</i> | ∈ | (mass, count) |

The above require some discussion. Firstly this list of constants is by no means intended to be exhaustive nor at this stage universal and other features may be imagined and included. Rather the above is intended to be illustrative. Secondly, all the constants in the above constitute the simple entities. For the formalisation given though, this represents an exhaustive listing of the entities with the exception that no root forms of words have been listed. The number of entities may and will be extended as required but as mentioned above will also require a listing and a type assignment. Further, given that all the variables are typed, the formalism has been restricted to a great extent to making only certain types of (meaningful) statements.

There is no reference to semantics or semantic features here, though these will be required at some point. This situation exists primarily because little work has been done in Word Grammar semantics (Gorayska 1985, Hudson, 1986a, 1986f) and my main concern has been to formalise that part of Word Grammar that was already in existence and test this formalisation against some data.

Fourthly, though it is not apparent from (8) to (13) above, all the simple entities are hierarchically organised in ways that will become apparent below. A trivial case is where we find a variable in one of the slots for the constants. For example, if we take

the typed variable, *Number*. Though the constants (*singular or plural*) do not naturally form a hierarchy, the variable, which ranges over the constants may be interpreted under Unification as a top node of a disjunctive hierarchy of depth one, with two terminal nodes: namely singular and plural. This is illustrated below:



As the simple entities are combined to form complex entities the hierarchical organisation leads to complex interactions which we will come to below. This then constitutes the syntax of the entities. We now need to see how the entities and the relations connect up to form propositions.

3.3 Propositions in Word Unification Grammar

Word Unification Grammar has two relations and two types of complex entity. It is the complex entities alone that will enter into the relations to form propositions, i.e. there will be no relations stated between simple entities and hence simple features. In line with the fact that there are only two relations in Word Unification Grammar is the fact that only two types of propositions are allowed. These will occupy the next two subsections.

3.3.1 Taxonomic Propositions

Taxonomic propositions obviously involve the ISA relation. Given that its use will be as a subsumption/subset relation we want to restrict it to stating relations between (complex) entities that are (known to be) of the same kind. That means that in line with the fact that we have two types of entity, we have two types of taxonomic propositions. These two types of taxonomic propositions reflect, in fact, the two central hierarchies given in Chapter 2, namely the Classificatory Hierarchy and the Relational Hierarchy.

In addition to those variables already mentioned above, I will adopt the convention of adding numbered forms of these variables such as *GF1* and *GF2* to denote the same kind of typed variable as *GF* where *GF1* may, but need not, differ from *GF2*. Where I use the

same variable in a single proposition it will be used to denote the same entity or entities as in unification. The syntax of these propositions then may be schematised as follows:

- (15) **Classificatory**
[Root, GF1, Agr] ISA [Root, GF2, Agr]
or more specifically,
[Root, [Type, Subcat, Fin], Agr] ISA
[Root, [Type2, Subcat2, Fin2], Agr]
for verbs and
[Root, [Type, Def, Count], Agr] ISA
[Root, [Type2, Def2, Count2], Agr]
for nouns.

- (16) **Relational**
a [Op, GR1, [Root, GF, Agr]] ISA [Op, GR2, [Root, GF, Agr]]

In order to get a feeling for the propositions, we will look at some specific examples of each. We turn first towards the lexical entries².

(17) **Some lexical entries for nouns**

- a [the, [determiner, definite, Countability], [3, Number, Case]]
- b [he, [pronoun, definite, count], [3, singular, nominative]]
- c [John, [proper, definite, count], [3, singular, Case]]
- d [cat, [common-noun, Definiteness, count], [3, Number, Case]]

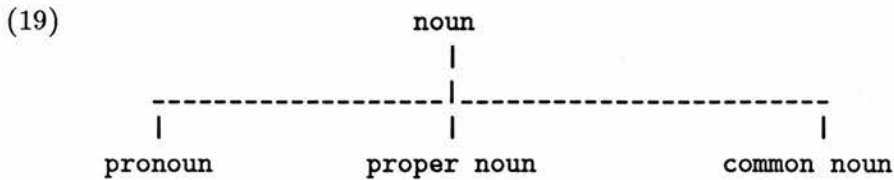
In (17) we see examples of some of the major types of nouns and a specification of their featural structure as given in the grammar. One controversial point in the above concerns the entries for common nouns. This involves whether as root forms they should be specified as singular in the lexicon by default and then overridden for plurals or whether the root form should be unspecified in much the same way as *Case* is. As there is little to choose between the approaches I will not dwell on it here. Needless to say though this point could be made for all un(der)-specified features.

(18) **Some lexical entries for verbs**

- a [sneeze, [normal, intransitive, Finiteness], Agr]
- b [love, [normal, transitive, Finiteness], Agr]
- c [give, [normal, ditransitive, Finiteness], Agr]
- d [have, [auxiliary, to-complement, Finiteness], Agr]
- e [can, [modal, infinitive-complement, Finiteness], Agr]

²We remember that the capitalised words in the following are the names of variables with possible instantiations given in (8) to (13) above

Here we have examples of some of the major classes of verbs. We now need to take a closer look at how these lexical entries relate to the rest of the grammar. This brings us then to the Classificatory propositions. Each of the propositions given in (17) and (18) provides a grammatical decomposition of a word's features (as far as possible) which is explicitly stated in the lexical entries. We now need to see how these relate to the Classificatory Hierarchy. Let us remind ourselves then of the kind of facts that we are trying to encode. Recall the example concerning the taxonomy of nouns given in (4) of the previous chapter and repeated below for convenience:



The advantage that such diagrams have is that they do not suffer from many of the inadequacies that the propositions of Word Grammar did. As a result, such diagrams present the relevant facts in a more neutral fashion³. These three facts may be encoded as follows:

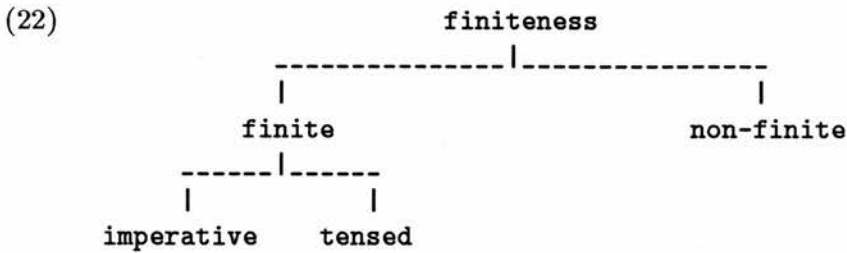
- (20)
- a [Root, [common-noun, Def, Count], Agr] ISA [Root, [noun, Def, Count], Agr]
 - b [Root, [proper, Def, Count], Agr] ISA [Root, [noun, Def, Count], Agr]
 - c [Root, [pronoun, Def, Count], Agr] ISA [Root, [noun, Def, Count], Agr]

The intended meaning of these propositions is that for example in (20a) any root form in the lexicon that is a common noun (or has the feature common noun) is also a noun. Verbs present an altogether more complicated picture. I take it that the task in hand is to attempt to formalise Word Grammar propositions such as those given in (B21) to (B28) of the fragment given in section 2.2.1 and repeated below (where we are not concerned here with has).

³The diagrams still need to be interpreted though.

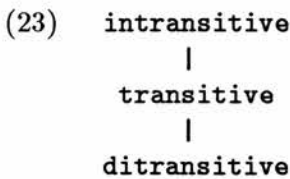
- (21) a B21 verb has (a finiteness)
 b B22 finiteness is (finite or non-finite)
 c B23 (finiteness of verb) is finite
 d B24 (finiteness of (verb (which has (a head)))) is non-finite
 e B25 (verb ((finiteness of which) is finite)) has (a tensedness)
 f B26 tensedness is (imperative or tensed)
 g B28 tense is (past or present)

My approach to formalising these statements was to try and construct a hierarchy for finiteness. Part of the hierarchy formed is given in (22):



Additionally, it was mentioned in section 2.2.5, that it would be an advantage for the formalism to have a hierarchically organised subcategorisation list not only in the interests of having a less diffuse, less repetitions grammar but also to take maximum advantage of the network as a taxonomic classifier and to draw on insights and established techniques in related frameworks (Pollard and Sag (1987), Shieber (1986), Flickinger, Pollard and Wasow (1985)). The function of the typed variable *Subcategorisation* is precisely to capture this. Diagrammatically, the kind of taxonomic facts that we want to encode are the shown in

(23)



The reason we might want to incorporate a taxonomy such as that given above is that all verbs possessing any one of these features will require a subject. Additionally, both transitive and ditransitive verbs require direct objects. Hence a taxonomy constructed along the above lines could be used to state such facts economically. How can the diagrammatically represented facts in (22) and (23) be incorporated into Word Unification Grammar propositions? The facts in (23) may be incorporated along the following lines:

(24) **Verbs Organised By Subcategorisation**

- a [Root, [verb, transitive, Fin], Agr] ISA
[Root, [verb, intransitive, Fin], Agr]
- b [Root, [verb, ditransitive, Fin], Agr] ISA
[Root, [verb, transitive, Fin], Agr]

The meaning of (24b) for example, is that any root form in the lexicon that is a *verb* (or has the feature *verb*) and *ditransitive*, is also a *transitive* verb, though given the strict co-occurrence conditions on features given in (8) to (13) we could substitute the typed variable *Type* here. The facts in (22) may be incorporated as follows:

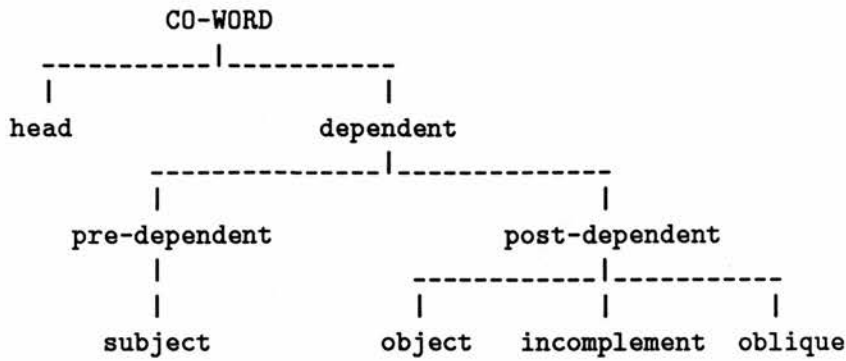
(25) **Verb Organised by Finiteness**

- a [Root, [verb, Subcat, finite], Agr] ISA
[Root, [verb, Subcat, finiteness], Agr]
- b [Root, [verb, Subcat, non-finite], Agr] ISA
[Root, [verb, Subcat, finiteness], Agr]
- c [Root, [verb, Subcat, tensed], Agr] ISA
[Root, [verb, Subcat, finite], Agr]

The meaning of (25c) for example is that any root form in the lexicon, that is a *verb* and *tensed* is also a *finite* verb. What we have then are lexical entries that are made up of lists of features, which enter into individual hierarchies. A lexical entry can now be seen as some path through the hierarchically organised feature set which compose the lexico-grammar (context free. Flickinger, Pollard and Wasow op. cit). The easiest way to understand this is in terms of cross-classification. For example, we could say that any verb is cross-classified for tense, number and person. Where these features are organised into separate hierarchies such cross-classification may be seen as joining up various points in the relevant hierarchies.

This brings us on to the Relational Hierarchy, which will obviously have many formal similarities in common with the Classificatory Hierarchy. The Relational Hierarchy is encoded in the relational propositions. The facts we are trying to encode here are the following:

(26) **The Relational Hierarchy**



These facts may be encoded in the following way:

- (27) [Op, subject, [Root, GF, Agr]] ISA [Op, predependent, [Root, GF, Agr]]
where *Op* is a variable over op(tonal) or ob(ligatory)

What this says is that a subject is also a predependent, all the other entities being variables. This is exactly the statement that we want. In such a way all the relational facts could be encoded. Let us now turn to the subcategorisation propositions.

3.3.2 Subcategorisation Propositions

Subcategorisation propositions involve the NEEDS relation which is used in the sense of one word *licensing* or *being licensed by* some other word. Propositions containing the NEEDS relation involve one lexical element from (7a) and one relational element from (7b). This is because we want to capture the dependency idea that two words enter into a grammatical relation and that this grammatical relation is determined by the head word in the pair. The syntax of these propositions then may be schematised as follows:

- (28) a [Root, GF, Agr] NEEDS [Op, GR, [Root2, GF2, Agr2]]
b [Root, GF, Agr] NEEDS [Op, head]

The structure of the left hand side, given above makes it possible to state quite complex, general and specific *preconditions*, expressed in terms of features, that a word must possess in order to subcategorise for another word. The structure of the relational element makes it possible to state similarly complex features of the word subcategorised for. These may be seen as *postconditions*. The two possibilities in (28) arise due to the fact that there

are two types of relational entities as given in (7b) and reflect the two sub-types of the dependency relation namely subcategorising for an element and requiring a head, which are two sides of the same coin. This then answers the question posed at the end of the last subsection. Encoding specific subcategorisation or licensing statements should be apparent given the above. Say, for example, we wanted to encode the the following fact, given in the fragment in 2.2.1:

(29) F28 (verb (tensedness of which) is tensed) has (a subject)

If this may be taken to mean that tensed verbs require an obligatory subject, then (29) may be encoded in Word Unification Grammar as in (30a):

- (30) a [Root, [verb, Subcat, tensed], Agr] NEEDS
 [ob, subject, [Root2, GF2, Agr2]]
 b [Root, [verb, Subcat, tensed], Agr] NEEDS
 [ob, subject, [Root2, [noun, Def, Count], Agr2]]
 c [Root, [verb, Subcat, tensed], [Number, Person, nil]] NEEDS
 [ob, subject, [Root2, [noun, Def, Count], [Number, Person, Case]]]
 d [Root, [verb, Subcat, tensed], [Number, Person, nil]] NEEDS
 [ob, subject, [Root2, [noun, Def, Count], [Number, Person, nominative]]]

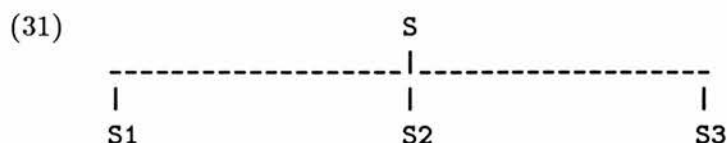
If we wanted to say something more specific than (29) such as the fact that a tensed verb takes a subject which is a noun then given our formalism this is unproblematic. Such a statement may be encoded as 30b) above. If further still we wanted to encode the fact that a tensed verb takes a subject that agrees with it in number and person, again this would be unproblematic. This may be achieved along the lines of (30c) by using the same variables in the agreement slot of the precondition of the NEEDS proposition as in the postcondition. If additionally we wanted to state that a tensed verb requires a subject that not only agrees with it in number and person but also has nominative case, then this may be encoded as in (30d). In such a way some of the advantages of the Word Unification Grammar formalism using Unification, particularly over the Word Grammar formalism should now be apparent.

3.4 Inference in Word Unification Grammar

Having discussed the syntax of the entities, relations and propositions, we now need to look at what I referred to as the operating principles in Word Grammar, making changes where necessary in order to accommodate the Word Unification Grammar formalism. There are two operating principles, one associated with each relation. Before we can look at the operating principles associated with NEEDS and hence at inheritance and selective inheritance, we need to look at the hierarchy and hence at the operating principles associated with ISA.

3.4.1 The Operating Principle For ISA

In order to make the hierarchies of Word Unification Grammar as simple and unproblematic as possible, they have been constructed along certain restricted lines. This has been done in order to avoid some of the problems that have beset other formalism particularly the problem of multiple inheritance. We recall that nothing was mentioned either positively or negatively about the cases of multiple inheritance in the Word Grammar framework. The approach taken here is to avoid multiple inheritance by excluding it in principle. In order to see how this is achieved we need to see the conditions for setting up hierarchies. In setting up a hierarchy, we go from super-class to sub-classes in the following way:



The conditions imposed on forming such a hierarchy are that the sub-classes must exhaust the super-class. What this would mean for the above hierarchy is that there must be no other entity (in the grammar) that is an S other than those given, i.e. S1, S2 and S3 exhaust S. Further the sub-classes must be mutually exclusive, so that S1 incompatible with both S2 and S3 and vice versa. If these conditions are maintained throughout every hierarchy then the problems associated with (Nixon) diamonds and multiple inheritance will be avoided by stipulation. In order to see the advantages of this we need to look at

how taxonomies set up in this way are used for simple and special inheritance. This will occupy the next section.

Before we see how these may be achieved in Word Unification Grammar and in the interests of completeness, we need a statement of the transitive closure of ISA. What we require of ISA is the following:

(32) *From*
 X ISA Y and
 Y ISA Z, *infer*
 X ISA Z.

This is the same as my formulation for Word Grammar given in (7) of the previous chapter. Having linked up the propositions to form the required network, we recall that the network may be used for inheritance in two distinct ways given in section 2.1.1. These were called the simple (or accumulative) and exceptional uses. In order to see how these may be incorporated into the Word Unification Grammar framework we need to look at the operating principles that govern NEEDS.

3.4.2 Inheritance in Word Unification Grammar

Consider the following structure, where the horizontal lines are NEEDS statements and the vertical lines are ISA statements.

(33) F2-----P
 |
 |
 F1-----Q
 |
 |
 r

A root form of a word, r , is mapped on to a set of lexical features $(f_i \dots f_n)$ where each f_i is a specific feature as given by the morphological component. This set of features is abbreviated to F1 in the above diagram. These features are hierarchically organised as sets of (connected) features as in (25) not as individual features. That is to say, there is no statement in the grammar that says f_i ISA $f_{i'}$. Instead, if we wanted to incorporate such a fact we would say $(f_i \dots f_n)$ ISA $(f_{i'} \dots f_n)$. This is represented in (33) by the vertical line from F1 to F2, which are abbreviations for sets of features.

The first type of inheritance we are interested in is the simple or accumulative inheritance. This is the unproblematic case where *r* in (33) would inherit both *P* and *Q*. Simple inheritance is most commonly used in cases where the hierarchical structure of the grammar is used to make economic subcategorisation statements such as the following for intransitive, transitive and ditransitive verbs:

- (34) a [Root, [verb, intransitive, Fin], [Agr] NEEDS
 [ob, subject, [Root2, [noun, Def, Count], [Person, Number, nominative]]]
 b [Root, [verb, transitive, Fin], Agr] NEEDS
 [ob, object, [Root2, [noun, Def, Count, [Person, Number, accusative]]]
 c [Root, [verb, ditransitive, Fin], Agr] NEEDS
 [ob, indirect object, [Root2, [noun, Def, Count, [Person, Number, accusative]]]

Let us look at the above one at a time. The first statement says that any Root that is a intransitive verb needs an obligatory subject that is a noun and that shares its features for *Person* and *Number* and has nominative case. The second statement says that any Root that is a transitive verb needs an obligatory object that is a noun in the accusative case. Given that any feature set containing the feature *transitive* ISA *intransitive* by (23) and (24) in the sense described for (33) this has the effect of requiring all transitive verbs to also take a subject that is a noun that agrees with it. The process whereby a ditransitive verb inherits the fact that it requires an object and a subject as well as an indirect object as given by (34c) can be easily imagined.

Having seen the usefulness of making grammatical statements in terms of economy and generality, we now need to see how this will work in the exceptional case. One advantage of the net formalism is that we can override certain general propositions by more specific propositions. There are many ways this can be achieved. Rather than having an extra relation DOES-NOT-NEED or NEEDS-NOT the approach taken in Word Unification Grammar is simply to add the prefix marker *non* to the right hand side of NEEDS statements. This trivially extends the syntax to include propositions of the following form:

- (35) a [Root, GF, Agr] NEEDS non([Op, GR, [Root2, GF2, Agr2]])
 b [Root, GF, Agr] NEEDS non([Op, head])

These propositions are similar to those in (27) above, except that the right hand side is preceded by *non*. We can now say in the example in (33) that where either $Q = \text{non}(P)$ or $P = \text{non}(Q)$ that in both cases Q overrides P . An example of this is the following:⁴

- (36) a [Root, [word|Y], Agr] NEEDS [ob, head]
b [Root, [verb, Subcat, tensed], Agr] NEEDS non([ob, head]).

The first statement says that any root that is a word (i.e. all words) NEEDS an obligatory head. The second statement says that any Root that has the features of being a tensed (intransitive) verb does not need an obligatory head. As the feature sets containing the features *verb* and *tensed* are more specific than the feature sets containing the feature *word* in the sense that (36b) ISA (36a), (36b) will override (36a) as (36b) ISA (36a). We have still not captured the facts about tensed verbs however. The correct statement about tensed verbs is that they may optionally have a head. This is required for embedded finite clauses. This fact may be incorporated as follows:

- (37) [Root, [verb, Subcat, tensed], Agr] NEEDS [op, head]

The statements in (36) and (37) combine to give us the desired results. It should be apparent from (36) and (37) that the element *non* is acting merely as a marker or a pattern matcher. Ideally (37) should override (36a) directly, by virtue of being a more specific statement about essentially the same facts. In practice though the way Word Unification Grammar overrides more general propositions such as (36a) is to have two statements at the relevant, more specific level. The *first* more specific statement is identical to the right hand side (or the post conditions) of the more general statement except that it contains the marker *non* (or vice versa). In other words the more general statement has to be specifically overridden. This is the relation that (36b) has to (36a). The *second* more specific statement whose left-hand side (or precondition) is identical to the first more specific statement (as in 36a) and (37)) is then used to make the correct statement.

We now need to state the conditions under which inheritance may take place and conditions under which inheritance may be *overridden*. This is called the Selective Inheritance

⁴I will adopt the PROLOG convention [X|Y] to mean the list with X as its head and Y as its tail, where Y may also be a list

Principle in Word Grammar. Informally what we want to say is that for any given element X, it inherits all the properties of its ancestors unless they are overridden.

- (38) **Selective Inheritance Principle**
If
X ISA Z and
Z NEEDS A and it is not the case that
OVERRIDDEN(A), then *infer*
X NEEDS A

In order for this to work we have to give some reasonably precise statement of the notion *overridden*. The intuition is that for A to be overridden in (38) there must be some Y in the hierarchy that is between X and Z that has the property non(A). This may be stated as follows.

- (39) *If*
X ISA Z and
Z NEEDS A and it is not the case that
 X ISA Y and
 Y ISA Z and
 Y NEEDS B and
 B = non(A) or
 non(B) = A then *infer*
X NEEDS A

This then is a statement of the default inheritance mechanism used in Word Unification Grammar. One final point needs to be made however. As P and Q in (33) are also hierarchically classified as are F1 and F2, we might want to incorporate into (39) the case where Q ISA P i.e. is more specific in some sense. In such a case we want r to inherit the more specific case namely, Q. Such cases are not included in the formalism for simplicity as the apparatus described above is sufficient (for present purposes at least).

3.4.3 Summary

The previous four sections have addressed the points in (2a) to (2d), excluding the Adjacency Principle. In doing so, we have now arrived at a formalism in which we have isolated certain primitive entities and used them to form more complex entities. The way in which complex entities may be formed from simple entities is strictly governed by the syntax given in (7) and the type restrictions given in (8) to (13). As such the formalism is

restricted to constructing only meaningful entities. The number of entities is also fixed and definite and we have a clear way of deciding, given our simple features whether a complex entity is included or excluded.

Further there is a fixed number of relations with a simple and fixed interpretation; these are the ISA relation which is simply a subset relation and the NEEDS relation which is simply the subcategorisation relation, with the addition that entities may “subcategorise” for heads. The isolation of primitive entities, the restrictions on forming complex entities and the isolation of primitive relations has made it possible to give an explicit syntax for the formation of propositions in Word Unification Grammar given in (15) and (16). This would appear to a long way towards restricting ourselves to only making meaningful propositions.

A further consequence of this formalisation is that it has been possible to give a clear statement of the operating principles governing the primitive relations. This is particularly so in the case of the Selective Inheritance Principle given in (39). In the process of achieving this it proved necessary to make use of and hence incorporate lexical entries. This had the effect of reducing the overall diffuseness of the grammar. A further advantage of having a more precise formalism is that it has been possible to make greater use of the taxonomy as a hierarchical classificatory device as was shown in (20) (24), and (25).

In conclusion we can say that all the problems with the Word Grammar formalism mentioned in (2) have been addressed resulting in the Word Unification Grammar formalism and in the process we have achieved subgoal (1a) of Chapter 1.

3.5 A Formalisation Of Dependency

In this section⁵ I will formalise the notion of dependency. The notion of dependency will be formalised initially without reference to the Adjacency Principle. I will then go on to give a statement of the Adjacency Principle. We will of course be interested in dependency structures that respect the Adjacency Principle. I shall also relate the formalisation to (13)

⁵I would like to thank Jerry Seligman for help with the more technical parts of this section

of Chapter 2 which is repeated here for convenience.

Robinson has the following, “informally presented” axioms of dependency grammars:

- (40) a “one and only one element may be independent;
b all others depend directly on some element;
c no element depends directly on more than one other; and
d if A depends directly on B and some element C intervenes between them (in linear order of string) then C directly depends on A or on B or on some other intervening element”

(Robinson 1970 p. 260)

In Chapter 2 it was stated that these Axioms were not sufficiently precise for our purposes.

Let us see how we can make them more so, bearing in mind that (40c) has been dropped.

A *dependency structure* $\langle E, d, r \rangle$ consists of a finite sequence, $E = \langle e_1 \dots e_n \rangle$ (of elements) together with an asymmetric, binary relation, d , (for *depends*) on $|E| = \{e_1 \dots e_n\}$ and a designated element, $r \in |E|$, (called the “root”) such that:-

A. for all $x \in |E|$, $\neg d(r, x)$

B. for all $x \in |E|$, if $x \neq r$ then there is some $y \in |E|$ such that

$d(x, y)$ and $x \neq y$

If $d(x, y)$ then we say that x *depends on* y or equivalently, y *is the head of* x . Given this it should be apparent that *A* above is the encoding of (40a) and *B* is the encoding of (40b) which to be more precise is rephrased to say that all the other non-root elements depend on some other element. i.e. no element may depend on itself.

In addition, given a subsequence E' of E , we define a binary relation s (for *subordinate*) on $|E'|$ by the following:-

C. for all $x, y \in |E'|$ $s(x, y)$ in $|E'|$ iff $\exists x_1, \dots, x_n \in |E'|$ such that

$d(x_i, x_{i+1})$ for $1 \leq i \leq n$ and $x_1 = x$ and $x_n = y$

i.e. s is the transitive closure of d restricted to $|E'|$.

If $s(x, y)$ in $|E'|$ then we say that x is *subordinate* to y in $|E'|$

This then is the definition of a dependency structure. We now need to restrict ourselves to dependency structures that respect the Adjacency Principle.

A dependency structure $\langle E, d, r \rangle$ satisfies the *Adjacency Principle* (where $[x, y]$ is the subsequence of E with end points x and y ⁶) iff

D. for all $x, y \in |E|$ if $d(x, y)$ then either

i) for all $z \in [x, y]$, either a) $z = y$ in $[x, y]$ or

b) $s(z, y) [x, y]$

OR

ii) there exists a $z \in [x, y]$ such that

a) $s(x, z)$ in $[x, z]$ or

b) $s(y, z)$ in $[x, z]$

This says that for a dependency relation to hold between two elements, x and y , where x depends on y , they must be either strictly adjacent (next to each other) as in Dia), or all the intervening elements must be subordinate to y , Dib) or there must be some intervening element that both x and y and every other intervening word is subordinate to. The relevant notion of intervening has had to be captured by relativizing subordinacy to a substring which has x and y as end points.

We have now fulfilled subgoal (1b) of Chapter 1 for the Word Unification Grammar formalism. We now need to look at how the above may be used in practice. This will concern the next two sections.

3.6 Definition of Grammar

Informally⁷ we say that the total structure of a word-string is expressed as the sum of the properties inherited by the words in the string, in the form of propositions in virtue of their lexical features and the connections made in the course of finding the correct heads and dependents. In other words dependency grammar may be seen as a mapping from strings of words to dependency structures that respect the Adjacency Principle. We can now attempt to formally state what a grammar will look like:

⁶ $[x, y] = [y, x]$ i.e. order is irrelevant in the definition

⁷see footnote 5

A word string, S , is some sequentially numbered collection of words:

Let $S = \{w_1, \dots, w_n\}$, where w is a word, and $n \geq 1$

We shall take as our example sentence *Mary loves her dog*. Hence we have the following:

$S = \{\text{Mary}_1, \text{loves}_2, \text{her}_3, \text{dog}_4\}$

F is a function from words to the set of their (partially specified) lexical assignments as in (33) above:

we say $F(w_i) = \{\hat{w}_i\}$

For our example we have the following:

$F(\text{Mary}_1) = \{[\text{Mary}, [\text{proper}, \text{definite}, \text{count}], [3, \text{singular}, \text{Case}]]_1\}$

which is abbreviated to $\{\hat{M}\text{ary}_1\}$

$F(\text{loves}_2) = \{\hat{\text{loves}}_2\}$

$F(\text{her}_3) = \{\hat{\text{her}}_3\}$

$F(\text{dog}_4) = \{\hat{\text{dog}}_4\}$

Some words will have more than one lexical assignment. In such a case, we need to distinguish the various lexical assignments.

we say $F(w_i) = \{\hat{w}_i, \hat{w}_i', \hat{w}_i'', \dots\}$

I will adopt the convention

$\hat{w}_i^* = \{\hat{w}_i, \hat{w}_i', \hat{w}_i'', \dots\}$

Let \hat{S} be the set of analysed words formed by applying F to every element in S

$\hat{S} = \{\{\hat{w}_1, \hat{w}_1', \hat{w}_1'', \dots\}, \{\hat{w}_2, \dots\}, \dots, \{\hat{w}_n, \dots\}\}$

or adopting the above convention

$\hat{S} = \{\hat{w}_1^*, \hat{w}_2^*, \dots, \hat{w}_n^*\}$

There are two types of subcategorisation requirement:; subcategorising for dependents, D , and “subcategorising” for heads, H . These are given below, where arrows represent *NEEDS* statements and point from heads to dependents. We will adopt the PROLOG convention of using capital letters to denote variables.

$\hat{w} \longrightarrow \langle \text{Op}, D(X) \rangle$

$\hat{w} \longleftarrow \langle \text{Op}, H \rangle$

where H is restricted to $\{\text{head}_1, \text{head}_2\}$ and

where D is some dependent name which in general will include subject, object and so on but which we will restrict to $\{\text{pre}, \text{post}\}$ and

where Op is a variable over either oblig(atory) or opt(ional), and

where X is some partially specified feature structure that is similar to (a lexical assignment like) \hat{w} except that it does not usually contain a specified word. This is because a head does not usually specify the exact word that it takes as a dependent, as we saw above.

Let G be a (partial) function from lexical assignments to subcategorisation requirements, i.e. a function that applies to the output of every element of the output of F to yield all the associated *NEEDS* statements of the form given above:

$$G(\hat{w}) = \{ \hat{w} \longrightarrow \langle Op, D_1(X) \rangle, \dots, \hat{w} \longrightarrow \langle Op, D_n(X) \rangle, \\ \hat{w} \longleftarrow \langle Op, H_1 \rangle, \dots, \hat{w} \longleftarrow \langle Op, H_m \rangle \}$$

In our example then we can say

$$G(\hat{loves}_2) = \{ \hat{loves}_2 \longleftarrow \langle \text{opt}, \text{head} \rangle, \\ \hat{loves}_2 \longrightarrow \langle \text{oblig}, \text{pre}([\text{Root}, [\text{noun}, \text{Def}, \text{Count}]]]) \rangle, \\ \hat{loves}_2 \longrightarrow \langle \text{oblig}, \text{post}([\text{Root}, [\text{noun}, \text{Def}, \text{Count}]]]) \rangle \}$$

Let \vec{S} be the set containing the subcategorisation requirements (both heads and dependents) determined by G for all the elements in \hat{S} .

$$\vec{S} = \{ G(\hat{w}_1), G(\hat{w}_1'), G(\hat{w}_1''), \dots, G(\hat{w}_2), G(\hat{w}_2'), G(\hat{w}_2''), \dots, G(\hat{w}_n), G(\hat{w}_n'), \\ G(\hat{w}_n'') \}$$

which is equivalent to

$$\vec{S} = \{ \{ \hat{w}_i \longrightarrow \langle Op, D(X) \rangle, \dots, \hat{w}_i \longleftarrow \langle Op, H \rangle \}, \dots \\ \{ \hat{w}_n \longrightarrow \langle Op, D(X) \rangle, \dots, \hat{w}_n \longleftarrow \langle Op, H \rangle \} \}$$

In our example above, we would have the following \vec{S} :

$$\vec{S} = \{ \hat{Mary}_1 \longleftarrow \langle \text{oblig}, \text{head} \rangle \\ \hat{loves}_2 \longleftarrow \langle \text{opt}, \text{head} \rangle \\ \hat{loves}_2 \longrightarrow \langle \text{oblig}, \text{pre}([\text{Root}, [\text{noun}, \text{Def}, \text{Count}]]]) \rangle \}$$

$$\hat{loves}_2 \longrightarrow \langle \text{oblig, post}([\text{Root}, [\text{noun}, \text{Def}, \text{Count}]]]) \rangle$$

$$\hat{her}_3 \longleftarrow \langle \text{oblig, head} \rangle$$

$$\hat{her}_3 \longrightarrow \langle \text{oblig, post}([\text{Root}, [\text{common-noun}, \text{Def}, \text{Count}]]]) \rangle$$

$$\hat{dog}_4 \longleftarrow \langle \text{oblig, head} \rangle \}$$

In addition we have the subset relation *ISA* as defined for Word Unification Grammar.
Let I_{pre} be the inference rule:

$$\frac{\hat{w}_i \longrightarrow \langle \text{Op, pre}(X) \rangle, \hat{w}_j \longleftarrow \langle \text{Op, H} \rangle, j < i, \hat{w}_j \text{ ISA } X}{\text{Depends}(w_j, w_i)} \text{-infer}$$

In our example this would allow us to make the following inference
Given

$$\frac{\begin{array}{l} \hat{loves}_2 \longrightarrow \langle \text{oblig, pre}([\text{Root}, [\text{noun}, \text{Def}, \text{Count}]]]) \rangle \text{ and} \\ \hat{Mary}_1 \longleftarrow \langle \text{oblig, H} \rangle \text{ and} \\ 1 < 2 \text{ and that} \\ \hat{Mary}_1 \text{ ISA } [\text{Root}, [\text{noun}, \text{Def}, \text{Count}]] \end{array}}{\text{Depends}(\hat{Mary}_1, \hat{loves}_2)} \text{-infer}$$

Let I_{post} be the inference rule:

$$\frac{\hat{w}_i \longrightarrow \langle \text{Op, post}(X) \rangle, \hat{w}_j \longleftarrow \langle \text{Op, H} \rangle, i < j, \hat{w}_j \text{ ISA } X}{\text{Depends}(w_j, w_i)} \text{-infer}$$

This allows us to make the following inference for example:
Given

$$\frac{\begin{array}{l} \hat{loves}_2 \longrightarrow \langle \text{oblig, post}([\text{Root}, [\text{noun}, \text{Def}, \text{Count}]]]) \rangle \text{ and} \\ \hat{her}_3 \longleftarrow \langle \text{oblig, H} \rangle \text{ and} \\ 2 < 3 \text{ and} \\ \hat{her}_3 \text{ ISA } [\text{Root}, [\text{noun}, \text{Def}, \text{Count}]] \end{array}}{\text{Depends}(\hat{her}_3, \hat{loves}_2)} \text{-infer}$$

Let C be a *choice function* which selects some \hat{w} in \hat{w}^*

Let \hat{S}_C denote the set formed by applying the choice function, C to every element in \hat{S} .

i.e.

$$\hat{S}_C = \{ C(\hat{w}_i^*) \dots C(\hat{w}_n^*) \}$$

In other words, \hat{S}_C selects one lexical feature structure for each word in S .

Let \vec{S}_C denote the subset of \vec{S} associated with those elements in \hat{S}_C , i.e.

$\vec{S}_C = \{ G(C(w_i)), \dots G(C(w_n)) \}$ i.e. those subcategorisation statements associated with the chosen lexical assignments in \hat{S}_C

Let \vec{I}_C be the set derived from the inference rules and from the subcategorisation statements in \vec{S}_C i.e.

$$\vec{I}_C = \{ \text{Depends}(w_i, w_j), \text{Depends}(w_n, w_m), \dots \}$$

We say that the string, S is a *Sentence* iff there is some choice function C such that

$$(\forall y \in \vec{S}_C) (\text{if } y \text{ is of the form, } \hat{w}_i \longrightarrow \langle \text{oblig}, D(X) \rangle \text{ then}$$

$$(\exists z \in \vec{I}_C) (\text{where } z \text{ is of the form } \text{Depends}(w_j, w_i))$$

and

$$(\forall y \in \vec{S}_C) (\text{if } y \text{ is of the form, } \hat{w}_i \longleftarrow \langle \text{oblig}, H \rangle \text{ then}$$

$$(\exists z \in \vec{I}_C) (\text{where } z \text{ is of the form } \text{Depends}(\hat{w}_i, \hat{w}_j))$$

What this says is that a string of words forms a sentence if and only if for every obligatory dependent required one has been found and for every obligatory head required one has been found i.e. every obligatory requirement has been met. Obviously this also allows any number of optional requirements to be satisfied as well.

Returning to our example, on the assumption that we only have one lexical assignment for each element in S , we have the following \vec{S}_C , and \vec{I}_C :

$$\vec{S}_C = \{ \hat{M}ary_1 \longleftarrow \langle \text{oblig}, H \rangle$$

$$\hat{lo}ves_2 \longleftarrow \langle \text{opt}, H \rangle$$

$$\hat{lo}ves_2 \longrightarrow \langle \text{oblig}, \text{pre}([\text{Root}, [\text{noun}, \text{Def}, \text{Count}]]]) \rangle$$

$$\hat{lo}ves_2 \longrightarrow \langle \text{oblig}, \text{post}([\text{Root}, [\text{noun}, \text{Def}, \text{Count}]]]) \rangle$$

$$\hat{h}er_3 \longleftarrow \langle \text{oblig}, H \rangle$$

$$\hat{h}er_3 \longrightarrow \langle \text{oblig}, \text{post}([\text{Root}, [\text{common-noun}, \text{Def}, \text{Count}]]]) \rangle$$

$$\hat{d}og_4 \longleftarrow \langle \text{oblig}, H \rangle \}$$

$$\vec{I}_C = \{ \text{Depends}(\hat{M}ary_1, \hat{lo}ves_2), \text{Depends}(\hat{h}er_3, \hat{lo}ves_2) \text{ Depends}(\hat{c}at_4, \hat{h}er_3) \}$$

Now for every *oblig* statement in \vec{S}_C , there is a *Depends* statement in \vec{I}_C . Our string of words in S is thus a sentence.

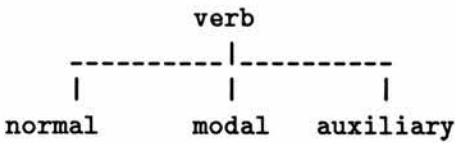
3.6.1 A Word Unification Grammar Fragment of English

In this section I would like to present a grammar fragment for English in order to show more specifically the details of the grammar. Rather than present a list I will introduce the propositions in chunks, discussing the innovations where they appear.

VERBS

Verb organised by Type

The hierarchy this encodes is the following



[Root, [normal, Subcat, Fin], Agr] ISA [Root, [verb, Subcat, Fin], Agr]

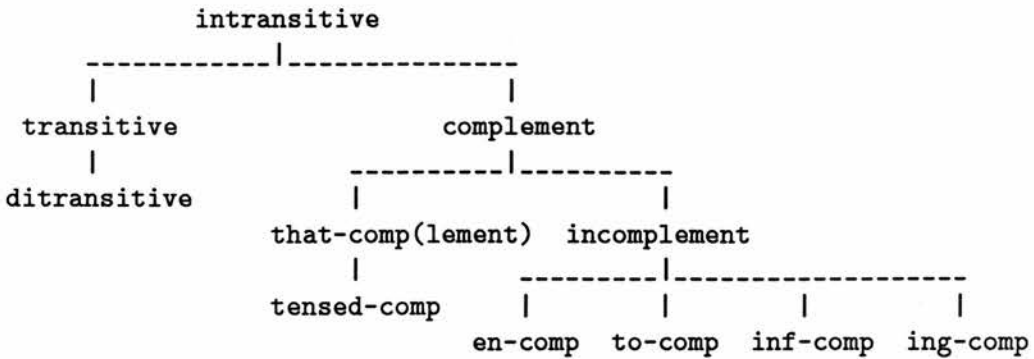
[Root, [modal, Sub, Fin], Agr] ISA [Root, [verb, Subcat, Fin], Agr]

[Root, [auxiliary, Subcat, Fin], Agr] ISA [Root, [verb, Subcat, Fin], Agr]

This and all the other hierarchies presented represent the current state of the grammar. There is no finality about them. They are the ones that are most useful for the structures that will be discussed below. If in order to increase the coverage of the data, some other organisation proved more useful it would be adopted. The grammar presented is merely a fragment.

Verbs Organised By Subcategorisation.

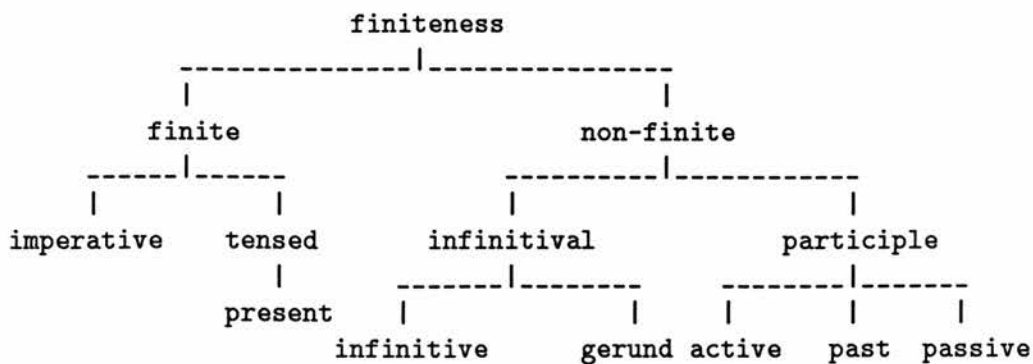
The hierarchy this encodes is the following



[Root, [verb, transitive, Fin], Agr] ISA [Root, [verb, intransitive, Fin], Agr]
 [Root, [verb, ditransitive, Fin], Agr] ISA [Root, [verb, transitive, Fin], Agr]
 [Root, [verb, complement, Fin], Agr] ISA [Root, [verb, intransitive, Fin], Agr]
 [Root, [verb, that-comp, Fin], Agr] ISA [Root, [verb, complement, Fin], Agr]
 [Root, [verb, tensed-comp, Fin], Agr] ISA [Root, [verb, that-comp, Fin], Agr]
 [Root, [verb, incomplement, Fin], Agr] ISA [Root, [verb, complement, Fin], Agr]
 [Root, [verb, en-comp, Fin], Agr] ISA [Root, [verb, incomplement, Fin], Agr]
 [Root, [verb, to-comp, Fin], Agr] ISA [Root, [verb, incomplement, Fin], Agr]
 [Root, [verb, inf-comp, Fin], Agr] ISA [Root, [verb, incomplement, Fin], Agr]
 [Root, [verb, ing-comp, Fin], Agr] ISA [Root, [verb, incomplement, Fin], Agr]

This section requires some explanation. This hierarchy is similar to the Relational Hierarchy given in Chapter 2. There it was mentioned that this hierarchy could also be thought of as the Subcategorisation Hierarchy. The reason for this is that each relation of the Relational Hierarchy is constrained as to what type of word may enter into that relation. For example, the relation *Subject* must be a noun. As such by ascribing words a feature from the Relational Hierarchy, we can use that feature for subcategorisation purposes. One difference between the Relational Hierarchy of Chapter 2, and the one just presented, is the inclusion of the ditransitive-transitive-intransitive branch. This is not essential, but was done to make greater use of the hierarchy as a taxonomic classifier. We could revert back to the features subject and object but we would then have to say for every verb that it required a subject.

Verb organised by Type, Subcategorisation, and Tense

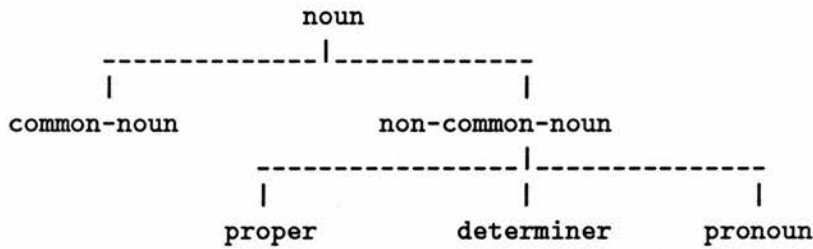


[Root, [verb, intransitive, finite], Agr] ISA [Root, [verb, intransitive, finiteness], Agr]
 [Root, [verb, intransitive, non-finite], Agr] ISA [Root, [verb, intransitive, finiteness], Agr]
 [Root, [verb, intransitive, tensed], Agr] ISA [Root, [verb, intransitive, finite], Agr]
 [Root, [verb, intransitive, imperative], Agr] ISA [Root, [verb, intransitive, finite], Agr]
 [Root, [verb, intransitive, present], Agr] ISA [Root, [verb, intransitive, tensed], Agr]
 [Root, [verb, intransitive, past], Agr1] ISA [Root, [verb, intransitive, tensed], Agr]
 [Root, [verb, intransitive, infinitival], Agr] ISA [Root, [verb, intransitive, non-finite], Agr]
 [Root, [verb, intransitive, participle], Agr] ISA [Root, [verb, intransitive, non-finite], Agr]
 [Root, [verb, intransitive, infinitive], Agr] ISA [Root, [verb, intransitive, infinitival], Agr]
 [Root, [verb, intransitive, gerund], Agr] ISA [Root, [verb, intransitive, infinitival], Agr]
 [Root, [verb, intransitive, active], Agr] ISA [Root, [verb, intransitive, participle], Agr]
 [Root, [verb, intransitive, passive], Agr] ISA [Root, [verb, intransitive, participle], Agr]
 [Root, [verb, intransitive, pastp], Agr] ISA [Root, [verb, intransitive, participle], Agr]

And finally to link the verbs up with the feature word

[Root, [verb, intransitive, finiteness], Agr] ISA [Root, [word, intransitive, finiteness], Agr]

We can do the same for nouns.



[Root, [common-noun, Def, Count], Agr] ISA [Root, [noun, Def, Count], Agr]
 [Root, [non-common-noun, Def, Count], Agr] ISA [Root, [noun, Def, Count], Agr]
 [Root, [proper, Def, Count], Agr] ISA [Root, [non-common-noun, Def, Count], Agr]
 [Root, [determiner, Def, Count], Agr] ISA [Root, [non-common-noun, Def, Count], Agr]
 [Root, [pronoun, Def, Count], Agr] ISA [Root, [non-common-noun, Def, Count], Agr]

and finally to link nouns up with words

[Root, [noun, Def, Count], Agr] ISA [Root, [word, Def, Count], Agr]

So far we have had little to say about the minor word classes concerning how entities are formed out of them and what features should be associated with them. For the minor word classes we limit ourselves to the following:

[Root, [preposition, nil, nil], Agr] ISA [Root, [word, nil, nil], Agr]

[Root, [adjective, nil, nil], Agr] ISA [Root, [word, nil, nil], Agr]

[to, [verb, inf-comp, to-infinitive], nil]

[that, [verb, Subcat, tensed], nil]

This then is a fragment of grammar concerned with taxonomy. We now need to include some subcategorisation statements. As the NEEDS statements will be those most frequently referred to into the grammar it will be useful to be able to refer to them:

- A. [Root, [word|X], Agr] NEEDS [oblig, head].
- B. [Root, [verb, intransitive, tensed], Agr] NEEDS not([oblig, head])
- C. [Root, [verb, intransitive, tensed], Agr] NEEDS [opt, head]
- D. [Root, [verb, intransitive, tensed], Agr] NEEDS [oblig, pre [Root2, [non-common-noun|B], [P, N, nominative]]
- E. [Root, [verb, transitive, Fin], Agr] NEEDS [oblig, post[Root2, [non-common-noun|B], [P, N, objective]]
- F. [Root, [verb, ditransitive, Fin], Agr] NEEDS [oblig, post[Root2, [preposition|X], Agr2]
- G. [Root, [verb, to-comp, Fin], Agr] NEEDS [oblig, post[to, [to|X], Agr]
- H. [to, [to-infinitive-comp, n], Agr] NEEDS [oblig, post[Root2, [verb, Subcat, infinitive], Agr]]
- I. [that, [verb, tensed-comp, tensed], X] NEEDS [oblig, post[Root2, [verb, Subcat, tensed], Agr]
- J. [Root, [preposition|X], Agr] NEEDS [oblig, post[Root2, [non-common-noun|Y], Agr]
- K. [Root, [determiner|X], Agr] NEEDS [oblig, post[Root2, [common-noun|Y], Agr]
- L. [Root, [verb, that-c, Fin], Agr] NEEDS [oblig, post[Root1, [verb, Subcat, tensed], Agr]

Bearing in mind our definitions lets us look at a few examples. For ease of exposition the choice function will be ignored, i.e it will be assumed that all words have only one lexical assignment. As such the symbol *C* will be dropped in the examples. One final point needs to be made. In our grammar, we have distinguished many Grammatical Relation in our Relation Hierarchy. These were used for the purposes of subcategorisation. However in our definitions, we only referred to *pre* and *post*. For the the purposes of the syntax of our grammar fragment, the only significant thing about the Grammatical relations is whether the are instances of pre or post dependents.

$$(41) S = \{John_1, sneezed_2\}$$

These words will receive the following lexical assignments in the grammar.

$$\hat{S} = \{ [John, [proper, definite, count], [3, singular, Case]-1, \\ [sneeze, [normal, intransitive, present], [3, singular, nil]-2] \}$$

We would also inherit from our grammar the following facts

$$\vec{S} = \{ [John, [proper, definite, count], [3, singular, Case]-1 \text{ NEEDS } [oblig, head] \text{ (from A)} \\ [sneeze, [normal, intransitive, present], [3, singular, nil]-2 \text{ NEEDS } [opt, head] \text{ (from C)} \\ [sneeze, [normal, intransitive, present], [3, singular, nil]-2 \text{ NEEDS } \\ [oblig, pre([Root2, [non-common-noun|X], [3,singular, nominative]])] \text{ (from D)} \}$$

From I_{pre} we can make the following deduction,

$$[sneeze, [normal, intransitive, present], [3, singular, nil]-2 \text{ NEEDS } \\ [oblig, pre([Root2, [non-common-noun|X], [3,singular, nominative]])] \text{ and } \\ [John, [proper, definite, count], [3, singular, Case]-1 \text{ NEEDS } [oblig, head] \text{ and } \\ [John, [proper, definite, count], [3, singular, Case]-1 \text{ ISA } \\ [Root2, [non-common-noun|X], [3,singular, nominative]] \text{ and } \\ 1 \text{ ; } 2$$

Depends([John, [proper, definite, count], [3, singular, nominative]-1,[sneeze, [normal, intransitive, present], [3, singular, nil])

This inference may then be placed in \vec{I} . For ease of exposition \vec{I} above will be simplified to the following

$$\vec{I} = \{ \text{Depends}(\text{John}, \text{sneeze}) \}$$

As there are no obligatory NEEDS statements in \vec{S} that do not have a counterpart in \vec{I} (though there is an optional statement in \vec{S} with no counterpart in \vec{I}) we can in accordance with our definition say that the string of words in S above forms a sentence.

Contrast the string *John sneezed* with the case where $S = \{ \text{sneezed}_1, \text{John}_2 \}$. In the latter case, the string would not have formed a sentence. This is because, despite the fact the both \hat{S} and \vec{S} would have been identical in both cases, there is no inference rule that would successfully apply in the second case. As such the conditions under which a string of words may form a sentence given in the previous section could not be met.

Let us work our way through some more examples. In the following I will ignore the fact that every tensed verb requires an optional head except where this head will be used.

$$(42) \quad S = \{ \text{John}_1, \text{loves}_2, \text{Mary}_3 \}$$

$$\vec{S} = \{$$

[John, [proper, definite, count], [3, singular, Case]-1 NEEDS [oblig, head], (from A)

[love, [normal, transitive, present], [3, singular, nil]-2 NEEDS

[oblig, pre([Root2, [non-common-noun|X], [3,singular, nominative]])] (from D)

[love, [normal, transitive, present], [3, singular, nil]-2 NEEDS

[oblig, post([Root2, [non-common-noun|B], [P, N, accusative]])] (from E)

[Mary, [proper, definite, count], [3, singular, Case]-3 NEEDS [oblig, head] (from A) }

$$\vec{I} = \{ \text{subject}(\text{John}, \text{love}), \text{object}(\text{Mary}, \text{love}) \}$$

$$(43) \quad S = \{ \text{John}_1, \text{put}_2, \text{Mary}_3, \text{on}_4, \text{the}_5, \text{chair}_6 \}$$

$$\vec{S} = \{$$

[John, [proper, definite, count], [3, singular, Case]-1 NEEDS [oblig, head], (from A)

[put, [normal, ditransitive, present], [3, singular, nil]-2 NEEDS

[oblig, pre([Root2, [non-common-noun|X], [3,singular, nominative]])] (from D)

[put, [normal, ditransitive, present], [3, singular, nil]-2 NEEDS

[oblig, post([Root2, [non-common-noun|B], [P, N, accusative]])] (from E)
 [put, [normal, ditransitive, present], [3, singular, nil]-2 NEEDS
 [oblig, post, [Root2, [preposition|X], Agr2], (from F)
 [Mary, [proper, definite, count], [3, singular, Case]-3 NEEDS [oblig, head] (from A)
 [on, [preposition, locative, nil], [nil, nil, nil]-4 NEEDS
 [oblig, post[Root2, [non-common-noun|X], Agr2], (from J)
 [on, [preposition, locative, nil], [nil, nil, nil]-4 NEEDS
 [oblig, head], (from A)
 [the, [determiner, definite, C], [3, singular, Case]-5 NEEDS
 [oblig, post[Root2, [common-noun|Y], Agr]. from (K)
 [the, [determiner, definite, C], [3, singular, Case]-5 NEEDS
 [oblig, head], (from A)
 [chair, [common-noun, definite, count], [3, singular, Case]-6 NEEDS
 [oblig, head] (from A) }
 $\vec{I} = \{ \text{subject}(\text{John}, \text{put}), \text{object}(\text{Mary}, \text{put}), \text{object}(\text{on}, \text{put}), \text{post}(\text{the}, \text{on}),$
 $\text{post}(\text{chair}, \text{the}) \}$
 (44) $S = \{ \text{John}_1, \text{wants}_2, \text{to}_3, \text{go}_4 \}$
 $\vec{S} = \{$
 [John, [proper, definite, count], [3, singular, Case]-1 NEEDS [oblig, head], (from A)
 [want, [normal, to-comp, present], [3, singular, nil]-2 NEEDS
 [oblig, pre([Root2, [non-common-noun|X], [3,singular, nominative]])] (from D)
 [want, [normal, to-comp, present], [3, singular, nil]-2 NEEDS
 [oblig, complement[to, [to|X], [3, singular, nil] (from G)
 [to, [to-infinitive, n], Agr]-3 NEEDS
 [oblig, inf-complement[Root2, [verb, Subcat, infinitive], Agr]] (from H)
 [to, [to-infinitive, n], Agr]-3 NEEDS
 [oblig, head], (from A)
 [go, [normal, i, infinitive], [nil, nil, nil]-4 NEEDS

[oblig, head], (from A)

$\vec{I} = \{ \text{subject}(\text{John}, \text{want}), \text{to-comp}(\text{to}, \text{want}), \text{infinitive-c}(\text{go}, \text{to}) \}$

(45) $S = \{ \text{John}_1, \text{wants}_2, \text{Mary}_3, \text{to}_4, \text{go}_5 \}$

$\vec{S} = \{$

[John, [proper, definite, count], [3, singular, Case]-1 NEEDS [oblig, head], (from A)

[want, [normal, to-comp, present], [3, singular, nil]-2 NEEDS

[oblig, pre([Root2, [non-common-noun|X], [3,singular, nominative]]) (from D)

[want, [normal, to-comp, present], [3, singular, nil]-2 NEEDS

[oblig, complement[to, [to|X], [3, singular, nil] (from G)

[Mary, [proper, definite, count], [3, singular, Case]-3 NEEDS

[oblig, head], (from A)

[to, [to-infinitive, n], Agr]-4 NEEDS

[oblig, inf-comp[Root2, [verb, Subcat, infinitive], Agr]] (from H)

[to, [to-infinitive, n], Agr]-4 NEEDS

[oblig, head], (from A)

[go, [normal, i, infinitive], [nil, nil, nil]-5 NEEDS

[oblig, head], (from A)

$\vec{I} = \{ \text{subject}(\text{John}, \text{want}), \text{object}(\text{Mary}, \text{want}), \text{to-comp}(\text{to}, \text{want}), \text{infinitive-c}(\text{go}, \text{to}) \}$

(46) $S = \{ \text{John}_1, \text{believes}_2, \text{that}_3, \text{Mary}_4 \text{ sneezed}_5 \}$

$\vec{S} = \{$

[John, [proper, definite, count], [3, singular, Case]-1 NEEDS [oblig, head], (from A)

[believe, [normal, that-comp, present], [3, singular, nil]-2 NEEDS

[oblig, pre([Root2, [non-common-noun|X], [3,singular, nominative]]) (from D)

[believe, [normal, that-comp, present], [3, singular, nil]-2 NEEDS

[oblig, post[Root2, [verb, Subcat, tensed], [3, singular, nil] (from L)

[that, [verb, tensed-comp, tensed], X]-3 NEEDS

[oblig, tensed-complement[Root2, [verb, Subcat, tensed], Agr] (from I)

[that, [that|X], X]-3 NEEDS

[oblig, head], (from A)

[Mary, [proper, definite, count], [3, singular, Case]-4 NEEDS

[oblig, head], (from A)

[sneeze, [normal, i, past], [P, N, C]-5 NEEDS

[oblig, pre([Root2, [non-common-noun|X], [3,singular, nominative]]) (from D)

[sneeze, [normal, i, past], [P, N, C]-5 NEEDS

[opt, head], (from A)

$\vec{I} = \{ \text{subject}(\text{John}, \text{believe}), \text{that-comp}(\text{that}, \text{believe}), \text{tensed-comp}(\text{sneeze}, \text{that})$

$\text{subject}(\text{Mary}, \text{sneeze}) \}$

(47) $S = \{ \text{John}_1, \text{believes}_2, \text{that}_3, \text{Mary}_4 \text{ sneezed}_5 \}$

$\vec{S} = \{$

[John, [proper, definite, count], [3, singular, Case]-1 NEEDS

[oblig, head], (from A)

[believe, [normal, that-comp, present], [3, singular, nil]-2 NEEDS

[oblig, pre([Root2, [non-common-noun|X], [3,singular, nominative]]) (from D)

[believe, [normal, that-comp, present], [3, singular, nil]-2 NEEDS

[oblig, post[Root1, [verb, Subcat, tensed], [3, singular, nil] (from L)

[Mary, [proper, definite, count], [3, singular, Case]-3 NEEDS

[oblig, head], (from A)

[sneeze, [normal, i, past], [P, N, C]-4 NEEDS

[oblig, pre([Root2, [non-common-noun|X], [3,singular, nominative]]) (from D)

[sneeze, [normal, i, past], [P, N, C]-4 NEEDS

[opt, head], (from A)

$\vec{I} = \{ \text{subject}(\text{John}, \text{believe}), \text{that-comp}(\text{sneeze}, \text{believe}), \text{subject}(\text{Mary}, \text{sneeze}) \}$

This then completes the examples for the grammar fragment. It will be apparent that the analyses presented leave a lot of questions open. This is because much work remains to be done. What I hope to have achieved however, is the following:

- (48) a a more rigorous formulation of Word Grammar
- b a more precise statement of Dependency, in terms of Dependency Structures
- c an explicit statement of the Selective Inheritance Principle
- d an explicit statement of the Adjacency Principle.
- e a formal statement of what a Dependency Grammar is
- f a statement of the language defined by the grammar

It is also my hope that in the course of presenting the grammar fragment and having worked through some examples, that I have demonstrated how some of the intuitions behind dependency as encapsulated in Word Grammar may be formally borne out. This chapter then concludes goals (1a) and (1b) of Chapter 1. We now move on to goal (1c) in the next two chapters, namely testing the formalism presented against a specific body of data

Chapter 4

Clitics

The phenomena of cliticisation present interesting analytical difficulties to any linguistic theory because it involves morphemes that are neither clearly independent words, in that their syntax is quite restricted, nor purely inflectional affixes, in that they have any syntax at all. The study of clitics, therefore, affords the possibility of insights into the interface of syntactic and morphological rules and principles. The phenomena seriously challenges the idea that there is an all or nothing distinction to be made between syntax and morphology (structuralists) or that the domain of morphology could be apportioned between syntax and phonology (generativists). The problems are made more perplexing by the pervasiveness of this phenomena throughout the languages of the world.

The aim of this chapter is to look, as descriptively as possible, at the problems presented by clitics in general, and more specifically at the clitic pronouns of Italian. Briefly, the major difficulties presented by clitics include the morphological/syntactic status of the clitic elements involved; the role played by clitics with regard to subcategorisation requirements and thematic structure especially in cases of clitic doubling (as in example (4) below): and how clitic placement and clitic movement interact with other syntactic processes. We will look at these and other problems in turn as they affect Italian later. First, we will look at little closer at the phenomenon of cliticisation in general.

4.1 Descriptive Generalisations About Clitics

Following Zwicky's (1977) seminal paper on the subject, I will distinguish three different types of clitics; *bound*, *simple* and *special*. As will be seen below the distinction between the latter two is not so much based on an opposition but should rather be seen as ends of a continuum.

4.1.1 Bound Clitics

Bound cliticisation involves cases where a morpheme that is always bound and always unaccented shows considerable syntactic freedom, in the sense that it can be associated with a wide variety of morphosyntactic categories. This contrasts on the one hand with the behaviour of affixes that tend to be associated with specific word classes and on the other with special clitics as we shall see below. I will adopt the standard term *host* to describe the word to which a clitic is attached. Frequently, such a bound word is semantically associated with an entire constituent yet is phonologically attached to only one word of the constituent. Usually the bound word is attached at the left or right margin of the constituent, standing outside inflectional affixes. An example of this type of cliticisation is the English possessive morpheme, 's. This is semantically associated with a noun phrase, and is attached phonologically to the last word which need not be either the head noun (1a) or even a noun at all (1b). It may even follow inflectional suffixes (1c).

- (1) a the Queen of England's hat
- b the woman I talked to's arguments
- c the woman I interviewed's arguments

In bound cliticisation there is no form that alternates with the clitic form, unlike cases of simple or special cliticisation, as we shall now see.

4.1.2 Simple Clitics

Simple cliticisation involves cases where a free morpheme, when unaccented, is phonologically reduced. The resultant form is then phonologically subordinated to a neighbouring

word. Cliticisation of this sort is usually associated with stylistic conditions, such as “casual” or “fast” speech. Examples of this can be found in the cliticisation of object pronouns in some varieties of English.

- (2) a he sees her
b he sees 'r

In such cases, there are both “formal”, *full* pronouns and “casual”, *reduced* pronouns, the full form being the one that appears in isolation or under emphasis. Significantly, this class of clitics has normal phonology, in that the rules relating the full and reduced forms are usually rules of general applicability in the language. Zwicky (1977) claims that this class of clitics also has the normal syntax, in that the reduced form occurs in the same positions as the full forms. It is probably more correct to say that simple clitics occur in a subset of the positions of their *full* counterparts. The point is that simple clitics have no *special* syntax.

4.1.3 Special Clitics

Special cliticisation involves cases where an unaccented, bound, *weak* form acts as a variant of a stressed, *strong* form with the same meaning and with *similar* phonological makeup. The clitic pronouns of Romance languages are standard examples of special clitics, though of course special cliticisation is not restricted to that class of words:

(3) **pronouns in French**
strong weak

moi	me	“me”
toi	te	“you”
lui	le	“him”

In this example, the strong forms are typically used when an accented pronoun is called for on semantic or syntactic grounds as in cases of emphasis or in isolation. Notice that though the strong and weak forms are phonologically similar, the precise phonological relationship between the weak and strong forms is (typically) not straightforward. It is doubtful that the two forms are related by phonological rules of any generality, unlike the relationship that simple clitics have to their strong counterparts. Also of note is the fact that under

certain conditions in some languages strong forms may co-occur with weak forms. This is called Clitic Doubling

- (4) je le vois lui
 I him-weak see HIM
 "I see him"

The above example also illustrates the fact that weak forms often display different syntax from their strong counterparts. The term *clitic* will be used to refer to cliticisation involving weak forms and the rest of this chapter will be devoted to an investigation of their special syntax.

4.1.4 Differences In The Order Of Clitics Relative To The Host

In strictly linear terms, we can distinguish three types of clitic; *proclitics*, *enclitics* and *endoclitics* (this latter term is due to Zwicky (1977)). Proclitics are those clitics that occur before their hosts, enclitics are those that occur attached to the end of their hosts, and endoclitics are those that occur within their hosts. The configuration of each type is given in the schema below (taken from Klavans 1979).

- (5) .
- | | | | | | | | | |
|------------|---|---------|---|------------|---|---------|---|-----------|
| [proclitic | + | [prefix | - | stem | - | suffix] | + | enclitic] |
| Group | | Host | | + | | + | | |
| | | | | \ | | / | | |
| | | | | endoclitic | | | | |

where "-" = affixation and
 where "+" = cliticisation

Arguments have been put forward by Klavans (1985) against the existence of endoclitics, analysing them instead as complex concatenations of either proclitics or enclitics. As I will not be concerned with those kinds of constructions, I will not give the arguments for or against, and no more mention will be made of them. From the above, then, we have the following typology of clitics:

(6)	I	II	III
	Bound	Simple	Special
Relationship to non-clitic variant	no variant	reduced	weak
A. proclitic			la mangio
B. enclitic	John's	I'd	mangiarla

In this investigation I will be mainly concerned with Italian (special) clitics. Italian has both weak proclitic and enclitic pronouns (i.e. types III A and III B) and it is these that I will be concerned with.

4.1.5 The Historical Perspective

Although this investigation will centre on clitics of type III A and B, it is important to bear in mind the close relationship between clitics of type II and III. As mentioned in the opening, simple and special clitics are best viewed as ends of a continuum rather than entities completely different in kind. This is due to the fact that type III are historically related to type II and are in a sense descended from them. This can lead to descriptive difficulties in analysing clitics, a point that is made by Zwicky:

“From the historical point of view, it should not be surprising that there is difficulty in drawing the line between simple and special clitics, since special clitics are often the remnants of an earlier system of simple clitics; this point is made by Givon (1971: 396-7) with respect to the clitic pronouns of modern French and Spanish, the ordering of which can be taken to reflect the object-before-verb order of earlier Romance... After the development from independent word to clitic, the next step is of course, the incorporation of clitics into morphology proper: what is a clitic at one stage is reinterpreted as a derivational or inflectional affix at the next.”

(Zwicky, 1977, p. 6)

However there is no clear evidence that the development is in one direction only or that having begun on the road from independent word to clitic, that the process may not be reversed.

4.1.6 The Classification Of Clitics

A question that naturally arises given the statement just made is which categories (nouns, verbs, and so on) can become clitics. From another perspective, this question may be seen

as enquiring as to what grammatical category to assign clitic elements. For the simple clitics, the answer is simple, namely that any word that can appear unaccented has the potential to cliticise to a neighbouring word. Obviously in such a case the word in question has the same grammatical category whether it is accented or not. The list of syntactic categories that can appear as *simple* clitics is therefore the same as the list of categories that we find cross-linguistically without accent. These include auxiliaries, personal pronouns, determiners, pre/postpositions, conjunctions, complementizers, and adverbs.

Given this it should be obvious from section 4.1.5 that the list of syntactic categories that can act as *special* clitics should be at least a subset of possible sources of simple clitics. But the answer is not always so simple. Some languages such as Tagalog, however, have very much richer systems of clitics. It is reported by Schachter (1974) to have some eighteen clitic particles in addition to agent and topic pronouns.

Further, there are also problems in deciding when the clitic has become totally incorporated into the inflectional morphology. At the point of incorporation into the morphology and points immediately prior to that point, assigning the clitic a syntactic category could present problems in terms of subcategorisation. It could be argued that Clitic Doubling is an example of this phenomenon. Additionally, while arguments can be put forward that clitic pronouns originate in (or belong in some sense to) the positions of their full NP counterparts and as such can be assigned the same syntactic category, for much richer and complicated cases, there may be no class of non-clitic constituents to which a clitic can be assigned, so that its syntactic source may not be at all clear. I will only be dealing with clitic pronouns, however, so this problem will not present difficulties.

4.2 The Pronominal System of Italian

In the remainder of this chapter, I will concentrate on the clitics of Italian, and at a restricted but significant set of Italian data involving clitics and cliticisation. Though, Italian has a complex system of clitics, I will strictly limit discussion to the personal pronominal clitics which will be listed below. The majority of the descriptive data concerning pronom-

inals presented below may be found in more detail in Brunet (1985) where many of the points raised receive a fuller discussion. I will not be concerned with any of the points of fine detail introduced there, but will be content to present a (slightly) simplified system of pronouns.

In the following sections, I will introduce the pronominal clitics in turn starting with the strong forms. Then I will introduce some of the syntactic phenomena in which these clitics are involved in as atheoretical and descriptive a way as possible. Finally I will summarise the problems that any account of the distributional behaviour of clitics must account for. When I come to give my own account it is against these benchmarks that it should be evaluated.

4.2.1 The Strong Pronominal Forms

Strong pronouns in Italian may only be used to refer to people. The following are subject pronouns of Italian:

(7) subject pronouns				
		sing.		pl.
1st person		io		noi
2nd person		tu		voi
3rd person	f.	lei	m/f	loro
	m.	lui		

Italian is a *pro-drop* language which generally speaking means that subjects are optional in sentences. In addition to the subject pronouns, Italian also has two forms of non-subject pronouns standardly called in the literature *tonic* and *atonic*, the tonic forms being strong forms and the atonic being the weak or clitic forms. I will keep the terms strong and weak (or clitic) as they seem clearer. We will look at both of these in turn.

(8) **the strong non-subject pronouns**

	object complement		accusative complement	
	sing.	pl.	sing.	pl.
1st person	me	noi	(di) me	noi
2nd person	te	voi	(di) te	voi
3rd person	f. lei	loro	(di) lei	loro
	m. lui	loro	lui	loro

The only difference in terms of form with the subject pronouns here occur in the first and second persons singular. There is some debate over the status of what I have called object pronouns. The debate concerns amongst other things whether they are in fact subject pronouns. As I will be concerned mainly with the weak personal pronouns, to enter into the debate here is unnecessary. I will now pass quickly on to the weak forms as it is these that will occupy us for the rest of this chapter.

4.2.2 The Weak or Clitic Pronominal Forms

The weak pronouns, unlike their strong counterparts may be used to refer to things as well as people. They are always used without prepositions and can be used as dative or accusative complements:

(9) **The clitic pronouns**

	dative		accusative	
	sing.	pl.	sing.	pl.
1st person	mi	ci	mi	ci
2nd person	ti	vi	ti	vi
3rd person	m. gli	loro	lo	li
	f. le	loro	la	le

From (9) it is apparent that *loro* is the form of both the masculine and feminine plural, dative clitic. It is also the same as the strong form of the non-subject third person plural pronouns given in (8) above as well as the third person plural subject pronoun, given in (7). Given this it is obvious that the third person plural clitic pronoun must have its own stress, if the strong forms have their own stress. Because the third person plural dative clitic has its own stress, unlike all the other clitic pronouns, it becomes questionable whether it is in

fact a clitic, and in fact its syntax resembles more closely that of its non-clitic counterparts. This could be due to factors mentioned in the historical perspective section namely that the third person dative clitic pronoun has not (yet) undergone or has resisted the relevant changes. This is further evidenced by the fact that the third person singular dative clitic *gli* is encroaching heavily upon the territory of its plural counterpart and appears to be slowly replacing it. As a result I will not be too interested in trying to include *loro* into my account.

4.2.3 The Form Of The Clitics In Clusters

When the dative clitic pronouns *mi*, *ti*, *ci*, *vi*, occur in conjunction with the accusative clitic forms *lo*, *la*, *li*, *le*, the final *i* of the dative clitic pronouns changes to an *e*. This orthographic change is accompanied by a similar phonological change. In such clusters, it is the dative that precedes the accusative.

(10) quando me+lo + dici tu, lo + credo
 when to-me+it say-2nd-sing you, it believe-1st-sing
 “when *you* tell me, I believe it”

(11) te lo prometto
 to-you it promise-1st-sing
 “I promise you (it)”

The two third person dative forms, *gli* (masculine) and *le* (feminine), when occurring in clusters assume a unique form *glie* which serves for both genders. There is also an accompanying phonological change that occurs that is reflected in the orthography.

(12) quest’ offesa, non glie+la + perdonero mai
 this insult not to-him/her+it will-forgive-1st-sing ever
 “I will never forgive him/her this insult”

4.3 The Syntax of Italian Clitics

Before we can look at the syntax of clitics it will be necessary to give a little background to Italian syntax in general. In the interests of simplicity the approach taken towards Italian is in some respects unorthodox. We will assume that nominal complements, (subject, object, etc) in Italian are free dependents of their verb, in the sense that their order relative to the

verb is unconstrained, provided that the Adjacency Principle is not violated. This gives us the following possible word orders for simplex sentences:

- (13) a S V O (io) mangio la torta
b S O V (io) la torta mangio
c O S V la torta (io) mangio
d O V S la torta mangio (io)
e V S O mangio (io) la torta
f V O S mangio la torta (io)

I eat the cake

In the simplex sentences above, though they differ in intonation patterns, all are acceptable to a large degree. Now while it is not true in general that nominal complements are free there are two reasons why making such an assumption will not affect the discussion. Firstly my interest is in clitics, and there are no subject clitics. My interest in subjects is therefore limited to their interaction with clitics which is limited.

Second, clitic movement, to the extent that it happens in Italian is bound, within a tensed clause. Within a tensed clause, the majority of the positions opened to clitic objects is a subset of those open to non-clitic objects. Those positions open to clitics but not to full NP's can arguably be sanctioned on other grounds. As the current account does not attempt to be exhaustive and the phenomena in question are limited in distribution, these details need not concern us here and the assumption that objects are free to occur in positions where they do not violate the Adjacency Principle will be maintained.

4.3.1 The Non-Exceptional Behaviour Of Clitics

The non-exceptional behaviour of clitics will be used to refer to cases where a clitic pronoun is used, and is attached to a verb that subcategorises for it. In such cases, there is no other complexity to be accounted for except its restricted position. With respect to the non-exceptional behaviour of clitics, the data that has to be accounted for include the following.

- (14) a io odio la torta
 I hate the cake
 b io la + odio
 I it hate
 "I hate it"
 c *io odio + la
 d *la + io odio
 e la + odio io
 f *odio io la +
 g *odio + la io

In (14a) we see a simplex sentence containing a transitive verb and two arguments. Given the assumption that nominal complements are free in Italian (14c), (14d), (14f) and (14g) represent positions that are not available to the clitic pronoun but would be available to the corresponding non-clitic NP. The sentences in (14) present us with our first problem. It seems reasonable to assume that the clitic is the object of the verb and we want to analyse it as such. Yet on the paradigm given in (13) how is it that the clitic cannot occur after the verb in (14c), (14f) and (14g) or before the subject, when the subject is before the finite verb (14d) as the non-clitic NP could?

It seems as though clitic objects are restricted to coming immediately before the verb. In fact, in simplex sentences it is standardly the case that the clitic forms of the personal pronouns (with the exception of *loro*) must occur procliticised on the finite verbs of which they are arguments.

However, when the verb is not tensed i.e. is in the infinitive, the gerundive or the imperative form, the clitic pronoun appears as an enclitic on the verb of which it is an argument. In the infinitival cases, the verb will typically lose its final vowel. I will ignore however, such phonological changes. We note the following:

- (15) a odio mangiare la torta
 hate-1st-sing to-eat the cake
 "I hate to eat the cake"
- b odio mangiar+la
 hate-1st-sing to-eat+it-clitic
 "I hate to eat it"
- c da la torta a me!
 give-2nd-sing-Imperative the cake to me!
 "give the cake to me!"
- d da+mmi la torta!
 give-2nd-sing-Imperative+to-me the cake!
 "give me the cake!"
- e camminava, mangiando la torta
 walk-3rd-sing-Imperfect, eating-Ger the cake
 "he was walking, eating the cake"
- f camminava, mangiando+la
 walk-3rd-sing-Imperfect, eating-Ger+it
 "he was walking, eating it"

In the non-exceptional case we can thus state quite simply that clitic pronouns must appear either as proclitics on tensed verbs of which they are arguments, in which case they are written as two words or as enclitic on (a subset of the) non-tensed verbs of which they are arguments, in which case they form a single orthographic unit with the verb. Again Brunet (op. cit.) notes some discrepancies but these will not concern us.

4.3.2 Clitic Climbing

At first glance, it seems as though clitics have the element that subcategorises for them as host and that they must either be proclitic or enclitic on the latter depending on its finiteness. However, the situation is not this simple. Irregularities appear with respect to certain syntactic processes, which chiefly for our purposes, involve cliticisation. Some main verbs that take infinitival complements allow a clitic pronoun, originating (in the sense of being subcategorised for) on the embedded infinitival complement, to be cliticised either to the embedded infinitival verb itself or to climb up and be cliticised onto the main verb. This phenomenon does not occur with all verbs. Other main verbs in the same constructions do not allow the clitic to climb and as such cliticisation is restricted to the embedded infinitival verb. The class of verbs that allows this Clitic Climbing will be called

restructuring verbs after the terminology introduced by Rizzi (1978). The contrast between restructuring and non-restructuring main verbs is illustrated in (16) and (17) below:

- (16) a Giorgio vuole mangiare la torta
 Giorgio wants to-eat the cake
- b Giorgio vuole mangiar+la
 Giorgio wants to-eat+it-clitic
- c Giorgio la + vuole mangiare
 Giorgio it-clitic wants to-eat
 "Giorgio wants to eat it"
- (17) a Giorgio odia mangiare la torta
 Giorgio hates to-eat the cake
- b Giorgio odia mangiar+la
 Giorgio hates to-eat+it-clitic
- c *Giorgio la + odia mangiare
 Giorgio it-clitic hates to-eat
 "Giorgio hates to eat it"

In the above, we find that some main verbs allow Clitic Climbing (16c) whilst others do not (17c). Clitic Climbing is not restricted to a subset of the main verbs though. We find the same phenomenon with the auxiliaries. In Italian, the major division among the class of verbs is between the two auxiliaries *avere* (to have) and *essere* (to be) and the main verbs. However within the class of main verbs, there appear certain irregularities in those verbs taking infinitival complements. In the simple case, where an auxiliary takes a participle of a transitive verb the situation is such that a clitic pronoun, originating on, or belonging to the participle, can only occur as a proclitic on the auxiliary itself and not on the participle of which it is an argument. No other order is permitted. Given that participles are non-tensed forms of the verb we might expect them to allow enclitics, but this is not the case. Auxiliaries thus also belong to the class of restructuring verbs. We can see the general pattern in (18) below:

- (18) a *ho mangiato + lo
 AUX-(have)-1st-sing eaten it
- b lo + ho mangiato
 it AUX-(have)-1st-sing eaten
 "I have eaten it"

Where the participle requires a verbal complement, we have a contrast between cases where that participle is a restructuring verb (19c) and cases where it is not (20c).

- | | | | | | | |
|------|---|-----------------------------|----------------|------------------|----------|-------|
| (19) | a | ho | voluto | mangiare | il | pane |
| | | AUX-1st-sing | want-past-part | to-eat | the | bread |
| | | “I wanted to eat the bread” | | | | |
| | b | ho | voluto | mangiar+lo | | |
| | | AUX-1st-sing | want-past-part | to-eat+it-clitic | | |
| | | “I wanted to eat it” | | | | |
| | c | lo + | ho | voluto | mangiare | |
| | | it-clitic | AUX-1st-sing | hated-past-part | to-eat | |
| | | “I wanted to eat it” | | | | |
| | | | | | | |
| (20) | a | ho | odiato | mangiare | il | pane |
| | | AUX-1st-sing | hate-past-part | to-eat | the | bread |
| | | “I hated to eat the bread” | | | | |
| | b | ho | odiato | mangiar+lo | | |
| | | AUX-1st-sing | hate-past-part | to-eat+it-clitic | | |
| | | “I hated to eat it” | | | | |
| | c | *lo + | ho | odiato | mangiare | |
| | | it-clitic | AUX-1st-sing | hate-past-part | to-eat | |

In (19c) and (20c) we find that the clitic object that “belongs” to *mangiare* has not cliticised onto it though this is a possibility as we see in (19b) and (20b). Instead it has climbed up beyond the past participle (which can not serve as a host) to the left of the first person singular present form of the auxiliary *avere*, namely *ho*. Here as we see, Clitic Climbing is licensed by some main verbs but not by others. The examples in (19) and (20) also suggest that what is required is a *hopping* not a *swooping* analysis of Clitic Climbing. From (18) we can see that the auxiliary *avere* allows cliticisation. A swooping analysis would thus license both (19c) and (20c) equally. A hopping analysis, on the other hand, would distinguish (19c) from (20c) in virtue of forcing the clitic to hop onto an intervening non-restructuring participle before moving onto the auxiliary (despite the fact that participles can not serve as hosts).

Further evidence for a hopping analysis comes from the prediction made by such an analysis that in verbal complexes a clitic in the course of its climb could cliticise onto any element that could serve as a host. We would also expect that a clitic could climb up to, but not onto or past, any non-restructuring verb. Confirmation of this is found in the

examples in (21) and (22) below where *voglio* and *dover* are (forms of) restructuring verbs but *odio* is not.

- | | | | | | |
|------|---|----------------------------|----------------------|------------------|----------|
| (21) | a | voglio | dover | mangiar+lo | |
| | | I-want | to-have-to | to-eat+it-clitic | |
| | b | voglio | dover+lo | mangiare | |
| | | I-want | to-have-to+it-clitic | to-eat | |
| | c | lo + | voglio | dover | mangiare |
| | | it-clitic + | I-want | to-have-to | to-eat |
| | | "I want to have to eat it" | | | |
| | | | | | |
| (22) | a | odio | dover | mangiar+lo | |
| | | I-hate | to-have-to | to-eat+it-clitic | |
| | b | odio | dover+lo | mangiare | |
| | | I-hate | to-have-to+it-clitic | to-eat | |
| | c | *lo + | odio | dover | mangiare |
| | | it_clitic + | I-hate | to-have-to | to-eat |
| | | "I hate to have to eat it" | | | |

As we see, the example in (22c) the clitic can climb up to but not past *dover*. It is obvious then that any solution needs to state more than that the matrix verb does or does not allow clitic climbing. The set of verbs that allow clitic climbing are the modals, aspectuals and auxiliaries which include the following:

- | | | | |
|------|---|----------|------------------------------------|
| (23) | a | essere | "to be" |
| | b | avere | "to have" |
| | c | dovere | "to have to" or "must" |
| | d | potere | "to be able" or "can" |
| | e | sapere | "to know" |
| | f | volere | "to wish" or "to want" |
| | g | fare | "to do" or "to make" |
| | h | lasciare | "to leave", "to allow" or "to let" |
| | i | sentire | "to feel" or "to hear" |
| | j | andare a | "to go to" |
| | k | stare a | "to be at" |
| | l | venire a | "to come to" |

4.3.3 Object Agreement

There is a further complication with the auxiliary constructions involving the interaction of Clitic Climbing and the form of the participle when the clitic is an object pronominal. This interaction is called Object Agreement. Before going in to detail it should be noted that this phenomenon is not strictly obligatory and seems to be in the process of change at

the moment, occurring more in writing. This is especially true when the first and second person clitics are involved.

There are four forms of the past participle in Italian. If we take a verb like *mangiare* it would have the following four past participle forms:

- (24) a mangiato eaten-masculine-singular
 b mangiata eaten-feminine-singular
 c mangiati eaten-masculine plural
 d mangiate eaten-feminine-plural

There are two conditions under which a participle will display agreement. In the first case, the forms in (24) in conjunction with an auxiliary (usually *avere*), may be used in cases when a clitic object pronoun is proclitic on the auxiliary. In such cases the participle obligatorily agrees with the clitic in number and gender. This is illustrated in (25) below.

- (25) a Lei ha mangiato la torta/le torte
 she AUX-3rd-sing eaten-neut the cake/s-fem-sing/pl
 "she ate the cake/s"
 b *Lei ha mangiata/e la torta/le torte
 she AUX-3rd-sing eaten-fem-sing/pl the cake/s-fem-sing/pl
 c Lei la/le + ho mangiata/e
 she it-3rd-sing/pl-fem AUX-3rd-sing eaten-fem-sing/pl
 "She ate it"
 d *Lei la/le + ho mangiato
 she it-3rd-sing/pl-fem AUX-3rd-sing eaten-neut
 "She ate it"

This phenomenon does not occur with the dative clitic and hence this is why it is called Object Agreement. If a non-clitic object NP is used, however, there is no agreement displayed on the participle even if the object appears before the auxiliary. In this case the participle always takes the masculine singular form. In fact, it could be argued that there are five forms of the past participle, the extra one being a neuter form that is used for cases other than the one in which the clitic object is proclitic on the auxiliary. This extra form would be identical with the masculine singular form. I will not be concerned to argue one way or the other on this point here but will adopt the fifth form for ease of exposition. In slightly more complicated cases containing an auxiliary, a participle and an infinitive we find the following situation.

- (26) la + ho potuta mangiare
 it-clitic-fem Aux-3rd-sing participle-can Infinitive-eat
 "I could eat it"

Here Object Agreement is again on the participle *potuta* even though it is not the verb that the clitic object belongs to. This shows that Object Agreement is not necessarily connected to subcategorisation, as the term might suggest.

There is a second case where we find agreement on the participle. Verbs that take *essere* as aspectual auxiliary inflect the participle for agreement with the subject as in (27).

- (27) a Lui é andato
 he AUX-3rd-sing go-3rd-sing-masc
 "he has gone"
 b Lei é andata
 she AUX-3rd-sing went-3rd-sing-fem
 "she has gone"
 c Loro sono andati/e
 they AUX-3rd-pl-masc/fem went-3rd-pl-masc/fem
 "they have gone"
 d *Lei é andato
 she AUX-3rd-sing went-3rd-sing-neut

The fact that verbs that take *essere* as auxiliary have subject agreement, while verbs that take *avere* display Object Agreement under certain circumstances, might lead one to believe that conflicts would be constantly appearing. In fact, such cases are extremely rare. This is largely due to the almost complementary distribution of the kind of verbs that each auxiliary may take. This will be discussed in a little more detail below where, we will see an interesting example.

4.3.4 Change Of Auxiliary

In Italian, all verbs either take *avere* or *essere* as aspectual auxiliary.

(28)	a	ha	voluto	una	macchina	nuova
		Aux-(to-have)-3rd-sing	wanted	a	car	new
	b	*é	voluto	una	macchina	nuova
		Aux-(to-be)-3rd-sing	wanted	a	car	new
		"he wanted a new car"				
	c	*ha	andato	a	Roma	
		Aux-(to-have)-3rd-sing	went	to	Rome	
	d	é	andato	a	Roma	
		Aux-(to-be)-3rd-sing	went	to	Rome	
		"Fabio went to Rome"				

However, some main verbs that usually take *avere* as auxiliary can optionally change their auxiliary and take *essere* when an embedded verb requires *essere*, (29a) and (29b). Other main verbs do not allow this, (29c) and (29d). (see Zubizaretta (1982, p156) or Burzio (1981, p148) for attempts at formulations of the rule of auxiliary assignment)

(29)	a	Maria	ha	voluto	andare	a	Roma
		Maria	Aux-(to-have)-3rd-sing	wanted	to-go	to	Rome
	b	Maria	é	voluto	andare	a	Roma
		Maria	Aux-(to-be)-3rd-sing	wanted	to-go	to	Rome
		"Maria wanted to go to Rome"					
	c	Maria	ha	promesso	di andare	a	Roma
		Maria	Aux-(to-have)-3rd-sing	promised	to-go	to	Rome
	d	*Maria	é	promesso	di andare	a	Roma
		Maria	Aux-(to-be)-3rd-sing	promised	to-go	to	Rome
		"Maria promised to go to Rome"					

This phenomenon is quite complex. I shall not attempt to account for the Change Of Auxiliary as I am more concerned with its interaction with clitics.

4.3.5 Interaction Of Clitic Climbing, Change Of Auxiliary And Object Agreement

While at first sight Clitic Climbing and Change of Auxiliary seem to be two completely unrelated phenomena, closer investigation reveals that those verbs that allow Clitic Climbing are the same set that allow Change of Auxiliary, namely the restructuring verbs. Further it is claimed by, amongst others, Rizzi (1979) that both phenomena interact directly with one another. Rizzi claims that Clitic Climbing will not occur if Change of Auxiliary has failed to occur. Hence he gives the following paradigm.

- (30)
- | | | | | |
|---|--------|----------------------------|--------|---------------------|
| a | Mario | avrebbe | voluto | venir+ci |
| | Mario | AUX-(avere)-3rd-sing-Cond | wanted | to-come+here-clitic |
| b | *Mario | ci + avrebbe | voluto | venire |
| | | | | |
| c | Mario | sarebbe | voluto | venir+ci |
| | Mario | AUX-(essere)-3rd-sing-Cond | wanted | to-come+here-clitic |
| d | Mario | ci + sarebbe | voluto | venire |
- “Mario would have wanted to come here”

However, there is an intriguing set of sentences that emerged while trying to create a conflict between subject and object agreement. From section 4.3.3 we know that in Italian, participial forms of verbs conjugated with the auxiliary *essere* must agree with the subject, whereas participles, which object clitics have climbed past, agree with that clitic. We also know from section 4.3.4 that, in certain constructions, verbs are allowed to change their auxiliary from *avere* to *essere*. Rizzi claims that Change Of Auxiliary and Clitic Climbing interact directly in the manner just described. However, let us look at the following sentences that allow Change Of Auxiliary:

- (31)
- | | | | | | |
|---|--|--------------|-----------------------|--------|-----------------|
| a | Lui | avrebbe | voluto | andare | a prender+la/lo |
| | he | Aux-have-fut | wanted-neuter | to-go | to-get+her/him |
| | “he would have wanted to go and get him/her” | | | | |
| b | Lei | avrebbe | voluto | andare | a prender+la/lo |
| | she | | | | |
| c | Loro | avrebbero | voluto | andare | a prender+la/lo |
| | they | | | | |
| d | Lui | sarebbe | voluto | andare | a prender+la/lo |
| | he | Aux-be-fut | wanted-masc-sing | to-go | to-get+her/him |
| e | Lei | sarebbe | voluta | andare | a prender+la/lo |
| | she | | wanted-fem-sing | | |
| f | Loro | sarebbero | voluti/e | andare | a prender+la/lo |
| | they-m/f | | wanted-masc-pl/fem-pl | | |

The above pairs of sentences, one with the auxiliary *essere* and the other with *avere*, all have the same meaning but the ones containing *essere* (to be) exhibit subject agreement. Now my interest in these sentences was to try and create a situation whereby the participle has to agree simultaneously with the subject and object, thereby creating a conflict and to see how this conflict would be resolved. This can be done by taking the sentences above containing *essere* and making the clitic *climb* from its embedded position to a position where it is attached to the auxiliary. For example:

- (32) a Lui la + sarebbe voluto andare a prendere
 b Lui la + sarebbe voluta andare a prendere
 c Lui lo + sarebbe voluto andare a prendere
 d *Lui lo + sarebbe voluta andare a prendere
- (33) a Lei la + sarebbe voluta andare a prendere
 b *Lei la + sarebbe voluto andare a prendere
 c Lei lo + sarebbe voluta andare a prendere
 d Lei lo + sarebbe voluto andare a prendere
- (34) a Loro la + sarebbe voluta andare a prendere
 b Loro la + sarebbe voluti andare a prendere
 c Loro la + sarebbe volute andare a prendere
 d Loro lo + sarebbe voluto andare a prendere
 e Loro lo + sarebbe voluti andare a prendere
 f Loro lo + sarebbe volute andare a prendere

These sentences are confusing for the native speaker due to the subject-object agreement conflict and so could be described as marginal. I have only starred those examples that are actually impossible, as the rest have no clear status. Even in the examples which contain no conflict of gender agreement on the participle, the sentences are odd. In fact, the construction is much more acceptable *without* the change of auxiliary having taken place as in (35) below:

- (35) a Lui la + avrebbe voluta andare a prendere
 b Lui lo + avrebbe voluto andare a prendere
 c Lei la + avrebbe voluta andare a prendere
 d Lei lo + avrebbe voluto andare a prendere
 e Loro la + avrebbero volute andare a prendere
 f Loro lo + avrebbero voluto andare a prendere

The use of *avere* here instead of *essere*, i.e. of not changing the auxiliary, has the added advantage of resolving the conflict. When conjugated with *avere*, past participles do not agree with the subject. As a result, the only agreement that needs to be made is with the object clitic. This, however, refutes Rizzi's claim that Clitic Climbing will not occur without Change Of Auxiliary. It should still be remembered though that the paradigm given in (30) above is correct, but notice that in that example it was not an object clitic that was used but the locative *ci*. Such clitics are not object clitics and are arguably neuter or unspecified for number, person and gender and so no conflict would be involved. This leaves open the question of why the Change Of Auxiliary is obligatory in the examples in (30).

This implies that Rizzi's generalisation that Clitic Climbing can only happen with Change of Auxiliary is incorrect or that at least it must be weakened to not include personal pronouns. Taking the weaker generalisation, i.e. the one that applies only to locatives, would have the consequence of differentiating between locatives and other pronouns and it is not clear at the moment if this is a realistic position to defend.

Examples (32) to (35) are interesting for various reasons. It has generally been assumed, at least since Rizzi, that i) the clitic pronouns discussed here come under the same or similar rules as those of the reflexive pronouns, the partitive pronoun *ne*, the impersonal pronoun *si* and the locative *ci* and *vi* not discussed and ii) Clitic Climbing and Change of Auxiliary interact as described above.

4.4 The Relative Ordering Of Clitics

There can also be more than one clitic in a sentence in Italian. This happens in cases where we have a ditransitive verb and we want to pronominalise both its objects. In such cases where there is more than one clitic in a tensed clause, the clitics form a complex and are inseparable, and as such do not have the individual freedom of movement that their non-pronominal counterparts enjoy. The relevant data are presented below:

(36)	a	non	posso	dir	+te	+lo	ancora
		not	can-1st-sing	say-Infinitive	to-you	it	yet
		"I can not tell it to you yet"					
	b	non	te +	lo +	posso	dire	ancora
	c	*non	lo	posso	dirte	ancora	
	d	*non	ti	posso	dirlo	ancora	

This has two implications; firstly, that even when the clitics are arguments of different verbs¹ they must form a complex and secondly, that if one of them moves then so must the other.

¹An example of this will be seen in (39) of the next chapter.

4.4.1 Summary and Benchmarks

To summarise, then, the following points must be addressed by any adequate account of the distributional properties of the pronominal clitics of Italian.

- (37)
- a Clitics pronouns can not occur in all the positions that full NP's can and are less free than the nouns they stand for
 - b The non-exceptional position of clitics is proclitic on tensed verbs that subcategorise for them and enclitic on non-tensed forms of the verbs. However they do not cliticise onto participles
 - c Under certain circumstances clitics can *climb*;
 - 1. with modals and aspectuals with no object agreement
 - 2. with auxiliaries, which also induce Object Agreement on participles which are not necessarily their heads
 - 3. other main verbs do not allow Clitic Climbing
 - 4. constructions that contain non-restructuring verbs allow Clitic Climbing up to but not including the non-restructuring verb
 - d Under certain circumstances some verbs allow a Change Of Auxiliary, others do not
 - e The conditions under which Clitic Climbing and Change Of Auxiliary interact
 - f The morphological/syntactic status of clitics in Italian
 - g The relative ordering of clitics is fixed and they form inviolable complexes
 - h Clitics are bound within a tensed clause

Chapter 5

The Syntax of Clitics In Word Unification Grammar

In this chapter, I will give an account of the phenomena presented in Chapter 4. Before we attempt to account for those phenomena, it will be useful to survey briefly previous accounts. This will be helpful not only in terms of drawing on previous insights, but also as a further metric of adequacy of the solution to be presented below.

5.1 Previous Analyses

The majority of the work done on clitics (in Italian) has been done within the Government and Binding framework (Chomsky 1981). There are four major solutions proposed to the problem of restructuring in general, though we are primarily interested in how restructuring is implicated in the syntax of clitics. These are (a) a restructuring rule, Rizzi (1978) (b) VP-movement, Burzio (1981) (c) base generation, Cinque (1980) Burzio (1986) and (d) parallel structures, Zubizarreta (1982) Manzini (1983). In the following brief presentation, background to the theory of Government and Binding theory will be assumed, though as we will be more concerned with any insights gained into the phenomena rather than particular formulations of those insights, a knowledge of the technical details of Government and Binding will not be necessary.

As was mentioned in Chapter 3, the terminology of “restructuring” and “restructuring verbs” goes back to Rizzi’s seminal work on the subject. In the examples of main verbs and auxiliaries (given in the previous chapter) that exhibit exceptional behaviour, Rizzi argues for a rule of restructuring governed by this subclass of verbs, suitably called restructuring verbs. This rule *optionally reanalyses* a terminal string $Vx (P) V$ as a single complex (where Vx is a restructuring verb, P is the infinitival complementiser, and V is the infinitival verb). The effect of this rule is to transform the underlying bisentential sentence signified by the two S nodes in (1a)¹ into a single verbal complex (1b) along the following lines:

- (1) a $[S[_{NP} \text{io}] [_{VP}[_{Vx} \text{voglio}] [S[_{NP} \text{e}] [_{VP} [V \text{mangiare}] [_{NP} \text{la}]]]]]$
 b $[S[_{NP} \text{io}] [_{VP}[_{Vx} \text{voglio mangiare}] [_{NP} \text{la}]]]$

In the above example, (1a) represents the regular syntactic structure and (1b) the structure after the rule of restructuring has taken place. Ignoring irrelevant detail, Rizzi assumes that a clitic originates in NP position and is then placed as far to the left of its immediately dominating VP as general conditions allow. This will result in the enclitic position on the infinitival (i.e. after *mangiare*) in 1a) as this is the furthest left position possible within the VP. In 1b) the clitic must move to the proclitic position (i.e. before *voglio*) due to the fact that after restructuring has taken place, this is the furthest left position.

Later developments in Government and Binding, especially those connected with the Projection Principle, ruled out restructuring-like analyses on theory internal grounds. Though the analysis was rejected the terminology persisted. There are other theory external problems with this account. Remembering the examples (31) to (35) of the previous chapter, and that one of the implications of the analysis in (1) above is that a clitic can not climb unless restructuring has taken place, we note that the analysis does not have empirical support.

In many ways the solutions given after Rizzi’s may be seen as attempts to encode his basic intuition within the terms of the developing theory. Only Burzio’s VP-movement analysis departed significantly from Rizzi’s basic approach as shall become apparent below.

¹In this structure $[e]$ is an empty category a theoretical primitive of Government and Binding the precise nature of which need not concern us here

The main problem appeared to be how to encode the optionality displayed in (1) into a unique lexical entry as required by Government and Binding theory. The details of the various analyses need not detain us but the general consensus of the approach taken, as it relates to the points in (37) of Chapter 3, is worth dwelling on.

On the question of the syntactic/morphological status of clitics, most treatments of Italian in Government and Binding assign a thematic role to the clitic pronoun. The notion of thematic role in Government and Binding is an incorporation into the syntactic component of what has been called participant roles by Jackendoff (1987). Two facts ensure that Government and Binding approaches consider clitics as independent words. Firstly, the kinds of things assigned thematic roles are generally considered to be words or phrases, not affixes. Secondly, given that syntax is seen as autonomous in Government and Binding and that Theta theory is an element of the syntactic apparatus, it must be the case that clitics enter into the syntax and hence are words.

With reference to the syntax of clitics, Burzio (1986) states that there are two major questions:

“One concerns the nature of the locality conditions holding between the clitic and the position which receives the theta-role. The other question is whether clitics are moved or base generated.”

(Burzio 1986, p221)

We may interpret this first question as asking what the configurational possibilities are between a clitic and an element that subcategorises for it (i.e. what positions may clitics move to) and the second as asking whether clitic (movement) should be handled with the same devices as other kinds of movement.

Burzio observes that cliticisation appears to be subject to the Specified Subject Constraint. With respect to movement, the Specified Subject Constraint says that only (parts of) subjects may be construed with elements outside of a clause (S-bar), or NP, that contains a (specified) subject (with the usual exceptions for *wh*-pronouns)². Leaving such details aside as what precisely a specified subject is in a pro-drop language such as Italian, the fact that clitics seem to obey the Specified Subject Constraint means that, as there

²I have not been concerned with clitic placement within NP

are no subject clitics, clitic movement is restricted to within S-bar. We may thus take the statement concerning clitics and the Specified Subject Constraint as trying to capture the fact that clitics are restricted to occurring within the same tensed clause as the elements that subcategorise for them.

On the question of movement or base generation, all the analyses since Rizzi, excluding Burzio (1981) but including Burzio (1986) have assumed some kind of base-generation. The accounts in terms of parallel structures may be seen as a way capturing the intuition that restructuring verbs are specified in the lexicon as both restructuring and non-restructuring verbs simultaneously; hence the term parallel structure. Other frameworks might have handled this by giving the restructuring verbs two lexical entries. Burzio (1981) on the other hand has a VP-movement analysis of cliticisation.

Jaegli (1982) and Borer (1981) have claimed that Romance object clitics *must* be base-generated. The examples they use come from cases of Clitic Doubling in Spanish. In such cases, it would appear that a clitic has no site available from which to move as the only potential site is occupied by the full NP, and hence must be base-generated in its surface position³. Burzio (1986) though, presents different arguments for base-generation and is more specifically concerned with the clitics of Italian. He states that

“... the base-generation and movement hypotheses actually predict different configurations of data with respect to complex predicates.”

(Burzio 1986, p223)

On theory internal argumentation which need not concern us here, he concludes like Borer and Jaegli that clitics must be base-generated. Base generation then seems to be the current consensus.

In summary, what we may extract from the Government and Binding accounts is that they treat clitics as words not affixes, that clitics and the verbs they are arguments of, must be within the same S-bar (for our purposes, tensed clause) and that their behaviour is to be explained with devices other than those that account for unbounded movement. Finally, that an analysis of restructuring seems to require certain verbs to be simultaneously

³Whether Clitic Doubling occurs in Italian or not is a contentious issue, the debate about which I do not wish to enter into.

restructuring and non-restructuring verbs, making certain constructions available to some verbs but not to others. In the following we will investigate these points.

5.2 Clitics As Words

What independent arguments may be put forward in favour of the treatment of the clitics of Italian as words? Firstly, given that the kind of clitics we are interested in are generally assumed to have developed historically from simple clitics, it seems reasonable to assume that at least at some stage these clitics must have been full words. It is then a question of deciding at which stage of development the clitics have reached in Italian. Obviously this must always be a language specific question. What we must do then is survey the evidence.

The clitics of Italian do exhibit some properties of affixes. For example it is generally assumed that morphemes which do not bear an independent accent are affixes, and as such clitics of Italian appear to be affixes. However, as we saw in the previous chapter, this generalisation does not hold for all the clitic pronouns of Italian, i.e. there are cases of pronouns in the clitic paradigm with independent stress. The third person plural pronoun *loro* does appear to have its own stress. Given that we are trying to evaluate at what stage of development the clitic system of Italian is, *loro* seems like evidence that the transformation of the system as a whole from words to affixes has not been quite completed.

As for the other clitic pronouns, while it is true that when they appear on their own they have no independent stress, it is not quite so true in clusters such as *glielo*, which would be quite difficult to articulate with no independent stress. This then could be seen as further evidence that the clitics have not been totally incorporated into the inflectional system and are hence, words. The most that can be said is that Italian is in the process of changing. But as historical change is never entirely predictable it is difficult to know with any surety which way Italian will eventually develop.

Another argument for analysing clitics as affixes is that it is assumed that parts of words, such as affixes, do not undergo rules of deletion under identity but that entire

words may:

- (2) a capisco Maria e credo Maria
b capisco e credo Maria
know-Pres-1st-sing and believe-Pres-1st-sing Maria
“I know and believe Maria”
c la credo e la capisco
her-clitic know-1st-sing and her-clitic believe-1st-sing
“I know and believe her”
d *la credo e capisco

Here we see that the clitic is behaving as if it were not an independent word. As I currently have no account of coordination in Italian it is difficult to know with any precision how or if to handle this phenomenon syntactically. I will suggest that it might have something to do with the fact that clitics have to be immediately adjacent to the head on which they are cliticised. On some reasonable account of co-ordination we would want to say that *la* has the same relationship to both verbs such that if it is a clitic of *credo* it must be a clitic of *capisco*. In (2d) however we see that *la* is separated from *capisco*, i.e. is not immediately adjacent to (one of) the heads on which it is cliticised. Alternatively, it could be claimed that these examples may be ruled out for phonological reasons. The point is that we do not seem obliged to treat clitics as affixes.

Much has also been made of the fact that pronominal clitics are more syntactically restricted than other full nouns; for example, they cannot be the head of a relative clause. However, against this it can be said that other non-clitic pronouns are syntactically constrained in similar ways. This does not mean to say that these pronouns are inflectional affixes. I would claim that in both cases all that needs to be done is to sub-divide the class of nouns into as many sub-classes as is required; those that take adjectives and those that don't; those that can be the head of relative clauses and those that can't. In short, I think there are no reasonably convincing reasons for taking the clitics of Italian to be affixes.

As is also well known, clitics do exhibit some properties of words. For example, unlike affixes, clitics in Italian are relatively syntactically free. As has been mentioned they can occur in various positions relative to the verb that subcategorises for them, depending on tense and mood. In addition, under the conditions described in section 4.3.2, they can

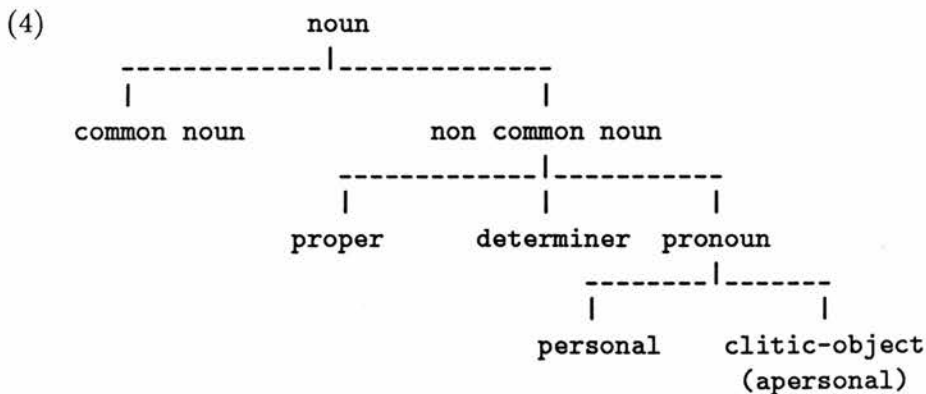
move up from a lower clause to a higher clause.

Most significantly of all, clitics may satisfy the subcategorisation requirements of verbs. If we take subcategorisation, for Italian at least if not cross-linguistically, to be a phenomenon restricted to words, this would strongly suggest that clitics of Italian are words.

- (3) a *io amo
 I love
 b io lo + amo
 I it/him-clitic + love

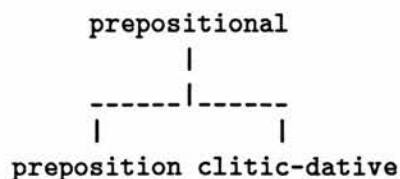
Finally on this point, I would like to mention that many of the problems caused by clitics and their syntactic/morphological classification are based on the assumption that syntax and morphology are completely separate and that there exist no borderline cases. This is not the outlook adopted by the current framework.

Given that we want to capture the fact that clitics in Italian are words, we have to decide what syntactic category to assign them and how to incorporate this fact into the grammar. The position taken here is to assign clitics the same category as the argument they replace. This appears correct for the object clitics which we may classify as nouns, but forces us to say that indirect object clitics are prepositions or at least prepositional (bearing in mind that we have no category prepositional phrase). This is slightly unsatisfactory but not problematic. What we then have to do is construct the relevant taxonomies that incorporate the node *clitic* at some point. The taxonomy used for nouns incorporating object clitics is the following:



The taxonomy for prepositions and hence dative clitics is the following:

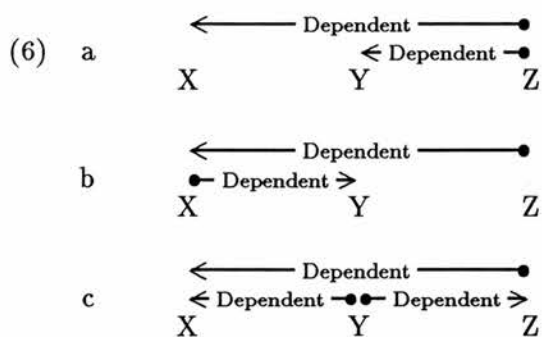
(5)



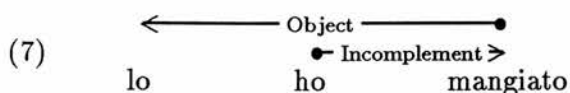
I have called the clitic object pronouns *apersonal* because whereas their strong counterparts may only refer to people, the clitics may be used for people, animals and things. Other taxonomies may be imagined such as a binary branching one but nothing empirical seems to hang upon it at this stage. The point is that some classification of the object clitics as nouns is possible. As these taxonomies lead up to the entity word we have capture the fact that clitics are to be analysed as words.

5.3 Problems Inherent In The Framework

Clitics in Italian cause some problems that are specific to the Word (Unification) Grammar framework. On the assumption that clitics differ minimally from other nouns, certain of the positions in which they are licensed violate the Adjacency Principle. Recall that the Adjacency Principle licenses the following structures:



If we now look at the structure assigned to a sentence containing a clitic in Italian such as *la voglio mangiare* (literally, it I-want to-eat) we have the following structure.



Though this sentence is actually grammatical, by violating the Adjacency Principle it is excluded by the grammar. These kinds of problems would not necessarily arise in frameworks where the Adjacency Principle does not play such a central role. What we need then is an analysis of clitics that not only accounts for their distribution and ordering, which any account must do, but also does so by respecting the Adjacency Principle.

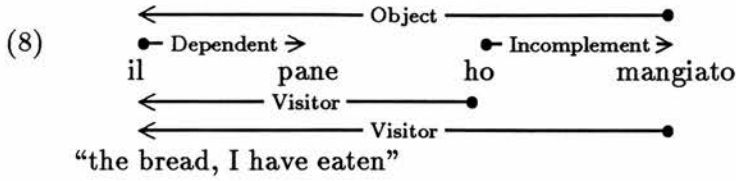
The alternative would be to relax our statement of the Adjacency Principle. But to allow in structures such as that assigned to (7) *la voglio mangiare*, would be to also allow in sentences such as *big the cat* (cf. example (17) Chapter 2) which have the same structure in Word (Unification) Grammar. Clearly, we would not want to do that. A second alternative would be to try and formulate a definition of the Adjacency Principle that is specific to Italian, but this too would be unsatisfactory. The Adjacency Principle is not the kind of thing that one would expect to be language specific. The final possibility is that the structure given above in (3) must, in some sense, be incomplete. That is to say that there must be other dependency lines that need to be included which have been left out. This will be borne out by what follows.

5.4 The Analysis And Possible Solutions

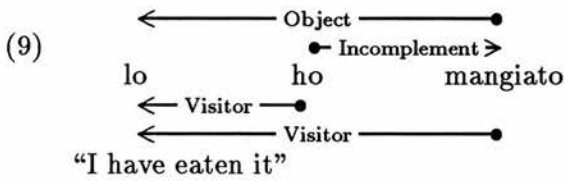
5.4.1 A Visitor Account

Two possible ways around the Adjacency Principle violation presented by clitics suggest themselves. A first alternative would be to make use of visitors which we recall is the apparatus used in Word Grammar to handle many kinds of “movement”. Recall from Chapter 2 that the idea behind the grammatical relation of visitor is that of a “moved” element which in looking for its head *visits* various possible candidates before finding the correct one. There are many advantages to using visitors. Firstly, as visitors are already part of the theoretical framework if a solution could be found that made use of them, it would be preferable to a solution that required the addition of extra devices to the framework. This would be especially so considering that the parallel of (7) above that contained an non-clitic object as in (8) would obligatorily require the use of visitors. This

is because although we have said that objects are free in Italian they are only free to the extent that they obey the Adjacency Principle. It is fairly clear in (8) that the dependency link between *il* and *mangiato* violates the Adjacency Principle without the visitor link from *ho* to *il*.



The use of visitors for (7) above would then give us the following structure which now no longer violates the Adjacency Principle:



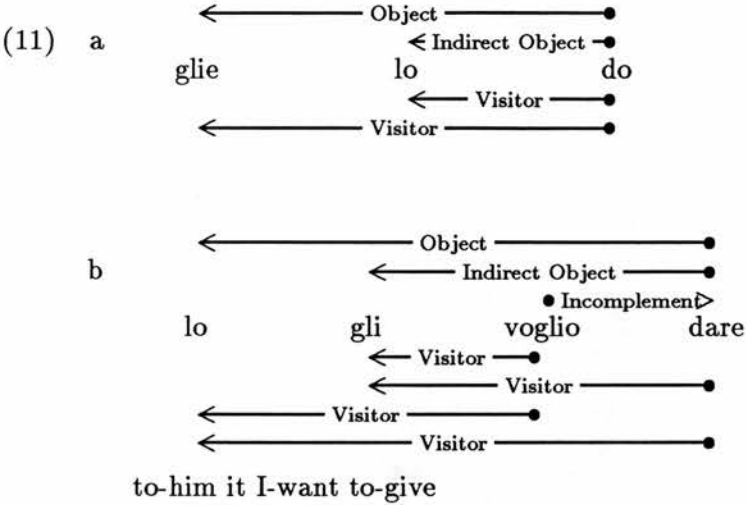
Assuming that an account of clitic movement using visitors were adopted what would the consequences of this be? It could be argued that a point in favour of the visitor analysis would be that it is compatible with the *hopping* nature of Clitic Climbing. This might make it easier to state the requirement that clitics may only hop from one restructuring verb to another, and would therefore make it easier to state how far up clitics could climb. The use of visitors might also be helpful in accounting for Object Agreement in Italian, for as we can see from (9) the clitic is a visitor and hence a dependent of the participle.

The use of visitors is not without its problems though. Any account in terms of visitors would have to explain two things. Firstly, given the parallels between (8) and (9) above, why in the case of the non-clitic noun is there no Object Agreement, bearing in mind that the relation of the participle to both nouns is the same?

- (10) a la torta ho voluto mangiare
 b *la torta ho voluta mangiare
 c la ho voluta mangiare
 d *la ho voluto mangiare

Secondly and more significantly, visitor links have only previously been used for unbounded dependency. Again given the parallels in structure between (8) and (9) we would need to

account for the fact that clitic movement is bounded within a tensed clause but that other kinds of movement are not. In other words it would require making a distinction in the visitor links between those used for clitic movement and those used for other kinds of movement. Additionally, the above account based on visitors also tells us nothing about why the clitics are so rigidly ordered in clusters, and the Adjacency Principle is no help in deciding such questions as we may observe below:



The point here is that the analysis based on visitors would have no immediate explanation for why when clitics move they are rigidly ordered with respect to each other unlike other moved elements. Given the other types of movement covered by visitors, this is not the behaviour we would expect. That is to say it is not the case that when two or more entities are dependents of the same word, that they are necessarily rigidly order with respect to one another.

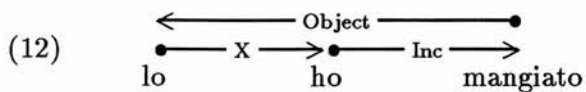
Were we to pursue the visitor approach then, we would have to find a way around the problem of ordering. The need for extra devices to account for the strict ordering is an indication that the mere appeal to the visitor link, in and of itself, is not going to be enough for a convincing account of the distributional phenomena of cliticisation. One such way of pursuing this analysis that would help overcome the ordering problem and the problem with Object Agreement mentioned above would be to say that there are two types of visitor link. The usual one for unbounded movement and a second one for bounded movement that also induces Object Agreement. The second, bounded, one would be used for clitics

and we would have a statement somewhere else in the grammar concerning relative linear ordering of clitics.

Another way of pursuing this analysis would be by making a distinction between full nouns and clitic nouns. Neither of the above approaches seem palatable. The first approach uses an ad hoc device, whereas the second creates a distinction for which there seems little independent justification. Further the appeal to the visitor approach was made in the hope of making clitic movement fall under more general kinds of movement. If adopting the visitor approach involves making all kinds of distinctions that were not previously required then this seems like strong evidence that the phenomena we are concerned with do not actually fall under more general kinds of movement. It thus seems as though this line of thought is not leading us anywhere and that a solution that does not make reference to visitors will have to be found. Interestingly though, this rejection of an account in terms of visitors is in line with the Government and Binding accounts given in section 5.1, which also reject accounts of clitic movement in terms of the same apparatus used for other kinds of movement and opts for a base generation account.

5.4.2 A Non-Visitor Account

An account that is not based on the use of visitor links would bring the alternative analyses into line with the intuition behind the Government and Binding approach that treats clitic movement differently than other kinds of movement. Given that we have relinquished accounts in terms of visitors there are just two alternatives left for making the structure in (7) compatible with the Adjacency Principle. One possibility would be to make the clitic the head of the verb on which it is cliticised. The resulting structure for (7) would be (12), where *X* is the link we are investigating:



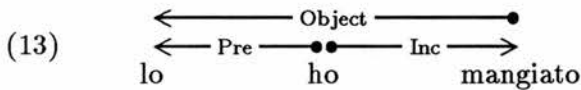
it I-have eaten

The problem with doing this is that although (12) doesn't violate the Adjacency Principle, the result is that there is now no root in the sentence. This goes against two traditions in

dependency theory. Firstly that a sentence should have a unique root and secondly that the root of such sentences is a tensed verb. This solution will therefore be rejected without further investigation. The alternative possibility for making the example in (7) adjacent will be pursued in the next section.

5.5 Bicephality or Double Headedness

A second way of making the clitic in the structure of (7) above respect the Adjacency Principle is to take the position that the clitic is a dependent of the auxiliary despite the fact that the auxiliary apparently does not subcategorise for it. In certain obvious ways given the nature of cliticisation, the clitic is dependent on the word to which it is cliticised. This is the approach I will take and I will outline it briefly below. Making the clitic dependent on *ho* in (7) would give us the following structure, with the new dependency arc we are investigating labelled *pre* for pre-dependent:



Though we no longer have an Adjacency Principle violation, how can we justify this analysis? The intuition is that cliticisation is a kind of subordination and therefore dependence of one word on another. So on this level of argumentation the link seems justified. It is also true to say that typically dependents take their position from their heads and again this is surely the case with the clitic and *ho*.

There is a problematic consequence of this approach. The clitic pronoun is obviously still an argument and therefore a dependent of *mangiare* in (13) so the object link is indispensable. But we are now claiming that the *pre(dependent)* link with the auxiliary is also indispensable, for without it the sentence will incorrectly be excluded by the grammar as it violates the Adjacency Principle. The effect of this is that we have now given the clitic a second (obligatory) head, namely *ho*. As such we must include in the grammar the fact that certain verbs such as *ho* subcategorise for clitics.

We have seen that individually each head can be motivated but we have still not addressed the fact that the clitic has two (obligatory) heads. Having two heads is not unique in Word Grammar structures, although it is not a feature of traditional dependency theory. In fact, as mentioned elsewhere, subject nouns in Control structures can have two or more heads. But there we are dealing with something slightly different. In Control structures it is not the subject noun that obligatorily requires the two heads subject nouns in English are perfectly grammatical (in other constructions) with only one head. Put another way, subjects nouns, like all other words only *obligatorily* require one head i.e. they are monocephalous, though they allow for the possibility of more. Clitics on the other hand appear to be obligatorily *bicephalous*, i.e. obligatorily require two heads.

It would not be possible to maintain that clitics only required one head and to construct a solution along the lines of that associated with Control sentences so that they end up with two heads. This is because if they only obligatorily required one head then we would expect them to occur in all the same places as their non-clitic counterparts. Given that they cannot, it must be the case that there are extra constraints on them; the principal one being bicephality.

Though unique in the respect of obligatorily requiring more than one head, it should not be all that surprising. Intuitively, by the very nature of clitics we know that they are subordinate and dependent upon the words on which they are cliticised. It is also fairly clear that a clitic must be a dependent of any word that it is an argument of. It could be argued that one of the links is syntactic and the other *clitic* link is morphological. To develop this suggestion further would take us too far afield from the central aims of this thesis and so it will not be pursued. It is only because in Italian there are structures in which the clitic host is distinct from the word of which the clitic is an argument, that we get this insight into the nature of clitics. In the next section I shall show the consequences of adopting a bicephalous approach and how the relevant facts may be encoded into Word Unification Grammar.

5.6 Bicephality in Word Unification Grammar

Before we proceed, rather than present a complete grammar of Italian we will assume that the grammar of Italian is constructed along similar lines to that of the grammar of English presented in Chapter 3. What that means is that there will be a taxonomy for tense, for subcategorisation and so on. The only significant differences would be the following: a) subjects are optional b) the obvious lexical differences, c) an extra feature for gender in the lexical entries, d) an extra feature for restructuring. We will assume, at this point, that there are no NEEDS statements other than those presented below. Other differences will be presented as we proceed through the account.

The line of thought pursued in the previous section has led to the following conclusions: Subjects are optional in Italian. This may be captured by adding the following proposition to the grammar:

- (14) [Root, [verb, intransitive, tensed, R], Agr] NEEDS
[op, subject, [Root2, [non-common-noun|B], [P, N, nominative]]

What this says is that any tensed, intransitive, verb may optionally take a subject that is below the node non common noun given in (4) above. This non common noun must also have nominative case. Unlike English a verb does not state which side its subject must appear and as such we say that subjects are free(-dependents). The *R* is a variable that might be instantiated to either *r*(estructuring) or *non-r*(estructuring). Other nominal dependents are also free though not optional:

- (15) i) [Root, [verb, transitive, Fin, R], Agr] NEEDS
[ob, object, [Root2, [non-common-noun | B], [P, N, objective]]
ii) [Root, [verb, ditransitive, Fin, R], Agr] NEEDS
[op, object, [Root2, [prepositional | X], Agr2]

Of note in (ii) is that the (indirect) object must be a prepositional type word and not simply a preposition. This allows for both the clitic and the non-clitic cases mentioned above. Clitics are marked in the grammar. This is reflected in the fact that they are the only entity that is obligatorily bicephalous, requiring TWO heads. We can capture this by adding the following statements to our grammar:

- (16) i) [Root, [word|X], Agr] NEEDS [ob, head]
- ii) [Root, [clitic|X], Agr] NEEDS [ob, head]

Within the context of the system described in Chapter 3 the statements in (16) will result in clitics obligatorily requiring two heads. It might be thought that as (16ii) repeats (16i) but at a more specific level that we would only want to inherit the lowest fact namely (16ii). But as there is no conflict between the statements in (16) within the terms of Chapter 3, there will be no overriding of propositions and hence the simple or accumulative form of inheritance will operate and the desired result of clitics requiring two heads will be achieved. It is also the case that all verbs that require nominal dependents in addition to the subject (i.e. are (di)transitive) have the following properties:

- (17) i) When finite they subcategorise for an optional clitic pronoun as a pre-dependent:
[Root, [verb, intransitive, finite, R], Agr] NEEDS
[op, pre-dependent, [Root2, [clitic | X], Agr2]
- ii) When in the infinitive they subcategorise for an optional clitic pronoun as a post-dependent
[Root, [verb, intransitive, infinite, R], Agr] NEEDS
[op, post-dependent, [Root2, [clitic | X], Agr2]
- iii) Additionally, some verbs that take verbal complements though they do not take nominal complements also subcategorise for an optional clitic pronoun in the same way as those given in i) and ii). These we call the restructuring verbs, represented by the feature *r*
[Root, [verb, Subcat, finite, r], Agr] NEEDS
[op, pre-dependent, [Root2, [clitic | X], Agr2]
[Root, [verb, Subcat, infinite, r], Agr] NEEDS
[op, post-dependent, [Root2, [clitic | X], Agr2]

In both (i) and (ii) the node referred to for subcategorisation is intransitive. It could be thought that this statement might allow intransitives to have (object or dative) clitics, which clearly they can not. This statement will not overgenerate as in such cases the verb will not have the second dependency link to license the clitic. We shall look at this more closely below.

If we made all verbs that take either an object or an indirect object (i.e. all verbs below the intransitive node) restructuring verbs as well, then (i) and (ii) above would reduce to (iii). What this means is that the simple facts of cliticisation in Italian can essentially be handled simply by two statements about subcategorisation and one involving an extra

No other order is permitted⁴ as the clitic may not receive two heads in other position except pre-verbally. So far at least the account has empirical coverage and achieves descriptive adequacy.

Given that the statements in (17) do not make reference to the type of clitic required, what is to stop transitive verbs taking a dative clitic? Dative clitics will have dative case assigned in the lexicon. This will not prevent them from being clitics of transitive verbs but it will prevent them from being direct objects which will be assigned accusative case. Therefore the dative clitic will not receive two heads in the transitive case which will correctly be predicted as ungrammatical. Let us now move to a slightly more complicated case, namely one in which a clitic is enclitic on an infinitival. We will take the following words:

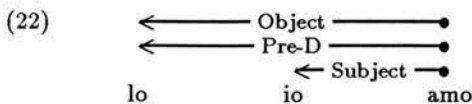
(23) lo, amo, mangiare

The relevant subcategorisational propositions inherited from the grammar are as follows:

- (24) a [lo, [clitic, def, Count], [3, sing, masc, acc]] NEEDS [ob, head],
 [lo, [clitic, def, Count], [3, sing, masc, acc]] NEEDS [ob, head]
 b [amo, [verb, indicative-present, infinitive-comp, non-r], [1, sing, nil, nil]] NEEDS
 [op, post-dependent, [Root, [verb, infinitive, Subcat, R], Agr]
 c [mangiare, [verb, infinitive, trans, r], Agr] NEEDS
 [op, post-dependent, [Root, [clitic | X], Agr2]
 [mangiare, [verb, infinitive, trans, r], Agr] NEEDS
 [op, object, [Root, [non-common-noun | Y], Agr]
 [mangiare, [verb, infinitive, trans, r], Agr] NEEDS [ob, head]

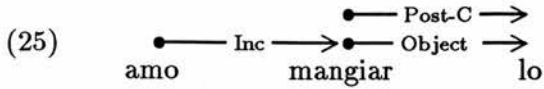
Here *amo* does not take a non-subject nominal complement (it takes an infinitival complement) and is not a restructuring verb and hence does not take an optional clitic pre-

⁴The following ordering is not excluded by the grammar but is nevertheless ungrammatical

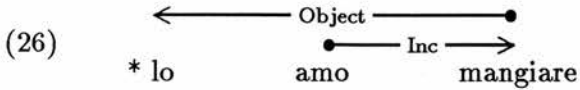


To make the correct predictions we have to stipulate that the pre-dependent link between the verb and the clitic requires immediate adjacency. On the assumption that the clitic link is a morphological like link this does not seem unreasonable.

dependent. The only way these propositions may be satisfied so that the clitic gets two heads is when the clitic is after the verb as in (25):



The order in (26) is excluded as the clitic can not get two heads assigned to it:



Thus accepting the argumentation presented above we have now succeeded in accounting for the non-exceptional behaviour of clitics.

5.6.2 Clitic Climbing

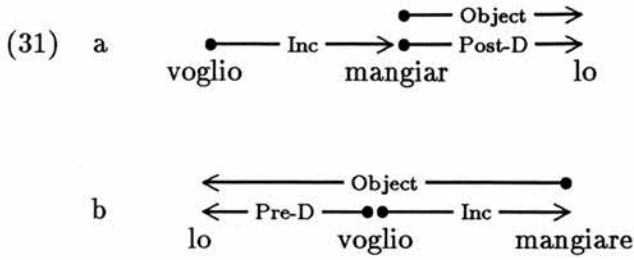
From the previous section it should be quite apparent how this solution may also be motivated to capture the simple facts of Clitic Climbing. The simple cases of Clitic Climbing are analogous to those in (25) and (26) with the exception that the matrix verb, in virtue of being a restructuring verb optionally subcategorises for a pre-dependent. Let us look at the example in (25) with a restructuring verb substituted for *amo* as in (27) below:

- (27) a voglio mangiar lo
 b lo voglio mangiare

The relevant subcategorisational propositions inherited from the grammar are as follows:

- (28) [lo, [clitic, def, Count], [3, sing, masc, acc]] NEEDS [ob, head],
 [lo, [clitic, def, Count], [3, sing, masc, acc]] NEEDS [ob, head],
- (29) [voglio, [verb, indicative-present, inf-comp, r], [1, sing, nil, nil]] NEEDS
 [op, post-dependent, [Root, [verb, infinitive, Subcat, R], Agr]
 [voglio, [verb, indicative-present, inf-comp, r], [1, sing, nil, nil]] NEEDS
 [op, pre-dependent, [Root, [clitic | X], Agr]
- (30) [mangiare, [verb, infinitive, trans, r], Agr] NEEDS
 [op, post-dependent, [Root, [clitic | X], Agr2]
 [mangiare, [verb, infinitive, trans, r], Agr] NEEDS
 [op, object, [Root, [non-common-noun | Y], Agr]
 [mangiare, [verb, infinitive, trans, r], Agr] NEEDS [ob, head]

Given the statements in (28) to (30) the only possible orderings are precisely those ones that we observe in (27) as can be seen from (31) below:



It might be thought that by having the option of a predependent and a post dependent simultaneously that the statements in (28) to (30) would overgenerate. In the examples in (31) it is not possible to have two clitics present simultaneously one at each end as the second clitic could only receive one head. Nor is it possible to have a non-clitic noun use up the extra optional dependency link as the restriction on these extra links is that they may only be used for clitic pronoun.

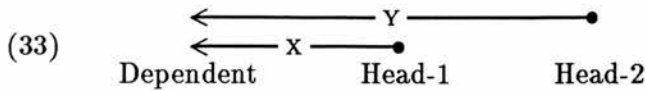
The account seems to have considerable empirical support. Before we go on to extend the coverage and include amongst other things cases of Object Agreement I would like to look closer at the theoretical implications and plausibility of this analysis by looking more generally into the nature of double dependencies, bicephality being a specific subcase.

5.6.3 Double Dependencies

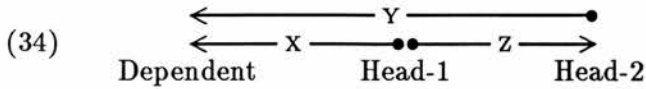
In what follows it must be remembered that the constructions Hudson is considering were cases of Control which involve modifier sharing not bicephality (the difference being one of obligatoriness). It should be expected however that many of the properties will be shared. Hudson (1986g) makes the following points concerning heads:

- (32) a “A word may have more than one head
 b A word may be separated from one of its heads by the other ... provided that the head furthest away from it depends on the other head.
 c A word which has two heads need have no semantic relation to the higher of them; i.e. a word takes some or all of its semantic relations from its lowest head
 d A word takes its position from its highest head.

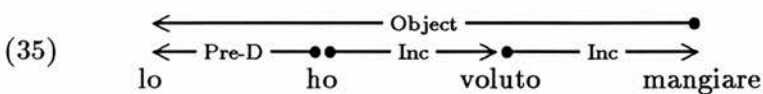
Pictorially (32b) says that in the structure:



Head-2 must also be a dependent of Head-1, i.e. the structure given above must include the additional arc labelled *Z*



This is the familiar situation we have for sentences such as *John wants to go*, where in (34), *John* would replace *Dependent*, *wants* would replace *Head-1* and *to* would replace *Head-2*⁵. The last two points in (32) say that if a word has a number of heads then the dependency lines are justified on a continuum that goes from mainly syntactic for the highest to mainly semantic for the lowest. This is important for the analysis I have proposed for clitics in Italian which involves obligatorily requiring all clitic pronouns having two heads. The relevant case is obviously where the clitic does not get both its heads from the same element, i.e. in Clitic Climbing. (In the non Clitic Climbing case there is obviously no conflict). Thus we have an example like the following:



“I wanted to eat it”

This looks promising; the clitic pronoun *lo* has two heads and in accordance with (32b) above, the head it is separated from is dependent (strictly speaking subordinate) on the head that the clitic is next to. Following (32c) and (32d) above we note as well that it takes its position from its highest head and so this relation is syntactic. *lo* also has a second head and is predominantly semantically (in the sense of being an argument) that it is related to this second, lower head. Thus the generalisations seem to be borne out.

⁵c.f. example (19) of Chapter 2

5.6.4 The Relative Ordering Of Clitics

Things unfortunately are not so simple. The problems with the solution only becomes apparent when trying to capture the rigid ordering of clitics in clusters. As we saw in Chapter 4, in Italian, the object clitics, *lo* or *la* for example always follows the indirect object clitics such as *gli*. Thus we have the following:

(36) *glie* + *lo* + *do*
 to-him it I-give

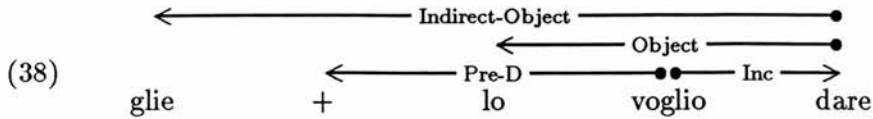
(37) *glie* + *lo* + *voglio* *dare*
 to-him it * I-want to-give

The first problem is how to assign the clitics the two heads they both require. The first head, they each individually get from being arguments of *dare*. From (36) we might be tempted to change our simple statement about tensed restructuring verbs optionally requiring a clitic pre-dependent to say that transitive verbs take an optional object clitic and that ditransitive verbs take both an optional object clitic and an optional indirect object clitic. This would mean that we could give each clitic two heads though we would still need some statement about clitic ordering.

However (37) presents further problems as the clitics are cliticised on to a verb that does not subcategorise for them. To handle this case along the lines just suggested, we would need additional apparatus that would percolate dependents of dependents (and subordinates) up to the matrix verb in order to know what kinds of clitics the matrix verb could take. Alternatively we could say that the restructuring verbs optionally take both an object and a dative clitic, and rely on the fact that any clitic not required would be ruled out on the basis that it would not get its required two heads much in line with the parallel case discussed at the end of section 5.6.2. Such a solution would not be satisfactory because there are many kinds of clitic in Italian that we have not looked at each of which would require of the restructuring verbs a stipulated link of its own. This kind of solution then is to be rejected.

Returning to examples (36) and (37) if we want to maintain the account of clitic placement given in the previous subsection and for it to be truly general, we are compelled

to make at least one of the clitics in (37) dependent on *voglio*. Given this how can the correct grammatical structure be assigned to that sentence? Well the structure assigned to (37) will be along the following lines:



As far as it goes, this analysis seems acceptable but we must bear in mind that we are looking for a situation in which both clitics have two heads. One way of doing this would be to make the *pre-clitic* link (which has been ambiguously placed between the two clitics for the meantime) be *absorbed* in some sense by the cluster. The result would be that both clitics have a head each and one shared head, giving them each both the heads they require.

This is not a very appealing solution. Either, the cluster is one clitic unit or it is two. If it is one unit and the cluster absorbs the pre-dependent arc then the cluster as a whole would have three heads. Though this would solve the ordering problem it creates more problems. Even if we could square the fact that the cluster has three heads with (16) above it would be difficult for the cluster not to additionally absorb both the object and indirect object dependency link without devices to prevent this and hence to be both accusative and dative. This might be further complicated in the case when the two elements of the cluster are arguments of different verbs and as a result we needed to separate the information in the cluster.

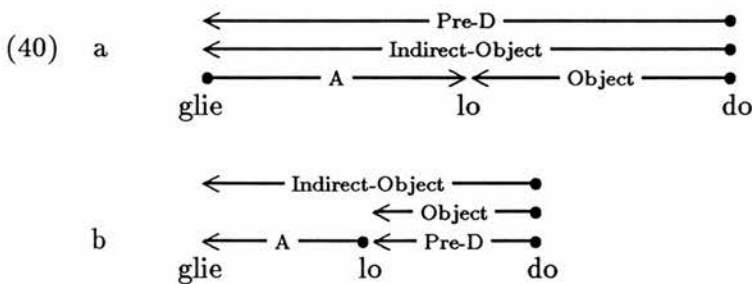
If, on the other hand, there are two units, then restricting the ordering remains problematic and some special device would be required to satisfy the head requirements, given that there are only three heads available. One possible solution would be to make the *pre-clitic* arc go to one of the clitics and then have the second clitic dependent on the first. Both would then each get the two heads they require and there would be no *absorbing* into the cluster. If this could be achieved it would be a far preferable analysis.

This analysis is a little controversial due to the fact that one is not accustomed to having objects depend on indirect objects or vice-versa. The position then needs some

defending. Two arguments come to mind. Firstly the relative ordering of pronominal clitics in Italian is strictly governed and cannot be changed. This is not true of the full nouns, and such ordering is a typical consequence of dependency. Secondly, in sentences with more than one clitic within a tensed clause, we find that the clitics must form a cluster, and that the cluster itself can climb but it can not be split. Thus we find the following paradigm:

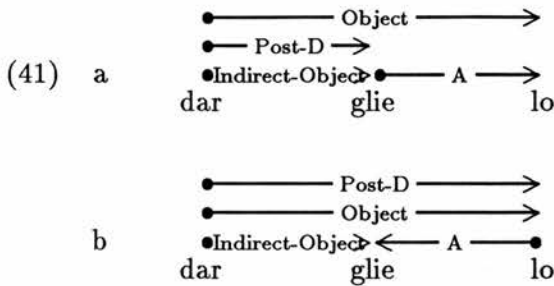
- (39) a glie + lo + voglio fare mangiare
 (to)-him it I-want to-make (to)-eat
 b voglio far + glie + lo mangiare
 I-want to-make (to)-him it (to)-eat
 c *voglio far + gli mangiar + lo
 I-want to-make (to)-him (to)-eat it
 "I want to make him eat it"

Clusters must be formed even where, as in (39c) the clitics are actually in the position you might normally expect them. Again this might suggest a certain dependence within the cluster (though it is not conclusive evidence). Setting dependency structures up in the correct way might ensure this. The question is which way round? To help us decide this lets look at the options for the simple sentence in (40). The link that we are currently investigating shall be labelled A.

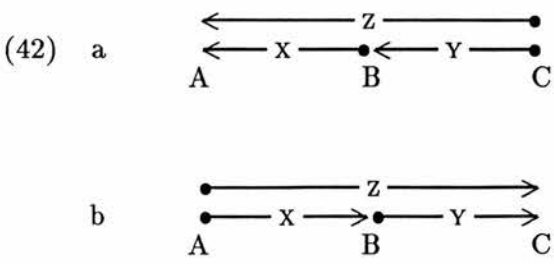


One might prefer analysis (40b) where the indirect object is dependent on the object as it, in a sense, makes the transitive case more basic than the ditransitive case. This would be achieved by saying that the object clitic subcategorises for an optional indirect object clitic. However an equal case could be made for (40a) on the grounds that the dative clitic is prepositional and it is prepositions that usually take nouns rather than the other way round.

How are we to decide between these two? This brings us back to the discussion on double dependencies in section 5.6.3. Remembering the criteria for words with two heads, we notice that (40b) violates condition (32b) of the constraints on double dependencies; that is to say *gli* is separated from one of its heads, namely *do* by its other head *lo* but the head that is furthest away from it, *do*, is not dependent on the head that it is not separated from, namely *lo*. In fact the reverse is the case, *lo* is dependent on *do*. The second alternative, (40a) violates none of the conditions. That is until we take a closer look. Remembering that clitics can also be enclitic on infinitivals in Italian we have the following enclitic counterparts to (40):

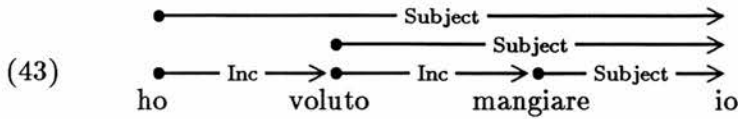


This time it is in (41a) that we have a violation of (32b). But in (41b) the option that was excluded in (40) we do not. This is unfortunate as the criteria have apparently ruled both out, putting the analysis into question. If we abstract away for a moment, and take a closer look at condition (32b) we see that what it is actually trying to achieve is to ensure there is no Adjacency Principle violation between a dependent and all of its heads. This became necessary when the notion in dependency theory that a word must only have one head was dropped by Word Grammar. However, if we look closely at its actual wording we see that it forbids structures of the following kind:



which as can be seen are mirror images of each other. Now this is quite strange to single out this structure for exclusion, especially as the Adjacency Principle is not violated.

Also, these are precisely the structures you would expect to find in languages that are prototypically used by dependency grammarians to motivate their theories, i.e. languages that are consistently head last such as Japanese, (42a) above, and languages that are consistently head first such as Welsh, (42b) above. Even in Italian, which allows the subject slightly more freedom than English we can find such structures:



It looks then as though condition (32b) above is stated on the basis of English alone and so is of little general use. This again leaves us with a choice of (40a) or (40b) as (32b) has been rejected. However, both (41a) and (41b) seem to violate (32c) that a word with two heads takes most of its semantics from its lowest head. This notion of *lowest head* again seems to be based on English alone. Given the structures in (42) a preferable statement of (32c) would refer to the notion of *furthest head*. This would be justified on the basis that *lowest head* refers specifically to head-first languages, whereas *furthest head* would be the more general notion. Having said this the notion of *furthest head* would still need formalising so that in examples like (40a) and (41b) we could make the relevant distinctions. Given the modifications just discussed both analyses presented are no longer excluded. The question of which clitic depends on which remains unresolved but the analysis is no longer too problematic. The point is that some decision can be made but we don't have enough evidence at the moment to decide.

In order to find more evidence, we need to change tack slightly. Hudson (1987a) presents counter arguments to Zwicky (1985b) for the notion of *head* and produces the following table to which I have added a column for Cl(itic) and V(erb), and gli and lo to represent dative and object clitics:

(44) .

	V+NP	P+NP	NP+VP	Det+N	Aux+VP	Comp+S	Cl+V	gli+lo
semantic functor	V	P	VP	Det	Aux	Comp	V	-
morphosyntactic locus	=	=	=	=	=	=	=	gli
subcategorisand	=	=	?	=	?	=	=	?
governor	=	=	=		=		(=)	?
distributional equivalent	=			=	=	=	=	-
obligatory constituent	=	=	=	=	=	=	=	-
ruler	=			=	=	=	=	-

"=" means "same as semantic functor"

If we bear in mind what was said about the form of clitics in clusters and take into account the second row of the table given above we finally have some foundation for deciding which analysis to choose in (40) and (41) above. In clusters it is always the dative clitic that changes its form. That it does not do so in isolation demonstrates that it is not the effect of the verb to which it is cliticised. If the dative clitic is the morphosyntactic locus in the cluster than it is probably, according to the table above, the head. As a result we might favour the analyses given in (40a) and (41a) above. It would also bring dative clitics in line with other prepositionals as mentioned above.

It must be remembered, though, that the table is merely a guide to tendencies and does not express necessary and sufficient conditions. So we can only be suggestive. If we have no strong grounds for choosing one way or another, and either choice is possible then the choice we make is essentially irrelevant.

5.6.5 Clitic Climbing Revisited

Let us return to the problem of ensuring that a clitic may only move up chains of restructuring verbs. The first thing we want to ensure is that, irrespective of where it originates, when a clitic is attached to a verb, that verb is a restructuring verb. This may be achieved by simply adding a restructuring feature to our entries for clitics and stipulating that the clitic's feature for restructuring has to unify with the restructuring feature on verbs. Inco-

porating the restructuring feature into the entries for clitics is quite simple and is shown below⁶:

(45) [Root, [clitic, Def, Count, +r], Agr]

We could then change the subcategorisation statements for restructuring verbs along the following lines:

(46) [Root, [verb, intransitive, tensed, R], Agr] NEEDS
[op, pre-d, [Root2, [clitic, Def, Count, R], Agr2]]

This basically says that any tensed verb may take a clitic as a predependent which shares its restructuring feature. This might seem too general but given that there are no -r clitics, unification failure will prevent non-restructuring verbs from taking a clitic as a dependent. This ensures that any verb that a clitic attaches itself to must be a restructuring verb. However, we have not yet shown how a clitic is prevented from climbing past a restructuring verb.

In order to achieve this, we must distinguish three kinds of restructuring verbs. Those are (i) main verbs that take nominal complements other than subjects (ii) main verbs that take verbal complements and (iii) auxiliaries (which also take verbal complements). Obviously the verbs that concern us are the ones that take verbal complements, namely (ii) and (iii) as it is only these that may form chains of verbs.

There are two ways we could prevent clitics hopping over restructuring verbs. Given that we want restructuring verbs to be able to combine with non-restructuring verbs where there are no clitics as in (47)

(47) ho odiato mangiare la torta
+r -r +r

we can not simply enforce agreement of restructuring features along strings of verbs as this would incorrectly rule out (47). Further if we can not enforce such an agreement rule, it will be difficult to say that the restructuring verbs of types (ii) and (iii) unlike those

⁶in the following -r = non-restructuring, +r = restructuring and R is a variable over +/- r and may be taken to mean unspecified for the feature restructuring

taking nominal complements, have the feature $+r$ once and for all. This is because in the example in (47) with *la torta* replaced by a clitic, we would have no means of preventing that clitic from being attached to the auxiliary *ho* given (46) above.

The first way around this involves having two entries in the lexicon for every restructuring verb, one with the feature $+r$ one with the feature $-r$ and then enforcing agreement of restructuring features along strings of verbs. There is a second approach though and one that does not involve the reduplication of lexical entries. The alternative approach says that while verbs that take nominal complements (excluding the subject) are specified as $+r$, the restructuring verbs that take verbal complements are actually unspecified for the feature restructuring, i.e. have the feature R , and get their restructuring feature instead from one of their dependents. This may be achieved as shown in (48):

- (48) [Root, [verb, infinitive-comp, Tense, R], Agr] NEEDS
 [ob, infinitive-comp, [Root2, [verb, Subcat, infinitive, R], Agr2]]

This says that any tensed verb that subcategorises for an infinitive agrees with that infinitival in the feature for restructuring. The term restructuring verb will now be used to refer to verbs that have the feature $+r$ or R . However, bearing in mind that some main verbs that take verbal complements are non-restructuring, but for all other purposes are identical to their restructuring counterparts, the statement in (48) will wrongly exclude structures like the following:

- (49) odio mangiare la torta
 -r +r

The example in (49) is excluded because the matrix verb may only combine with words that share its restructuring feature as stated in (48). Obviously (48) is too general. The solution is to override (48) with a more specific statement for non-restructuring verbs as in (50) below:

- (50) [Root, [verb, infinitive-comp, Tense, non-r], Agr] NEEDS
 [ob, infinitive-comp, [Root2, [verb, Subcat, infinitive, R], Agr2]]

Here there is no enforced agreement between restructuring features and so the example in (49) is no longer excluded. I have yet to demonstrate explicitly how this prevents clitics from hopping over non-restructuring verbs. The explanation is as follows. As we

may see from (50) non-restructuring verbs do not unify with the restructuring feature of their dependents. We also know from (46) that they may not accept clitics as dependents themselves. On the other hand, restructuring verbs that take verbal complements are unspecified for the feature restructuring and unify with their dependents feature. By not accepting clitics as dependents it is obviously the case that clitics may climb at least up to but not onto non-restructuring verbs. This is precisely what we want in cases of verbal strings containing restructuring and non-restructuring verbs.

Given that all restructuring verbs that take verbal complements are unspecified for restructuring but take their feature from their dependents we have three possible scenarios. Firstly, the verbal complement of such a restructuring verb is again a restructuring verb that takes a verbal complement. In this case the restructuring features will unify but will not be instantiated. In this case there has been no change and as such will fall under the following two cases. The second case is where a restructuring verb takes a non-restructuring verb as a complement. In this case, the feature for restructuring on the (unspecified) restructuring verb will unify and the matrix verb will take the feature -r. In such a way the matrix verb too will be prevented from taking a clitic dependent by virtue of being -r. This will be achieved despite the fact that the (unspecified) restructuring verb will subcategorise for an optional clitic, but once it has a non-restructuring dependent, that clitic will have to have the feature -r by (46). As was mentioned above, there are no such clitics.

The final case is where the verbal complement of the restructuring verb is a verb that takes a nominal complement and is hence assigned the feature +r. This is the same as the previous case except that here the matrix verb gets the feature +r under unification. As the only verbs to get assigned the feature +r are verbs that take nominal complements, and the only verbs that get assigned -r are those non-restructuring verbs that take verbal complements, the remainder being unspecified for restructuring, all cases have been covered. Let us look at some examples where we will concentrate on restructuring features only:

- (51) a odio mangiar lo
 -r +r +r
 b voglio mangiar lo
 R +r +r
 c ho odiato mangiar lo
 R -r +r +r
 d ho voluto mangiar lo
 R R +r +r

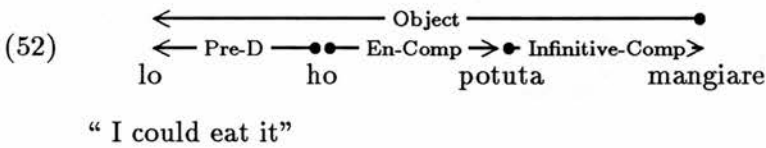
The example in (51a) falls under the rule given in (50) whereby a non-restructuring verb may take a verbal complement that is either restructuring or not. If the verbal complement is +r then any clitics that might be present may not cliticise onto the matrix verb by (46). In (51b) we have the case of a restructuring verb that takes a verbal complement. All such verbs are unspecified for the feature +/- r. This is symbolized by the capital R. In combining with a verb that is +r as in (51b) the variable *R* becomes instantiated to the feature of its dependent namely +r. As such the clitic may attach to the verb that subcategorised for it, as in (51a), or to the matrix verb. In (51c) the matrix restructuring verbs assumes the feature -r and so Clitic Climbing is prevented. In (51d) both restructuring verbs get their feature for restructuring from *mangiar* and hence Clitic Climbing is possible. As only verbs that take nominal complements are specified as +r and only the non-restructuring verbs are specified as -r, all the others being unspecified for restructuring, it should be obvious that Clitic Climbing will be permitted up to but not onto the highest non-restructuring verb in a string. This in effect allows Clitic Climbing to the left of the highest restructuring verb in a chain. This is in line with Rizzi's original intuitions without the stipulation that Clitic Climbing is obligatory for reasons that we saw above.

We have also captured the hopping nature of clitic climbing and justified the unusual position of having a restructuring slot for verbs that were not generally considered to be classified as such, i.e. transitive and intransitive verbs and of actually making them restructuring verbs i.e. assigning them the feature +r rather than -r as might have been thought. Making what were considered the restructuring verbs as unspecified for the feature +/- r also captures in a very simple way the intuition behind all the restructuring analyses

especially when seen in terms of parallel structures. That is being both restructuring and non-restructuring verbs simultaneously.

5.6.6 Object Agreement

The analysis given above for the non-exceptional behaviour of clitics and for Clitic Climbing can be seen to be more or less well motivated. To accommodate Object Agreement we must slightly change the account. Normally in Italian objects do not agree with verbs or vice versa. However, clitic objects (and not dative clitics) agree with past participles. This happens irrespective of whether they are arguments of that participle or not. The relevant example is repeated below with the dependency arrows included:



Given that here the agreement is between a clitic and a participle of a verb that does not subcategorise for that clitic, any account in terms of argument structure will have to be rejected, i.e the clitic in (52) is not an object of the participle. The intuition seems to be that the auxiliary acts as an intermediary between the clitic (when there is one) and the participle. As the auxiliary itself agrees with the subject but not with the clitic it must have an additional set of special agreement features, *Agr2*, for participle agreement which are different from its actual agreement features, *Agr1*. This gives us the following entries for auxiliaries in Italian:

- (53) [ho, [auxiliary, en-comp, indicative-present, R, Agr2], Agr1]

We can then use this extra agreement slot to enforce agreement with the clitic and the participle (provided of course that R becomes instantiated to +r). This is captured in the following.

- (54) [Root, [auxiliary, en-comp, tensed, R, Agr2], Agr1] NEEDS
 [op, pre-d, [clitic, Def, Count, R], Agr2]

- (55) [Root, [auxiliary, en-comp, tensed, R, Agr2], Agr1] NEEDS
 [ob, post-d, [verb, Subcat, past-participle, R], Agr2]

The statements in (54) and (55) will correctly enforce Object Agreement between a clitic and a participle. Let us look at an example:

- (56) a *la ho mangiata*
 b [la, [clitic, Def, Count, +r], [3, sing, fem, acc]]
 [mangiata, [verb,t,participle, +r], [3, sing, fem, nil]]
 c [ho, [auxiliary, en-comp, tensed, R, Agr2], [1 ,sing, Gend, nil] NEEDS
 [op, pre-d, [clitic, Def, Count, R], Agr2]
 [ho, [auxiliary, en-comp, tensed, R, Agr2], [1, sing, Gend, nil] NEEDS
 [ob, post-d, [verb, Subcat, past-participle, R], Agr2]

In the above, (56b) represents the lexical assignments given to *la* and *mangiato*. (56c) represents the *NEEDS* requirements of the auxiliary, *ho*. In combining with *la*, the remaining statement of (56c) is instantiated to (57) below:

- (57) [ho, [auxiliary, en-comp, tensed, +r, [3, sing, fem, acc], [1, sing, Gend, nil]]
 NEEDS
 [ob, post-d, [verb, Subcat, past-participle, +r], [3,sing, fem, acc]]

In other words (57) will only be met by a third person, singular, restructuring, participle. Fortunately, *mangiata* is precisely this. With any other form of the participle unification will fail.

In this chapter then I have accounted reasonably comprehensively for the phenomena (37a) to (37c) and (37f) and (37g) given at the end of the previous chapter. The phenomena of Change Of Auxiliary referred to in (37d) and (37e) has proven too complex to include in the present study. I have not attempted (37h). In the process I have also completed reasonably successfully the goal in (1c) of Chapter 1.

5.7 Advantages and Disadvantages of The Approach

There are some problems with the analysis however. There is no account of clitic movement within NP, nor of Change Of Auxiliary, nor of why clitics must form clusters even if they are arguments of different verbs as in (39), nor as we saw of making clitic movement bound within a tensed clause.

We now move on to a more specific problem concerning the analysis of Object Agreement. Firstly, the statements in (54) and (55) enforce agreement between the participle and *any* clitic, including the dative clitic. This is obviously incorrect. We need some way of capturing the fact that the participle agrees with the clitic if and only if the clitic is the object clitic. One way to solve this problem is by performing the same trick of feature passing with the clitics that was achieved with the auxiliary. Namely to add a feature slot for this special agreement to clitics. This would now transform (45) into (58) and (54) into (59):

(58) [Root, [clitic, Def, Count, R, Agr2], Agr1]

(59) [Root, [auxiliary, en-comp, tensed, R, Agr2], Agr1] NEEDS
 [op, pre-d, [clitic, Def, Count, R, Agr2], Agr3]

Adopting (58) and (59) would achieve the required results for the following reasons. Firstly, in the object clitic case *Agr1* and *Agr2* in (58) would be the same. In the dative clitic case, on the other hand, *Agr1* in (58) would be specified as neuter. Given that it is *Agr1* that is used to induce agreement on the participle as in (58) this will achieve the right results, namely the use of the neuter participle with dative clitics and an agreeing participle with the object clitics. In cluster we would have to make the object clitic the head of the dative clitic so that the object clitic was always the link with the verb and hence the clitic whose special agreement slot was relevant for Object Agreement.

There remains, however, one problem. This is when there is no clitic. The problem is how to prevent the participle randomly inflecting for agreement. This is a possibility because if we recall, the special agreement slot of the auxiliary is unspecified and as such will unify with anything. If there is no clitic then there will be nothing else to ground the relevant agreement variable and hence nothing to prevent the unification with an inflected participle form going ahead, licensing (60) below:

(60) *ho mangiata la torta

One way to stop this spurious agreement would be two entries for the auxiliary, one a restructuring, the other a non-restructuring. The non-restructuring case would not take

a clitic and would take the third person masculine form (or alternatively the neuter form mentioned above) of the participle. The restructuring auxiliary case would obligatorily take a clitic and also enforce the special agreement on the participle.

This is not very satisfactory, There are two points to make in defence of the overall account. Firstly Object Agreement is a peripheral phenomena, that is in many regions it is receding. Secondly, this data is problematic for any framework to account for.

There are several advantages to the analysis just presented. Firstly, there has been no need to make a distinction in the grammar between proclitics and enclitics. Secondly clitics have been classified as a subset of the class of pronouns and prepositions and their status as words have been captured. These include the fact that they can fulfill the subcategorisation requirements of verbs. This is line with the Government and Binding analysis described above.

The analysis in terms of bicephality has allowed us to capture the non-exceptional behaviour of clitics with the minimum of changes to the grammar. In fact all that was required was three extra NEEDS propositions and the addition of a feature for Object Agreement. Requiring clitics to obligatorily have two heads and only allowing the correct subset of verbs to sanction the second dependent link means that the limited distributional positions have been correctly captured. The approach is also in line with the Government and Binding analysis that treats clitic movement differently to other kinds of movement.

In clitic clusters, by making one clitic dependent on the other we have managed to treat single clitics and clitic clusters uniformly, rather than make clitic clusters separate entries. The analysis given above has also allowed us to capture the rigidly governed ordering of clitics without appeal to any extra devices.

The framework adopted has also allowed us to state the generalisation that finite verbs take optional proclitics, and that non-finite verbs with the exception of past participles take enclitics with great ease, by simply overriding the relevant propositions at the past participle node.

Using the Word Unification Grammar notion of optionality has allowed us to capture

Rizzi's simple intuition that certain structures are available to some verbs but not to others quite simply and without appeal to restructuring yet managing to capture the fact that in both restructured and non-restructured cases we are dealing with the same lexical item. It could also be seen as compatible with a base-generation analysis. Further, we have been able to give the analysis in terms that captured the exceptional behaviour and its hopping style.

In conclusion, then, the account presented above compares favourably with previous accounts not only in terms of coverage but in terms of incorporating certain basic insights into the phenomenon into the analysis. Additionally, when compared against the benchmarks presented at the end plural formation Chapter 4, all the phenomena except those involving Change Of Auxiliary have been accounted for. The fact that Change Of Auxiliary has not been incorporated into the analysis has to do with the time and space available to develop an account. We can confidently say that not only has goal (1c) of Chapter 1 been achieved but that it has been reasonably successfully achieved. We now move on to goal (1d) of Chapter 1. This will occupy the next chapter.

Chapter 6

Word Unification Grammar and Categorial Grammar

The aim of this chapter¹ is to compare and contrast Word Grammar and Categorial Grammar. There are several reasons why comparison would be advantageous. Firstly, there are very many diverse frameworks around, the relations between which are unclear. Given that the comparison is successful in the sense that translations may be made, whole bodies of work may be consolidated and unified. Further, results and insights gained through pursuing one framework may be transferable to another.

Comparison also has the advantage that sets of frameworks become more accessible to researchers in both the frameworks compared and different analyses of data may be compared more precisely. Further still, as part of the process described by Chomsky in Chapter 1, translation between frameworks helps illuminate inadequacies not only in the analysis of data but in the frameworks themselves. Even if translation is not possible, comparison will still help to make it clear where the similarities and differences lie, and what the advantages and disadvantages of pursuing a given framework may be.

¹I would like to thank Guy Barry and particularly Mark Hepple for discussion and comments on this chapter. Any errors are of course my own.

6.1 Categorical Grammar

A Categorical Grammar as originally proposed by Ajdukiewicz (1935) consists of two components. The first is a categorial lexicon. The second consists of a small number of rule schemata called combination rules. In its original form, this component was restricted to a single non-directional rule of Functional Application.² Categorical Grammars of this form will be called Applicative Categorical Grammars.

In the categorial lexicon, each entry includes a category, defining the kind of constituent (if any) with which the word in question can combine and the kind of constituent that results, in such a way that,

“two words or expressions belonging to the same class can be substituted for one another in a context possessing unified meaning without that context becoming an incoherent word pattern and losing unified sense”

(Ajdukiewicz 1935, p208)

The Applicative Categorical Grammar notation has the same weak generative capacity as the phrase structure rules of other grammars. In a Categorical Grammar, functions and arguments are distinguished in the lexicon. In the simple case, arguments bear atomic categories and functions bear compound categories of the form X/Y where X and Y vary over any category, atomic or compound. Items having the category X/Y or $(W/X)/Y$ and so on are functions over Y . Such functions may be seen as mapping between entirely

²A combination rule is simply a rule that determines how words may combine with each other to form larger strings. As we will see below the single non-directional rule of application becomes two directional rules in later developments; forward application (FA in examples (22) and (24)) and backward application (BA in examples (22) and (25)). What is meant by this is that given the following sentence and the assignment of categories,

NP	S\NP/NP	NP
John	loves	Mary

loves may forward combine with *Mary* as symbolised by the forward leaning slash. The result of this combination is a category for *loves Mary* of $S\NP$. This may then backward combine with *John* as again indicated by the backward leaning slash. The result of this application is a string with the category S . We will see further examples of combination rules below.

syntactic domains. However the categories can also be thought of as a shorthand for the semantics of the entities in question.

“The numerator will be the index of the semantic category to which the whole expression composed of the functional sign plus its arguments belongs, while in the denominator appear, one after the other, the indices of the semantic categories of the arguments, which together with the semantic functor combine into a significant whole.”

(Ajdukiewicz p210, 1935)

So, syntactic rules and categories correspond with rules of semantic interpretation. The function categories can therefore be thought of as mapping between semantic representations. Originally, the combinatory rule of Functional Application was non-directional. This, then, is a Categorical Grammar of the Ajdukiewicz type.

However, in order to write real grammars for strongly configurational languages (like many natural languages) this basic apparatus is both too restrictive and not restrictive enough. It is not restrictive enough in virtue of the fact that both the rule of Functional Application and the category labels are non-directional, allowing many reductions to be made that are unwanted. It is too restrictive in virtue of the fact that category labels place an ordering on which element must be combined with first and so on. The existence in natural language of unbounded movement, discontinuous constituents, extractions and so on (i.e. constructions in which elements that belong together in the semantics are separated) make it clear that something is required to augment the basic ingredients of function and argument categories plus Functional Application prescribed by Ajdukiewicz.

Those who have attempted to cope with these phenomena within Categorical Grammar have tried to generalise or extend the basic apparatus by the addition of various operations on the functions and arguments, over and above simple Functional Application. These additions include Functional (Partial) Composition (Steedman 1982, 1985, Dowty 1988) Type Raising (Dowty 1988) Wrapping (Bach 1979, 1988) Combinators (Steedman 1988) or the use of the Lambek Calculus (Lambek 1961, Moortgat 1987) among others. It is the inclusion of one or more of these extensions that mark the change from an Ajdukiewicz type Applicative Categorical Grammar to more flexible manifestations and also its applications

from logical languages to natural languages. Categorical Grammars incorporating any of the above extensions will be referred to as Flexible Categorical Grammars.

These various attempts at extending basic Categorical Grammar have given rise to many competing varieties of the theory. It will be impossible to discuss them all here. I will take it that these extensions are different attempts to capture the data of natural language but all incorporating the same basic intuitions about the grammar and the lexicon.

6.2 Comparison

Having given a brief overview of Categorical Grammar, I will turn to a discussion of the relationship between Categorical Grammar and Word (Unification) Grammar.

6.2.1 Lexicalism

The lexicalist character of Categorical Grammar arises, in part, from the fact that a word is assigned a syntactic category in the lexicon. Of course by referring predominantly to words and their requirements, Word Unification Grammar is also lexicalist in this sense. The crucial difference is that in Categorical Grammar, specific information about word order, amongst other things, is encoded into the lexical entries. As a result the syntactic component may be reduced to very general mechanisms for combining elements. A Categorical Grammar thus condenses much of the work done by the syntactic rules in Phrase Structure Grammars into the lexicon, and hence into lexical entries. It is in this sense that Categorical Grammar may be said to be lexicalist.

It thus appears that there is agreement on where the burden of the grammar should lie; namely the lexicon. But precisely how this is achieved and what should be incorporated leads to differences in realisation.

6.2.2 Constituency

The advantage of the Categorical Grammar notation is that each entry contains both categorial and subcategorial information simultaneously. In this respect it has less redundancy

than most Phrase Structure Grammars that incorporate both Phrase Structure rules and subcategorisation information.

Although it is claimed that Categorical Grammar is a Phrase Structure Grammar, it differs from them in many respects. When Applicative Categorical Grammar is augmented with devices like Functional Composition to a Flexible Categorical Grammar, where and what the *phrases* will be is to a large extent a matter of choice. Indeed the notion of *phrase* in Flexible Categorical Grammars bears little resemblance to the notion of *phrase* in standard Phrase Structure Grammars. There are two ways to reduce the example in (1) below to NP, given (for example) Functional Composition:

- (1) NP/N N/N N
 determiner adjective common noun

On the one hand, we could combine the determiner and the adjective first by Functional Composition; on the other hand we could combine the adjective and the common noun first by Functional Application. Both methods of reduction will ultimately lead to NP. This would not be an option in a standard Phrase Structure Grammar. Further, given that the determiner and the adjective may combine to form an element of the exactly the same category as a determiner it is unclear if what is involved is even the more general part:whole relation. This situation holds for many extensions of Applicative Categorical Grammar.

Given these differences with standard Phrase Structure Grammars it is hard to reconcile a Flexible Categorical Grammar with traditional notions of constituency. As a consequence, and to avoid confusion, the term *phrase* will be employed for those structures arrived at by using Phrase Structure Grammars and the term *word string* will be used for structures arrived at by using a Flexible Categorical Grammar. There will of course, be some overlap. In the move away from constituency to more flexible groupings of words, Flexible Categorical Grammars and Word Grammar seem to be treading similar paths.

6.2.3 Classification

Another way in which Categorical Grammar differs from Phrase Structure Grammars involves the classification of lexical items. Whereas in standard Phrase Structure Grammars the syntactic type of any item (i.e. whether it is a verb or a noun or whatever) is apparent from its label, and the number of possible labels is limited, in Categorical Grammar this is not the case. The total set of possible categories is recursively defined from a set of base categories. There is no reference to, or restriction of, the word class of an element being classified by a category,³ in the sense of the classificatory hierarchy given in Chapter 2. One consequence of this is that the word class of every category must be stipulated. In Word (Unification) Grammar as we have seen every word enters into the classificatory hierarchy and hence its syntactic type is more apparent, allowing amongst other things, cross categorial generalisations to be made. This brings Word (Unification) Grammar more into line with Head-Driven Phrase Structure Grammar, and Unification Categorical Grammar which can be seen as Categorical Grammar like approaches incorporating inheritance.

6.2.4 Heads

In section 6.2.2 it was stated that Applicative Categorical Grammars are similar in some respects to Phrase Structure Grammars. If this is the case, then the categories of Categorical Grammar seem most similar to the pre-X-bar theory Phrase Structure rules. In X-bar theory it is stipulated that all phrase structure rules instantiate the following schema:

$$(2) \quad X^m \longrightarrow \dots X^n \dots$$

where m and n refer to the number of bars and possibly $n = m$

The move to X-bar theory was made because it was realised that rules of the form $NP \longrightarrow \dots \text{Noun} \dots$ did not capture the intuition that an NP had to crucially contain a noun, a problem that the $N^m \longrightarrow \dots N^n \dots$ notation was designed to remedy. This being the case it would seem that Applicative Categorical Grammar has not encoded and therefore has missed some of the insights of X-bar theory.

³Bouma (1988) referring to Hoeksema (1984) has attempted to overcome this problem by using the notion of *reducibility* whereby the syntactic class of a word may be derived from its category

In order to see the significance of X-bar theory we need to look at some of the deeper, linguistic motivations behind it. One of the main reasons for setting up X-bar theory was the recognition by transformationalists of the importance of the notion of head as used by more traditional theories. This brings us to the most central point in any discussion of dependency and Categorical Grammars.

Word Grammar, as we have seen, takes the dependency notion of head as basic. In an important article (countering Zwicky (1985b)) Hudson (1986) attempts to unify the properties of the notion of head and to demonstrate its centrality to linguistic theory. Though there exists no absolute definition of head one of its most significant properties is that it is usually the semantically central element of a pair of words. In his defence Hudson takes Zwicky's characterisation which is as follows:

“we could take it that the head/modifier distinction is at root semantic; in a combination $X + Y$, X is the 'semantic head' if speaking very crudely, $X + Y$ describes a kind of the thing described by X ”

(Zwicky, 1985b)

This could be likened to the Categorical Grammar approach which, we remember, uses notions of semantic functor and argument. Here, then, is the potential point of convergence. Word Grammar has heads and dependents; Categorical Grammar has functions and arguments. In Word Grammar, the head of one word may be the dependent of another and vice versa. In Categorical Grammar a similar relation between functions and arguments exists. In order to pursue this, we need to see if, where and why the assignment of head in Word Grammar differs from the functor of Categorical Grammar. That is to say, we need to see if the functors of Categorical Grammar correspond with the heads of Word Grammar and if the dependents of Word Grammar correspond with the arguments of Categorical Grammar.

The assignment of head in Word Grammar, as we saw from the table in (44) of Chapter 5, is based on the possession of some subset of the features that characterise heads. The first row of that table has the feature “semantic functor”, but in dependency this feature is not identical with the notion head. In a Categorical Grammar lexicon what is the basis

for assigning an entry in the lexicon a function category? One of the deciding factors seems to be that if an element obligatorily subcategorises for another then it is assigned a function category. This is identical with subcategorising for obligatory dependents in Word Grammar. It is also in line with the idea that heads are the subcategorisand in a head dependent pair, as presented in the table in the Chapter 5. For example, the category for transitive verbs may be something like the following:

(3) S\NP/NP

This can be taken to mean that a verb is something that requires two NPs to make a sentence. In such cases of obligatory subcategorisation, we see a perfect matching of functor category and the notion head in Word Grammar.

Optionality is not encoded in a Categorical Grammar *per se* because the categorial information about what arguments a function takes, once given in the lexicon, is fixed. There are two types of optionality that need to be considered, optional arguments and optional adjuncts. Categorical Grammar handles various argument options that a word might have by having multiple categorial entries in the lexicon. That is to say it makes the options for combination obligatory in a sense by including a different entry in the lexicon for each possible option. If for example we wanted to say that a verb like *eat* is intransitive but takes an optional object, Categorical Grammar would capture these facts by having two entries in the lexicon for the verb *eat* as in (4a) and (4b). Alternatively this could be handled by a lexical rule if it was felt that optional transitivity was a general enough process. Word Unification Grammar on the other hand could handle it by use of the feature *op* as in (4c) or again, if the process was felt to be general enough, by creating a further node in the subcategorisation hierarchy such as *optional transitive*.

(4) a eat S\NP (intransitive)
 b eat S\NP/NP (transitive)
 c [eat, [normal, intransitive, Fin], Agr] NEEDS
 [op, object([Root, [noun...], Agr2])]

To a certain extent the differences between the two approaches is just a question of how one chooses to express the disjunction of information. However, in line with examples like

(4a) and (4b) it is acceptable in Categorical Grammar that for any new construction that confronts a lexical item, a different category can be postulated for it. This has the result that almost all items have multiple categories. The problem with such an approach is how to decide when we are dealing with essentially one element as in (4) or when we are dealing with two (homonymy, synonymy, etc). However, we are concerned with points of similarity and difference here, rather than with problems specific to either. The important point then is that, as with obligatory arguments, the functors of Categorical Grammar in the case of optional arguments and the heads of Word Grammar are co-extensional.

The second kind of optionality mentioned involves adjuncts (e.g. adjectives, adverbs, relative pronouns etc.). Because adjuncts, unlike arguments, may usually occur an unrestricted number of times, they present slightly different problems. The usual way adjunct optionality is handled in Categorical Grammar is by making adjuncts functions over the elements they are adjuncts of. What this means is that in Categorical Grammar adjuncts are not subcategorised for by any element, but subcategorise for elements that they are adjuncts of. In Word Grammar, on the other hand, adjuncts are considered optional dependents⁴ (that may occur an unrestricted number of times) and therefore elements that are subcategorised for. This contrast is evident in the following:

- (5) a prenominal adjectives are assigned the category N/N i.e are functions over nouns
 b [Root, [common noun...],Agr] NEEDS
 [op, adjunct[Root, [adjective...], Agr]]

What we see in the above is that in Categorical Grammar adjectives “subcategorise” for common nouns as in (1) above, whereas in Word Grammar it is the other way around, namely that common nouns subcategorise for adjectives. This has a corollary that in Word Grammar the common noun is the head of adjective. Given that we are attempting to compare the functors of Categorical Grammar with the heads of Word Grammar, this would seem to be a case where functors and arguments are out of step with heads and dependents; that is to say in Word Grammar adjuncts are dependents whereas in Categorical Grammar

⁴q.v. (22) of Chapter 2.

adjuncts are functors and apparently heads.⁵

Let us look in more detail at the motivation for the respective approaches, with specific reference to the relation between adjectives and common nouns in English. We can argue on syntactic grounds that the common noun in the common noun/adjective relationship is the head. Firstly, a common noun may appear without a modifying adjective but not vice versa. Secondly, when the adjective *is* present it seems reasonably clear that it takes its position from the common noun. These are some syntactic arguments for making the common noun the head.

In terms of functor and argument semantics it seems clear that we want the adjective to be a functor over common nouns and not vice versa. This is because we want to say that nouns are not obliged to take adjectives and as such are *complete* without them. The alternative would be to allow common nouns to be incomplete in the sense that they would be functors that need not take an argument. Given that we want to allow for an unspecified number of adjectives to occur, we have a second reason for making adjectives functions over common nouns.

Excluding adjuncts, however, there seems to be a large deal of overlap between the head of Word Grammar and the functor of Categorical Grammar. How can this conflict be resolved? One way is suggested by Bach (cited by Bouma (1988)). Bouma discusses the possibility of deriving the notions of *head*, *complement* and *modifier* from the category labels as follows:

- (6) a HEAD
In a construction consisting of a functor, F, and an argument, A, the head constituent is F, unless it is endotypic (that is, F is of the form X/X or X\X) in which case, A is the head.
- b COMPLEMENT
An argument which is not a head is a complement
- c MODIFIER
A functor which is not a head is a modifier

If (6) above holds in general, it seems possible from the category assigned to adjectives and other modifiers, which all have roughly the same form, namely X/X, to derive the fact

⁵Although Head-Driven Phrase Structure Grammar adopts a largely Categorical Grammar approach to subcategorisation, it resembles Word Grammar in its treatment of adjuncts.

that though a given word might be a semantic functor it is not necessarily the head.

If we reinterpret this in a Word Grammar framework we are presented with a novel way of looking at the common noun/adjective relationship, and at the notion of subcategorisation. Given the approach just described, all that labels such as N/N seem to be saying is that modifiers subcategorise for their heads rather than the other way around. This goes against traditional concept of subcategorisation where it is only dependents that are subcategorised for. We are now faced with the suggestion that subcategorisation and dependency are no longer identical, i.e. it is possible for one element to subcategorise for another and for the first to be a dependent of the second. We thus require some more general notion that subsumes subcategorisation (a relation between heads and dependents) and that of requiring a head. This is precisely what we have in the NEEDS relation of Word Unification Grammar. There we find statements of both kinds, namely those found in (7):

- (7) a X NEEDS head
- b X NEEDS dependent.

So it seems that the NEEDS relation incorporates both the notion of subcategorisation and of requiring a head. How can we incorporate the discussion about adjuncts into the Word Unification Grammar framework? Recall that Chapter 3 we saw two distinct types of subcategorisation proposition:

- (8) a [Root, GF, Agr] NEEDS [Op, GR, [Root2, GF2, Agr2]]
- b [Root, GF, Agr] NEEDS [Op, head]

As mentioned there, the reason for the difference between the right hand sides of (8a) and (8b) is due to the fact that dependents do not usually specify requirements on their heads. On the contrary, heads specify the properties of their dependents. The approach discussed in Bouma (1988) suggests that we should extend the right hand side of 8b) for adjuncts along the lines of (8a). The effect of this would be that adjuncts could then specify properties of their heads. We could then say that the functors of Flexible Categorical Grammar were identical with the subcategorisers of Word Unification Grammar. Further, what is the head in one framework is the head in the other. In other words the grammars would be strongly equivalent.

To develop this approach further would require more time than is available and all I can hope to do here is to be suggestive. However, if this approach can be made to work, there seems to be little of substance to the differences between heads and functors, and in fact some insights have been gained in the comparison.

6.2.5 Further Differences

There are further differences connected with the category labels assigned in Categorical Grammar. Firstly a Categorical Grammar label imposes an ordering on which element (dependent) must be found first, Word Grammar does not. More generally, Categorical Grammar gives a name to the substructures it creates after combination, whereas Word Grammar does not.

It thus seems that Categorical Grammar places a great deal of information into the categories assigned to words. The information that they contain includes at least categorial and subcategorial information, an ordering on dependents, and semantic information. In Word Grammar these different types of information are handled separately.

Further, as we saw in Chapters 3 and 5, Word Grammar also allows for the possibility that one word may have two heads, and for the possibility that in such a case one of the heads may be predominantly semantic. A case in point are the examples of Control. In Categorical Grammar two functions may not take the same argument. The desired effect is achieved in Categorical Grammar analyses of Control where we find complex semantic labels creating the necessary semantic links between controller and controlled without requiring that elements be the argument of more than one functor. None of the points mentioned present serious theoretical differences,

A final difference is that in Word Grammar Grammatical Relations are basic whereas in (some types of) Categorical Grammar they may be derived. As I have had very little to say about Grammatical Relations in Word Unification Grammar, and as the choice between them being basic or derived is not strictly tied to dependency, I will be content with merely noting the difference here.

We could say in summary then that the differences between (Flexible) Categorical Gram-

mars and Word (Unification) Grammar are the following:

- (9) Categorical Grammar
 - a has no intrinsic (hierarchical) classificatory system of word type as basic though they may be derived
 - b encodes optionality of arguments by having multiple entries
 - c has no intrinsic notion of head, complement or adjunct though again these may be derived
 - d encodes optionality of adjuncts by making them functions over their heads.
 - e places an ordering on dependents
 - f creates sub-structures when combinations have been made
 - g has no intrinsic notion Grammatical Relations as basic though again these may be derived
 - h has a simple Adjacency Principle of strict adjacency

6.3 A Categorical-Like Word Grammar

Let us pursue the comparison begun in the previous section by developing a Categorical Grammar that is minimally different from a Word (Unification) Grammar. In order to do this we will ignore the differences in inheritance and Grammatical Relations which are specific to Word Grammar. Having done this we may then transfer the solution for clitic placement that I have constructed in Word Grammar to a Categorical Grammar and then to compare the solutions.

6.3.1 Clitic Placement

Firstly we should remember that the crucial part of the solution to the placement of clitics was based on the idea that clitics obligatorily require two heads. This is not directly translatable, because head requirements are not explicitly stated in Categorical Grammar.

However given the following list of categories:

- (10) The set of categories is the smallest set such that
 - a) $\text{Basecat} = \{V, N, P, \text{CN}, \text{cl}, \text{Det}\}$
 - b) $X \in \text{Basecat} \longrightarrow X \in \text{Cat}$
 - c) $X, Y \in \text{Cat} \longrightarrow X \backslash Y, X / Y, X.Y \in \text{Cat}$The use of the non-directional slash, |, is notational and is used to cover cases of both the forward slash, /, and the backslash, \.

we can assign clitics the following categories in the lexicon

- (11) a Accusative clitics will be assigned the category N.cl
- b Dative clitics will be assigned the category P.cl
- c Clitic Clusters will be assigned the category P.N.cl

What is intended by the category X.Y is a product; i.e. N.cl is the product category Noun *and* clitic. In Categorical Grammar it could be argued that category labels may be interpreted as implicitly requiring functors to combine with. By giving clitics a product category, I would argue that we have captured the intuition that clitics require two heads. We will see this more clearly in the examples below. The semantics associated with these product categories is as follows; the *cl*, in virtue of being merely a syntactic connection between words, is assigned some arbitrary constant, α , which has no effect. The other part(s) of the category label are assigned the same semantics as their non-clitic counterpart. This assignment of categories is not too difficult to justify given the fact that for clitic clusters such as *glielo* (*gli + epenthetic vowel + lo* we will probably need to assign a (product) category consisting of a noun and a preposition (phrase).

One significant difference between the lexical analysis presented above and the analysis presented in Chapter 5 concerns the category assigned to the clitic cluster in (11c). The assumption we have been working under so far is that by assigning clitics a product category consisting of two separate categories, we have captured the fact that each clitic needs two heads. However, by assigning the cluster a product category consisting of three elements as in (11c) we are effectively saying that clitic clusters only require three heads, not four as in the Word Unification Grammar solution. This recalls the discussion in Chapter 3 about how to deal with such clusters. In this new framework, however, there are no statements about heads. Having one clitic depend on another clitic, which was the approach adopted in Chapter 3, seems unnecessary and unjustified. The correct solution to adopt here seems to be that however many clitics there are they need just one extra head (i.e. a host to cliticise onto).

Before we continue we will need to state some reduction rules:

- (12) Forward Application (FA)
- | | | | | |
|-----|---|---|---|----|
| X/Y | + | Y | ⇒ | X |
| f | | a | | fa |

- (13) Backward Application (FA)

$$\begin{array}{ccc} Y & + & X \backslash Y \Rightarrow X \\ a & f & fa \end{array}$$
- (14) Forward Composition (FC)

$$\begin{array}{ccc} X/Y & + & Y/Z \Rightarrow X/Z \\ f & g & \lambda a.f(ga) \end{array}$$
- (15) Backward Composition (BC)

$$\begin{array}{ccc} Y \backslash Z & + & X \backslash Y \Rightarrow X \backslash Z \\ f & g & \lambda a.f(ga) \end{array}$$
- (16) Forward Product 1 (FP1)

$$\begin{array}{ccc} X/Y & + & Y.Z \Rightarrow X Z \\ f & a.b & fa b \end{array}$$
- (17) Forward Product 2 (FP2)

$$\begin{array}{ccc} X/Z & + & Y.Z \Rightarrow X Y \\ f & a.b & fb a \end{array}$$
- (18) Backward Product 1 (BP1)

$$\begin{array}{ccc} Y.Z & + & X \backslash Y \Rightarrow Z X \\ a.b & f & b fa \end{array}$$
- (19) Backward Product 2 (BP2)

$$\begin{array}{ccc} Y.Z & + & X \backslash Z \Rightarrow Y X \\ a.b & f & a fb \end{array}$$

It might be thought that extra rules will be required to cover the case of the clitic cluster where we have categories of the form $X.Y.Z$. But as the product is commutative, $X.Y.Z$ is equivalent to $X.(Y.Z)$ and $(X.Y).Z$ which will then make them fall under the rules (16) to (19) given for products.

6.3.2 The Non-Exceptional Behaviour of Clitics

Before we can look at a solution for clitic placement in Italian, we need the following rule to capture the fact that Italian is a pro drop language. This may be stated as follows:

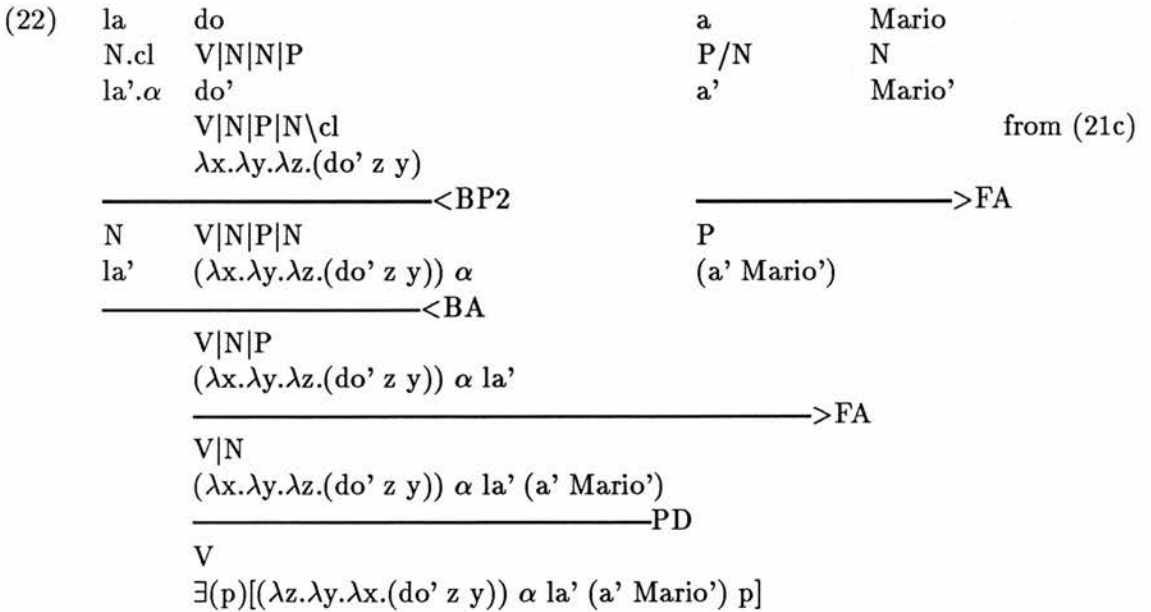
- (20) Pro-Drop (PD)

$$\begin{array}{ccc} V|N & \Rightarrow & V \\ f & & \exists x[fx] \end{array}$$

Then we need the following lexical rules for transitives and ditransitives, to capture the fact that they may optionally take pre-clitics when finite, represented by the feature $+f$, for finite:

- (21) a
$$\frac{V_{+f,+r} | N | N}{S} \longrightarrow \frac{V_{+f,+r} | N | N \backslash \text{cl}}{\lambda x.S}$$
 where x is not free in S
- b
$$\frac{V_{+f,+r} | N | N | P}{S} \longrightarrow \frac{V_{+f,+r} | N | N | P \backslash \text{cl}}{\lambda x.S}$$
 where x is not free in S
- c
$$\frac{V_{+f,+r} | N | N | P}{S} \longrightarrow \frac{V_{+f,+r} | N | P | N \backslash \text{cl}}{\lambda x.\lambda y.\lambda z.Syz}$$
 where x is not free in S

What the above rules mean are that any verbs of the form given in (21) that are also finite may take a proclitic. The vertical slashes abbreviate the fact that the arguments may either come before or after the verb and so is the translation of the assumption that nominal dependents are free. ⁶ The assignments in (21) also capture the fact that cliticisation is a purely syntactic phenomenon by vacuous abstraction of the clitics over functions that take them. These rules, in addition to some reasonable assignment of categories to words, will give us the following possible derivations:

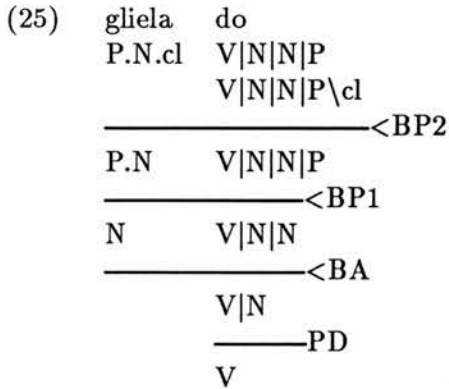
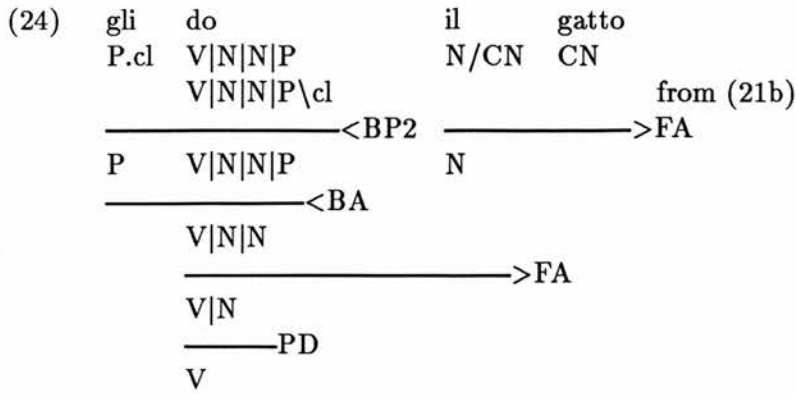


(23) below represents the semantics with β reduction.

$$(23) \quad \exists(p)[\text{do}' (\text{a}' \text{ Mario}') \text{ la}' p]$$

We also have the following derivations:

⁶The two different orderings in (21b) and (21c) are required for cases where we only have one proclitic (either the P or the N) and a preceding argument. There is also a certain amount of overlap in the application of these rules.



From the examples above the intuition that there is a purely syntactic relation involved in cliticisation has been captured by the separating off of the clitic link. In the above cases this might seem strange but in cases of Clitic Climbing this analysis seems less so. Let us now move on to infinitivals. In order to capture the fact that clitics may be enclitic on infinitivals we need rules of the following kind (with semantics along the lines of (21) above):

- (26) a $V_{-f,+r} | N | N \rightarrow V_{-f,+r} | N | N/cl$
b $V_{-f,+r} | N | N | P \rightarrow V_{-f,+r} | N | N | P/cl$
c $V_{-f,+r} | N | N | P \rightarrow V_{-f,+r} | N | P | N/cl$

In (26) as in (21) the rules presented are optional. It thus seem that we have captured the non-exceptional behaviour of clitics very simply. Let us move onto the exceptional behaviour.

6.3.3 Clitic Climbing

In this section we will have to make use again of the feature +/-r(estructuring) Assume that we have the lexical assignment of categories to verbs as in (27):

- (27) a voglio $\rightarrow V_R | N / (V_R | N)$
 b odio $\rightarrow V_{-r} | N / (V_R | N)$
 c ho $\rightarrow V_R | N / (V_R | N)$

As in Chapter 5, transitives and ditransitives will be assigned the feature +r, non-restructuring verbs that take verbal complements will be assigned the feature -r and restructuring verbs taking verbal complements will be assigned the variable R. From (27) we can also see that restructuring verbs that take verbal complements take their value for restructuring from their complement while non-restructuring verbs that take verbal complements place no such restrictions on the kind of verb they may combine with. Additionally, clitics will have a feature for restructuring that they will pass on to the verbs that they are cliticised onto. We now need a rule that will allow restructuring verbs like those in (27) to take clitics. This is given in (28):

- (28) Clitic Climbing Rule
 $V_{+r} | N / (V | N) \rightarrow V_{+r} | N / (V_{+r} | N) \setminus cl_{+r}$

Given that verbs like those in (27a) will be assigned a variable for their restructuring value they will match (optional) rules like (28). The rule in (28) will then work in exactly the same way as that described for the non-exceptional behaviour of clitics. Some examples are given below:

- (29)
- | | | | |
|---------------|--|------------------|----------------------|
| la | voglio | mangiare | |
| $N.cl_{+r}$ | $V_R N / (V_R N)$ | $V_{+r} N N$ | |
| la'. α | voglio' | mangiare' | |
| | $V_{+r} N / (V_{+r} N) \setminus cl_{+r}$ | | |
| | $\lambda x.voglio'$ | | |
| | | | |
| | | | <BP1 |
| N | $V_{+r} N / (V_{+r} N)$ | | |
| la' | $(\lambda x.voglio') \alpha$ | | |
| | | | |
| | | | >FC |
| | $V_{+r} N N$ | | |
| | $\lambda z.((\lambda x.voglio') \alpha (mangiare' z))$ | | |
| | | | |
| | | | <BA |
| | $V_{+r} N$ | | |
| | $(\lambda z.((\lambda x.voglio') \alpha (mangiare' z))) la'$ | | |
| | | | |
| | $voglio'(mangiare' la')$ | | — β -reduction |

$$\begin{array}{rcl}
 (30) & * & \begin{array}{ccc}
 \text{la} & \text{odio} & \text{mangiare} \\
 N.cl_{+r} & V_{-r}|N/(V_R|N) & V_{+r}|N|N \\
 \hline
 & & \text{>FC} \\
 & & V_{-r}|N|N \\
 \hline
 & & \text{<BP1} \\
 cl_{+r} & & V_{-r}|N
 \end{array}
 \end{array}$$

In (29) and (30) above we see how the simple data of Clitic Climbing may be handled by the new framework. We also see the usefulness of distinguishing the two categories associated with the clitic and hence the value in saying that clitics need two heads. In (29) we see how *la* has a syntactic link with *voglio* and a semantic link with *mangiare*. In (30) we see that the reduction of the categories can only proceed so far before halting as there is no rule that will allow any category with the feature *-r* to take a clitic. Now we move onto more complicated strings. The rule for clitic movement in Italian says that a clitic can climb only up chains of restructuring verbs (+r) but not past a non-restructuring verb (-r). Again this may be captured by adopting the solution in terms of variable sharing presented in Chapter 5. This will give us the following derivations:

$$\begin{array}{rcl}
 (31) & * & \begin{array}{ccc}
 \text{la} & \text{ho} & \text{odiato} & \text{mangiare} \\
 N.cl_{+r} & V_R|N/(V_R|N) & V_{-r}|N/(V_R|N) & V_{+r}|N|N \\
 & V_{+r}|N/(V_{+r}|N)\backslash cl_{+r} & & \\
 \hline
 & & \text{<BP1} & \text{>FC} \\
 N & V_{+r}|N/(V_{+r}|N) & & V_{-r}|N|N
 \end{array}
 \end{array}$$

In (31) we can see that only restructuring verbs will allow Clitic Climbing. In (31) the derivation cannot proceed due to a mismatch in the values for restructuring i.e *ho* requires a complement that is *+r* but *odiato* is *-r* and so the sentence is correctly predicted as ungrammatical. In many ways this is clearer than the solution presented in Chapter 5 though the details are more or less equivalent.

6.3.4 Object Agreement

The Word Unification Grammar analysis of clitic placement in Italian is thus simply translated into the Categorical Grammar formalism developed. To incorporate the Object Agreement solution we require slightly more complicated feature assignments to the categories. Our earlier analysis involved a special agreement slot in the auxiliary and the clitic which

will not be necessary in the Categorical Grammar formalism. Let us look now at how this solution may be translated. We have the following assignment of categories to auxiliaries.

$$(32) \text{ ho} \longrightarrow Aux_{R,+f,Agr} | N / (V_{R,Agr2} | N)$$

In (32) *Agr* stands for the agreement displayed on the auxiliary, whereas *Agr2* stands for the special Object Agreement on the participle. The categories assigned to clitics must be amended to the following

- (33) a $N_{Agr} \cdot cl_{Agr,+r}$, for the object clitic
 b $P_{Agr} \cdot cl_{neut,+r}$, for the dative clitic
 c $P_{Agr} \cdot N_{Agr2} \cdot cl_{Agr2,+r}$, for the clitic cluster of object and dative clitic.

In (33) the agreement feature that will be used for Object Agreement is the one that is situated on the *cl* of the category. In (33a) and (33c) the agreement feature on the clitic is identical with the agreement feature on the noun. In (33b) the agreement feature on the clitic is instantiated to neuter. We could then have a rule for Clitic Climbing for auxiliaries along the following lines:

$$(34) Aux_{R,+f} | N / (V_{R,neut} | N) \longrightarrow Aux_{+r} | N / (V_{+r,Agr} | N) \setminus cl_{+r,Agr}$$

Let us look at an example:⁷

(35)	la	ho	potuta	mangiare
	$N.cl_{+r,3,s,f}$	$Aux_{R,+f} N / (V_{R,neut} N)$	$V_{R,3,s,f} N / (V_R N)$	$V_{+r} N N$
		$Aux_{+r} N / (V_{+r,Agr} N) \setminus cl_{+r,Agr}$		
	N	$Aux_{+r} N / (V_{+r,3,s,f} N)$		
		$Aux_{+r} N / (V_{+r} N)$		>FC
		$Aux_{+r} N N$		>FC
		$Aux_{+r} N$		<BA
		$Aux_{+r} N$		

This then mirrors the solution given in Chapter 5 and gives the right results. One major difference in this solution is that we have two categories for the clitic whereas in Chapter

⁷Only the agreement features on the *cl* part will be shown on the clitic as it is this that is used in Object Agreement. In the example *3* = third person, *s* = singular, and *f* = feminine

5 we only had one. This was originally adopted to incorporate the fact that clitics needed two heads. The advantage this has had is that many of the technical problems associated with Object Agreement have not arisen. These, include the problem of deciding which clitic was the head of which, the problematic case of ensuring that when there was no clitic the participle was neuter, amongst others. It thus appears that approaching a categorial analysis from a dependency point of view rather than directly from function and argument point of view has led to interesting solutions of the data.

We have now completed goal (1d) of Chapter 1. All that remains to be done is the implementation. This is to be found in Appendix A

6.4 Conclusions

The goal of this thesis was to investigate Word Grammar. The area that seemed to need most work was in the formalisation of the basic concepts before an evaluation of the framework could be attempted. This not only included work on the specifics of Word Grammar but also on the notions of Dependency, the Adjacency Principle, a statement of what a dependency grammar is and a statement of the language defined by the grammar. All these I have attempted to clarify with varying degrees of success. In developing Word Grammar it was found that the move to Word Unification Grammar, i.e. the incorporation of Unification amongst other things, was not only useful for the theory itself but in making use of standard concepts, brought Word Grammar more into line with other frameworks. This new framework was then applied with some success to the data of cliticisation in Italian.

The fact that all this was possible strongly suggests that the traditional concepts of dependency can be made specific and precise enough for the purposes of formal linguistics, in terms of formally specifying a dependency grammar. Further, the analysis of data suggests that such concepts appear to be quite useful to the formal grammarians and the descriptive linguist alike. That the relevant notions of dependency may be formally specified, has as one significant implication that such grammars may be implemented and

as such of interest to computational linguists and cognitive scientists alike. Though I make no great claims for my own framework, and consider it merely as a first stage in the endeavor, I hope this work completed so far has been suggestive enough of promise of approaches based on dependency.

Implementation

```
% Below is a listing of the grammar fragment of Italian that was
% implemented and the simple shift-reduce parser that was employed by
% the parser.
```

The Grammar

The Relations

```
% In order to get the relations to work in the correct way in Prolog
% it proved necessary to split them into a base case and a recursive
% case. They were then defined as operators to make the Prolog
% statements more legible.
```

```
:-op( 500, xfx, isa).
:-op( 400, xfx, base_isa).
:-op( 600, xfx, needs).
:-op( 500, xfx, base_needs).
```

```
% The following statements define the hierarchy for the basic word
% classes of Italian;
```

Verbs

```
% The statements for verbs are of the following schematic type;
% Verb isa [Word,[Type,Subcat,Finiteness,Restructuring],AGR].
% These are the statements where verbs are organised by Type
% n = normal i.e non-modal and non-auxiliary
% Capital letters are used for variables
```

```
[X,[n|Y],Z] base_isa [X,[verb|Y],Z].
[X,[modal|Y],Z] base_isa [X,[verb|Y],Z].
[X,[auxiliary|Y],Z] base_isa [X,[verb|Y],Z].
```

```
% Verbs organised by Type & Subcat
```

```
% i = intransitive
% t = transitive
% di = ditransitive
```

```

% inf_c = infinitival complement
% en_c = en_complement
% r = restructuring
% non_r = non-restructuring

```

```

[X,[verb,i_v|Y],Z] base_isa [X,[verb,i|Y],Z].
[X,[verb,t|Y],Z] base_isa [X,[verb,i|Y],Z].
[X,[verb,di|Y],Z] base_isa [X,[verb,t|Y],Z].
[X,[verb,inf_c|Y],Z] base_isa [X,[verb,t_v|Y],Z].
[X,[verb,en_c|Y],Z] base_isa [X,[verb,t_v|Y],Z].

```

```

% Verb organised by Type, Subcat, & Tense
% This is some of the tenses of Italian though not all were used

```

```

[X,[verb,i,finite|Y],Z] base_isa [X,[verb,i,finiteness|Y],Z].
[X,[verb,i,non_finite|Y],Z] base_isa [X,[verb,i,finiteness|Y],Z].
[X,[verb,i,tensed|Y],Z] base_isa [X,[verb,i,finite|Y],Z].
[X,[verb,i,imperative|Y],Z] base_isa [X,[verb,i,finite|Y],Z].
[X,[verb,i,present|Y],Z] base_isa [X,[verb,i,tensed|Y],Z].
[X,[verb,i,imperfect|Y],Z] base_isa [X,[verb,i,tensed|Y],Z].
[X,[verb,i,past|Y],Z] base_isa [X,[verb,i,tensed|Y],Z].
[X,[verb,i,future|Y],Z] base_isa [X,[verb,i,tensed|Y],Z].
[X,[verb,i,indicative_present|Y],Z] base_isa [X,[verb,i,present|Y],Z].
[X,[verb,i,indicative_imperfect|Y],Z] base_isa [X,[verb,i,imperfect|Y],Z].
[X,[verb,i,indicative_past|Y],Z] base_isa [X,[verb,i,past|Y],Z].
[X,[verb,i,indicative_future|Y],Z] base_isa [X,[verb,i,future|Y],Z].
[X,[verb,i,infinitival|Y],Z] base_isa [X,[verb,i,non_finite|Y],Z].
[X,[verb,i,participle|Y],Z] base_isa [X,[verb,i,non_finite|Y],Z].
[X,[verb,i,infinitive|Y],Z] base_isa [X,[verb,i,infinitival|Y],Z].
[X,[verb,i,gerund|Y],Z] base_isa [X,[verb,i,infinitival|Y],Z].
[X,[verb,i,active|Y],Z] base_isa [X,[verb,i,participle|Y],Z].
[X,[verb,i,passive|Y],Z] base_isa [X,[verb,i,participle|Y],Z].
[X,[verb,i,pastp|Y],Z] base_isa [X,[verb,i,participle|Y],Z].

```

Nouns

```

% The statements for nouns are of the following schematic type;
% Noun isa [Word,[Type,Definiteness,Countability],AGR].
% Clitics as was discussed in chapter 5 required additional slots for
% restructuring and special Object Agreement. These features were
% not required by other nouns so have not been included in the
% schematization above

```

```

[X,[common_noun|Y],Z] base_isa [X,[noun|Y],Z].
[X,[non_common_noun|Y],Z] base_isa [X,[noun|Y],Z].
[X,[proper|Y],Z] base_isa [X,[non_common_noun|Y],Z].
[X,[determiner|Y],Z] base_isa [X,[non_common_noun|Y],Z].
[X,[pronoun|Y],Z] base_isa [X,[non_common_noun|Y],Z].

```

```

[X,[personal|Y],Z] base_isa [X,[pronoun|Y],Z].
[X,[apersonal|Y],Z] base_isa [X,[pronoun|Y],Z].
[X,[clitic,R,agr([objective|Z])|Y],agr([objective|Z])] base_isa
[X,[apersonal,R,agr([objective|Z])|Y],agr([objective|Z])].

```

Prepositions

```

[X,[prepositional|Y],Z] base_isa [X,[word|Y],Z].
[X,[preposition|Y],Z] base_isa [X,[prepositional|Y],Z].
[X,[clitic|Y],agr([oblique|Z])] base_isa
[X,[prepositional|Y],agr([oblique|Z])].

```

Adjectives

```

[X,[adjective|Y],Z] base_isa [X,[ad_word|Y],Z].
[X,[adverb|Y],Z] base_isa [X,[ad_word|Y],Z].

```

% These statements relate all the major word classes to the Word node

```

[X,[verb,i,finiteness|Y],Z] base_isa [X,[word,i,finiteness|Y],Z].
[X,[noun|Y],Z] base_isa [X,[word|Y],Z].
[X,[ad_word|Y],Z] base_isa [X,[word|Y],Z].
[X,[prepositional|Y],Z] base_isa [X,[word|Y],Z].

```

Some Lexical Entries

% At present there is no morphological component and so the various
% forms of all the words have been put in the lexicon. This will
% change with further developments

% Also lexical entries of this form are required by the parser but
% will not be used in the theory. This is because it is not the case
% that ISA here is used as a subset relation.

```

il base_isa [il,[determiner,def|A],agr([Case,3,singular,m])].
la base_isa [la,[determiner,def|A],agr([Case,3,singular,f])].
lo base_isa [lo,[clitic,r,AGR,def],agr([objective,3,singular,m])].
le base_isa [le,[clitic,r,AGR,def],agr([objective,3,plural,f])].
li base_isa [li,[clitic,r,AGR,def],agr([objective,3,plural,m])].
gli base_isa [gli,[clitic,r,AGR,def],agr([oblique,3,N,G])].
le base_isa [le,[clitic,r,AGR,def],agr([oblique,3,singular,f])].

ragazzo base_isa [ragazzo,[common_noun,X,count],agr([Case,3,singular,m])].
ragazzi base_isa [ragazzi,[common_noun,X,count],agr([Case,3,plural,m])].
ragazza base_isa [ragazza,[common_noun,X,count],agr([Case,3,singular,f])].
ragazze base_isa [ragazze,[common_noun,X,count],agr([Case,3,plural,f])].
dick base_isa [dick,[proper,def],agr([Case,3,singular,m])].
io base_isa [io,[pronoun,def],agr([subjective,1,singular,X])].
mi base_isa [mi,[pronoun,def],agr([objective,1,singular,X])].
tu base_isa [tu,[pronoun,def],agr([subjective,2,singular,X])].

```

ti base_isa [ti, [pronoun, def], agr([objective, 2, singular, X])].
 lei base_isa [lei, [pronoun, def], agr([subjective, 3, singular, f])].

 lui base_isa [lui, [pronoun, def], agr([subjective, 3, singular, m])].
 loro base_isa [loro, [pronoun, def], agr([subjective, 3, plural, X])].
 loro base_isa [loro, [pronoun, def], agr([objective, 3, singular, X])].

 brutto base_isa [brutto, [normal|Y], agr([X, 3, singular, m])].
 brutti base_isa [brutto, [normal|Y], agr([X, 3, plural, m])].
 brutta base_isa [brutto, [normal|Y], agr([X, 3, singular, f])].
 brutte base_isa [brutto, [normal|Y], agr([X, 3, plural, f])].

 a base_isa [a, [preposition, locative], agr([oblique|X])].
 in base_isa [in, [preposition, locative], agr([objective, X, Y, Z])].

 voglio base_isa
 [voglio, [modal, inf_c, indicative_present, X], agr([subjective, 1, singular, Y])].

 ho base_isa
 [ho, [auxiliary, en_c, indicative_present, R, AGR], agr([subjective, 1, singular, X])].

 sono base_isa
 [sono, [auxiliary, en_c, indicative_present, r], agr([subjective, 1, singular, X])].

 odio base_isa
 [odio, [n, t, indicative_present, Y], agr([subjective, 1, singular, X])].

 odio base_isa
 [odio, [n, inf_c, indicative_present, non_r], agr([subjective, 1, singular, X])].

 do base_isa
 [do, [n, di, indicative_present, r], agr([subjective, 1, singular, X])].
 odiare base_isa
 [odiare, [n, inf_c, infinitive, non_r], Y].

 dare base_isa [dare, [n, di, infinitive, r], Y].

 amo base_isa [amo, [n, t, indicative_present, r], agr([subjective, 1, singular, X])].
 ama base_isa [ama, [n, t, indicative_present, r], agr([subjective, 3, singular, X])].
 amare base_isa [amare, [n, t, infinitive, r], Y].

 amate base_isa [amate, [n, t, pastp, r], agr([X, Y, plural, f])].
 amati base_isa [amati, [n, t, pastp, r], agr([X, Y, plural, m])].
 amata base_isa [amata, [n, t, pastp, r], agr([X, Y, singular, f])].
 amato base_isa [amato, [n, t, pastp, r], agr([X, Y, singular, m])].

 date base_isa [date, [n, di, pastp, r], agr([X, Y, plural, f])].
 dati base_isa [dati, [n, di, pastp, r], agr([X, Y, plural, m])].
 data base_isa [data, [n, di, pastp, r], agr([X, Y, singular, f])].
 dato base_isa [dato, [n, di, pastp, r], agr([X, Y, singular, m])].

odiate base_isa [odiate,[n,inf_c,pastp,non_r],agr([X,Y,plural,f])].
 odiati base_isa [odiati,[n,inf_c,pastp,non_r],agr([X,Y,plural,m])].
 odiata base_isa [odiata,[n,inf_c,pastp,non_r],agr([X,Y,singular,f])].
 odiato base_isa [odiato,[n,inf_c,pastp,non_r],agr([X,Y,singular,m])].

volute base_isa [volute,[n,inf_c,pastp,R],agr([X,Y,plural,f])].
 voluti base_isa [voluti,[n,inf_c,pastp,R],agr([X,Y,plural,m])].
 voluta base_isa [voluta,[n,inf_c,pastp,R],agr([X,Y,singular,f])].
 voluto base_isa [voluto,[n,inf_c,pastp,R],agr([X,Y,singular,m])].
 dovute base_isa [dovute,[n,inf_c,pastp,R],agr([X,Y,plural,f])].
 dovuti base_isa [dovuti,[n,inf_c,pastp,R],agr([X,Y,plural,m])].
 dovuta base_isa [dovuta,[n,inf_c,pastp,R],agr([X,Y,singular,f])].
 dovuto base_isa [dovuto,[n,inf_c,pastp,R],agr([X,Y,singular,m])].

% THESE ARE THE BASIC [SYNTACTIC AT THE MOMENT] REQUIREMENTS OF THE WORDS.
 % T isa member of Type
 % S isa member of Subcat
 % F isa member of Finiteness
 % R isa member of Restructuring

[X,[word|Y],Z] base_needs [ob-head].
 [X,[verb,i,tensed|Y],Z] base_needs not([ob-head]).
 [X,[verb,i,tensed|Y],Z] base_needs [op-head].
 [X,[verb,i,tensed|Y],Z] base_needs [op-subject,[A,[non_common_noun|B],Z]].
 [X,[verb,t|Y],Z] base_needs
 [ob-object,[A,[non_common_noun|B],agr([objective|C])]].

[X,[verb,di|Y],Z] base_needs
 [ob-object,[A,[prepositional|B],agr([oblique|C])]].

[X,[auxiliary,en_c,F,R,AGR],Z] base_needs
 [ob-en_complement,[A,[verb,i,pastp,R|B],AGR]].

[X,[verb,i,tensed,r],Z] base_needs [op-pre_clitic,[A,[clitic|B],C]].
 [X,[verb,i,tensed,r,AGR],Z] base_needs [ob-pre_clitic,[A,[clitic,r,AGR|B],C]].
 [X,[verb,i,infinitive,R|Y],Z] base_needs
 [op-enclitic,[A,[clitic,R|B],C]].

[X,[modal,inf_c,F,R],Y] base_needs
 [ob-inf_complement,[A,[verb,i,infinitive,R|Z],agr([objective|B])]].

[X,[n,inf_c,F,R],Y] base_needs
 [ob-inf_complement,[A,[verb,i,infinitive|Z],agr([objective|B])]].

[X,[preposition|Y],Z] base_needs [ob-complement,[A,[non_common_noun|B],Z]].
 [X,[clitic,r,Agr|Y],agr([objective|Z])] base_needs
 [op-pre_clitic,[A,[clitic,r,Agr|B],agr([oblique|C])]].
 [X,[determiner|B],Z] base_needs [ob-complement,[A,[common_noun|B],Z]].


```

        subordinate(InterWord-0,NextWord-M,Newer)),!,
        0 is M - 1.

% The second statement deals with the case where an element
% intervening between a head and a dependent is the head of both.

adjacent(Word-N,NextWord-M,Newer):-
    enumerate(N,0,M,N),
    subordinate(Word-N,InterWord-0,Newer),
    subordinate(NextWord-M,InterWord-0,Newer).

% This is an implementation of the notion of subordinacy.

subordinate(InterWord-0,Word-N,Newer):-
    match(Newer,Dependent-of-Word-N-is-InterWord-0).

subordinate(InterWord-0,Word-N,Newer):-
    match(Newer,Dependent-of-Word-N-is-OtherWord-Z),
    subordinate(InterWord-0,OtherWord-Z,Newer).

```

The Parser

```

% The following is an implementation of the definition of a grammar as
% given in section 3.6. It might be useful to refer to that section
% to get a clearer understanding of the following.

% The following contains the main elements of the program;

parse(S,X) :-
    number(S,NumberedS),
    morphanalall(NumberedS,MorphdS),
    add_constraints(MorphdS,C),
    reduce_constraints(C,N), !,
    solve_constraints(N,X).

% In the above, the predicate parse takes a string of words
% S and will eventually return an analysis X of that
% string. The way it does this is given by the predicates that
% follow it.
% The predicate number, takes the string of words S and
% numbers them. This is the list NumberedS. This is
% equivalent to S of section 3.6
% The predicate morphanalall morphologically-analyses-all the
% numbered words in the string. At the moment this does very little
% work as the morphological analyses are given. MorphS is
% equivalent to S-hat of section 3.6
% The predicate add_constraints inherits all the NEEDS
% statements associated with every element in S. C in
% the above is equivalent to S-vec of section 3.6
% The predicate reduce constraints is used to get rid of

```

```

% negative statements of the form X NEEDS non(Y) This is not a
% necessary step and has no equivalent in section 3.6. It just makes
% tracing the program a little easier.
% The predicate solve_constraints is where all the actual work
% of parsing is done and so it is this that I will be centrally
% concerned to explain. The way in which all the predicates excluding
% solve_constraints are defined is listed below;

number(_, [], []).
number(List, NumberedList):-
    number(1, List, NumberedList).
number(WordNumber, [Word|T], [Word-WordNumber|T2]) :-
    NextWordNumber is WordNumber+1,
    number(NextWordNumber, T, T2).

add_constraints([], []).
add_constraints([H|T], [NH|NT]):-
    addonec(H, NH),
    add_constraints(T, NT).

addonec([], []).
addonec([[Word, [Z|Y], X]-N|T], [Properties|NT]):-
    bagof([Word, [Z|Y], X]-N-needs-Prop, [Word, [Z|Y], X] needs Prop, Properties),
    addonec(T, NT).

reduce_constraints([], []).
reduce_constraints([H|T], [N|T1]):-
    reduce(H, N),
    reduce_constraints(T, T1).

reduce([], []).
reduce([A|B], [C|N]):-
    atleast(A, C),
    reduce(B, N).

atleast([], []).
atleast(X, B):- match(X, Z-not(ZZ)), !,
    remove(Z-not(ZZ), X, A),
    atleast(A, B).

atleast(X, X).

unifymine([], _, []).
unifymine([A|X], Y-P, [A|B]):-
    A = Y-P,
    unifymine(X, Y-Q, B).

% The following is a listing of the predicate solve_constraints

solve_constraints(Y, X):-
    toprove(Y, [], [], X).

% The first part is a shift-reduce parser, with the reduce step done first.

```

```

toprove(Input,Stack1,Found,Final):-
    reducex(Stack1,Stack2,Found,Newfound),
    toprove(Input,Stack2,Newfound,Final).
toprove([W1|Rest],Stack,X,Y):-
    member(W,W1),
    unifymine(W,I-P,WW),
    toprove(Rest,[WW|Stack],X,Y).

% These next two clause are what is called termination clauses, i.e.
% how the program knows when to stop either because it has been
% successful or because it has been unsuccessful
% The first clause says that if there are no more words on the stack
% and you have found all the relations among the words in the string,
% and yet there are still obligatory NEEDS statements around then this
% is not a parse and so fail.

toprove([], [First|Rest], A, A):-
    ((match([First|Rest], [X-N-needs-[ob-Y-Z,W]]),!,fail);
    member(First,FF),
    match(FF,X-N-needs-[ob-Y-Z,W]),
    !,fail.

% Otherwise succeed and write out the results. This is equivalent to
% the definition of Sentence given in section 3.6.

toprove([], [First|Rest], A, A):-
    nl,nl,write(First),
    nl,nl,write(Rest).

% Lets now see how reductions may be made. We have two clauses. These
% are equivalent to the rule $I_{pre}$ and $I_{post}$

% REGULAR HEAD ---> DEP

reducex([Second,First|Rest], [Z|Rest], Found,Newfound2):-
    match(Second,Y-M-needs-[P-F-head,[]]),
    match(First,X-N-needs-[O-post-Dependent,Y]),
    adjacent(X-N,Y-M,Found),
    addlist(Found,[Dependent-of-X-N-is-Y-M],Newfound),
    write([Dependent-of-X-N-is-Y-M]),nl,
    remove(Y-M-needs-[P-F-head,[]],Second,Secondnew),
    remove(X-N-needs-[O-post-Dependent,_],First,Firstnew),
    ((reducex([Secondnew,Firstnew|Rest],[Z|Rest],Newfound,Newfound2));
    addlist(Newfound,[],Newfound2),
    addlist(Secondnew,Firstnew,Z)).

% The first clause says that if the one word NEEDS a head , and some
% other word on the stack NEEDS a post dependent of the form of the word
% that NEEDS a head (denoted by the sharing of the variable Y\ but
% see the definition of match), and the two words are adjacent
% in the sense defined above, and with respect to the list of statements

```

```

% already found, denoted by Found, then add to our list of found
% things the dependency relation between these two words and remove the
% two NEEDS statements.. If there are no more relations to be found
% between these two words that join the lists of their remaining
% requirements if any together.

% The second clause does precisely the same thing but looks for {\it
% pre dependent relations.

% REGULAR DEP <--- HEAD

reducex([Second,First|Rest],[Z|Rest],Found,Newfound2):-
    match(First,X-N-needs-[P-F-head,[]]),
    match(Second,Y-M-needs-[O-pre-Dependent,X]),
    adjacent(X-N,Y-M,Found),
    addlist(Found,[Dependent-of-Y-M-is-X-N],Newfound),
    write([Dependent-of-Y-M-is-X-N]),nl,
    remove(X-N-needs-[P-F-head,[]],First,Firstnew),
    remove(Y-M-needs-[O-pre-Dependent,_],Second,Secondnew),
    ((reducex([Secondnew,Firstnew|Rest],[Z|Rest],Newfound,Newfound2));
    addlist(Newfound,[],Newfound2),
    addlist(Secondnew,Firstnew,Z)).

```

% This then is the current implementation of Word Unification Grammar.
 % A grammar fragment was presented in section 3.6.1. The similarity
 with the grammar presented in chapter 3 should be apparent.

Bibliography

- Ajdukiewicz, K.** (1935) *Syntactic Connexion*, in McCall, S. *Polish Logic: 1920 - 1939*, pp. 207 - 231, OUP, Oxford.
- Anderson, J.** (1977) *On Case Grammar: Prologemena To a Theory of Grammatical Relations*, CROOM.
- Anderson, J.** (1979) *Serialisation, Dependency and The Syntax of Moru*, *Studies in Linguistics* XXXIII, pp. 1 - 23.
- Anderson, J.** (1984) *Case Grammar and The Lexicon*, in *Occasional Papers in Linguistics*, University of Ulster.
- Bach, E. W.** (1979) *Control In Montague Grammar*, *Linguistic Inquiry* 10, pp. 515 - 533.
- Bach, E. W.** (1988) *In Defense of Passive*, *Linguistics and Philosophy* Vol 3.
- Baschung, K. Bes, G. G. Corluy, A. and Guillotin, T.** (1986) *Auxiliaries and Clitics in French UCG Grammar*, Unpublished paper, Department of Linguistics, University of Clermont-Ferrand.
- den Besten, H. and Edmondson, J. A.** (1981) *The Verbal Complex in Continental West Germanic*, in Abraham, W. (ed), *On the Formal Syntax of The West-germania*, Benjamins.
- Bloomfield, L.** (1933) *Language*, Holt.
- Borer, H.** (1981). *Parametric Variation in Clitic Constructions* in Borer, H. (1984) *Parametric Variation*. Foris, Dordrecht.
- Boudon, J.** (1988) *The Uses of Structuralism*, translated by Vaughan, M. Heineman, London.
- Bouma, J.** (1988) *Modifiers and Specifiers*, *Linguistics* Vol. 26, pp. 21 - 46.
- Brachman, R. J.** (1979) *On the Epistemological Status of Semantic Netyworks*, in Findler, N. V. (ed), *Associative Networks and Use of Knowledge By Computers*, New York Academic Press.
- Brachman, R. J.** (1983) *What ISA Is and Isn't*, Fairchild Technical Report No. 638, FLAIR Technical Report No. 15.
- Brachman, R. J.** (1985) *I Lied About The Trees Or, Defaults and Definitions in Knowledge Representation*, *AI Magazine* 80 - 92.

- Brachman, R. J. Fikes, R. E. and Levesque, H. J.** (1983) *KRYPTON: A Functional Approach to Knowledge Representation*, in Proceedings of the 3rd Annual Meeting of the American Association for Artificial Intelligence, Washington, DC.
- Bresnan, J.** (1985) *The Mental Representation of Grammatical Relations*, MIT Press.
- Brunet, J.** (1985) *Grammaire Critique de l'Italien*, Vol. 8, Presses Universitaires de Vincennes, Paris.
- Burzio, L.** (1981) *Intransitive Verbs and Italian Auxiliaries*, PhD Thesis, MIT.
- Burzio, L.** (1986) *Italian Syntax*, D. Reidel, Holland.
- Cassadio, C.** (1987) *Extending Categorical Grammar (An analysis of Word Order and Cliticisation in Italian)*, in Ruzsa, I. and Szabolcsi, A. Proceedings of the 1987 Debrecen Symposium on Logic and Language, Budapest.
- Cinque, G.** (1980) *On Extraction from NP in Italian*, Journal of Italian Linguistics, 5, 47 - 99..
- Chomsky, N.** (1957) *Syntactic Structures*, Mouton, The Hague.
- Chomsky, N.** (1965) *Aspects of The Theory of Syntax*, MIT Press.
- Chomsky, N.** (1968) *Language and Mind*, Harcourt Brace, New York.
- Chomsky, N.** (1975) *Reflections on Language*, Pantheon Books, New York.
- Chomsky, N.** (1980) *Rules and Representations*, Blackwell, Oxford.
- Chomsky, N.** (1981) *Lectures on Government and Binding: The Pisa Lectures*, Foris.
- Chomsky, N.** (1986) *Knowledge of Language: Its Nature Origin and Use*, Praeger, New York.
- Clocksink, W. F. and Mellish, C.** (1981) *Programming in Prolog*, Springer.
- Doyle, J.** (1978) *Truth Maintenance Systems For Problem Solving*, MITAI Technical Report, AD-A054-826.
- Dowty, D.** (1978) *Governed transformations as lexical rules in a Montague Grammar*, in Linguistic Inquiry Vol. 9, pp. 393 - 426.
- Dowty, D.** (1982) *Grammatical Relations and Montague Grammar*, in Jacobson and Pullam (eds), The Nature of Syntactic Representation, Reidel.
- Dowty, D.** (1988) *Type Raising, Functional Composition, and Non-Constituent Conjunction*, in Oehrle, R. Bach, E. and Wheeler, D. (eds), Categorical Grammars and Natural Language Structures, Reidel.
- Dowty, D. R. Wall, R. E. and Peters, S.** (1981) *Introduction to Montague semantics*, Reidel.
- Etherington, D. W. and Reiter, R.** (1983) *On Inheritance Hierarchies With Exceptions*, in Proceedings of the 3rd Annual Meeting of the American Association for Artificial Intelligence, Washington, DC.

- Fikes, R. and Kehler, T.** (1985) *The Role of Frame Based Representation in Reasoning*, CACM, pp. 904 - 920.
- Fillmore, C.** (1968) *The Case for Case*, in Bach and Harms (eds), *Universals in Linguistic Theory*, Holt.
- Flavell, J. H.** (1963) *The Developmental Psychology of Jean Piaget*, VNR, New York.
- Flickinger, D.** (1987) *Lexical Rules in the Hierarchical Lexicon*, PhD Thesis, Department of Linguistics Stanford University.
- Flickinger, D. Pollard, C. and Wasow, T.** (1985) *Structure-Sharing in Lexical Representation*, ACL23, pp. 262 - 267
- Gaifman, H.** (1965) *Dependency Systems and Phrase Structure Systems*, in *Information and Control* 8.
- Gazdar, G. Klein, E. Pullam, G. and Sag, I.** (1985) *Generalised Phrase Structure Grammar*, Blackwell, Oxford.
- Gazdar, G. Pullum, G. K.** (1981) *Subcategorization, Constituent Order And The Notion 'Head'*, in Moortgat, M. van der Hulst, H. and Hoekstra, T. (eds), *The Scope of Lexical Rules*, pp. 107 - 123, Foris.
- Gorayska, B.** (1985) *The Semantics and Pragmatics of English and Polish With Reference to Aspect*, PhD thesis, University College, London.
- Harris, Z.** (1951) *Methods In Structural Linguistics*, University of Chicago Press, Chicago.
- Hays, D. G.** (1964) *Dependency Theory: A Formalism and Some Observations*, *Language* 40, pp. 511 - 525.
- Hayes, P. J.** (1977a) *In Defence of Logic*, IJCAI77, pp. 559 - 565.
- Hayes, P. J.** (1977b) *On Semantic Nets, Frames and Associations*, IJCAI77, pp. 99 - 107.
- Hellwig, P.** (1986) *Dependency Unification Grammar*, COLING, pp. 195 - 198.
- Hendrix, G.** (1979) *Encoding Knowledge in Partitioned Networks*, in Findler, N. V. (ed), *Associative Networks*, New York Academic Press.
- Hoeksema, J.** (1984) *Categorial Morphology*, Unpublished dissertation, University of Groningen.
- Hudson, R. A.** (1976a) *Arguments for A Non-Transformational Grammar*, University Press, Chicago.
- Hudson, R. A.** (1976b) *Conjunction-Reduction, Gapping and Right-Node Raising*, *Language* 52, pp. 535 - 562.
- Hudson, R. A.** (1979) *Panlexicalism*, mimeo.
- Hudson, R. A.** (1980a) *Sociolinguistics*, CUP, Cambridge
- Hudson, R. A.** (1980b) *Constituency and Dependency*, *Linguistics* 18. pp. 179 - 198.

- Hudson, R. A. (1980c) *A Second Attack on Constituency: a Reply to Dahl*, *Linguistics* 18, pp. 489 - 504.
- Hudson, R. A. (1982) *Incomplete Conjuncts*, *Linguistic Inquiry* 13, 547 - 550.
- Hudson, R. A. (1983a) *Word Grammar*, Proceedings of the 13th International Congress of Linguists, The Hague.
- Hudson, R. A. (1983b) *Towards a Cognitive Linguistics*, Working Papers of The London Psycholinguistic Research Group 5, pp. 35 - 47.
- Hudson, R. A. (1984a) *Word Grammar*, Blackwell, Oxford.
- Hudson, R. A. (1984b) *Multiple (alias Parasitic) Gaps*, mimeo.
- Hudson, R. A. (1985a) *Grammatical Relations*, mimeo.
- Hudson, R. A. (1985b) *On the So-called COMP-trace effect*, mimeo.
- Hudson, R. A. (1985c) *A Socially and Psychologically Plausible Theory of Language Structure*, Georgetown University Round Table on Language and Linguistics. Also in Schiffrin, D. (ed), *Meaning, Form and Use in Context: Linguistic Applications*, Georgetown University Press, Washington D. C. pp. 150 - 159.
- Hudson, R. A. (1985d) *Arguments for Events*, mimeo.
- Hudson, R. A. (1985e) *Sociolinguistics in Grammar*, SWoPLaL 2.
- Hudson, R. A. (1985f) *The Limits of Subcategorisation*, in *Linguistic Analysis*, Vol. 15, no.4 pp. 233 - 255.
- Hudson, R. A. (1985g) *Some Basic Assumptions About Linguistic and Non-Linguistic Knowledge*, *Quaderni Di Semantica VI*, pp. 284 - 287.
- Hudson, R. A. (1985h) *"Interrogative Clause" or "Interrogative Verb"*, mimeo.
- Hudson, R. A. (1986a) *Frame Semantics, Frame Linguistics, Frame ...?*, in *Quaderni di Semantica*, Vol. VII, pp. 95 - 111.
- Hudson, R. A. (1986b) *The Comp-Trace and that-trace Effects*, mimeo.
- Hudson, R. A. (1986c) *Sociolinguistics and The Theory of Grammar*, *Linguistics* 24, pp. 1053 - 1078.
- Hudson, R. A. (1986d) *A Prolog Implementation of Word Grammar*, *Speech Hearing and Language: Work in Progress 2*, Dept. Of Phonetics and Linguistics, UCL, London. pp. 133 -150.
- Hudson, R. A. (1986e) *Phonology in Word Grammar*, mimeo.
- Hudson, R. A. (1986f) *Word Grammar and Relevance Theory*, mimeo.
- Hudson, R. A. (1986g) *The Comp-trace Effect and Grammatical Relations*, mimeo.
- Hudson, R. A. (1987a) *Zwicky On Heads*, *Journal of Linguistics* 23, pp. 109 - 132.
- Hudson, R. A. (1987b) *Recent Developments in Dependency Theory*, in Jacobs, J. et al. (eds), *Syntax. An International Handbook of Contemporary Research*. Walter de Gruyter.

- Hudson, R. A.** (1987c) *Grammar, Society, and The Pronoun*, in Steele, R. and Treadgold, T. (eds), *Language Topics: Essays in Honour of Michael Halliday*, Amsterdam, Benjamins, pp. 493 - 505.
- Hudson, R. A.** (1987d) *ISA versus IS (A)*, mimeo.
- Hudson, R. A.** (1988a) *Coordination and Grammatical Relations*, *Journal of Linguistics*, 24.
- Hudson, R. A.** (1988b) *The English Passive*, distributed by LAUD.
- Hudson, R. A.** (1988c) *Extraction and Grammatical Relations*, University of Thessaloniki: Working Papers in Linguistics, *Lingua* 76, pp. 177 - 208.
- Hudson, R. A.** (1988d) *Identifying the Linguistic Foundations For Lexical Research and Dictionary Design*, in Walker, D. et al. (eds), *Proceedings of The Pisa Workshop on the Automated Lexicon*. Also in *International Journal of Lexicography* 1.3.
- Hudson, R. A.** (1989) *Gapping and Grammatical Relations*, *Journal of Linguistics*.
- Hudson, R. A.** (1991 forthcoming) *A Word Grammar of English*, Oxford, Blackwell.
- Hudson, R. A. and Langendonck, W. van** (1987) *Word Grammar*, in Droste, F. (ed), *Mainstreams in Linguistics*, Amsterdam and New York: Benjamins.
- Israel, D. J.** (1980) *What's wrong with Non-Monotonic Logic?*, in AAAI-80.
- Jackendoff, R. S.** (1977) *X-Bar Syntax: A Study of Phrase Structure*, MIT Press.
- Jackendoff, R. S.** (1987) *The Status of Thematic Relations in Linguistic Theory*, *Linguistic Inquiry* Vol. 18, pp. 369 - 411.
- Jaeggli** (1982). *Topics in Romance Syntax*, Foris, Dordrecht.
- Kay, M.** (1979) *Functional Grammar*, in *Proceedings of The Fifth Annual Meeting of The Berkeley Linguistic Society*.
- Kay, M.** (1985a) *Unification Grammar*, in Dahl, V. and Saint-Dizier, P. (eds), *Natural Language Understanding and Logic Programming*, Northholt.
- Kay, M.** (1985b) *Parsing in Functional Unification Grammar*, in Dowty, D. R. Karttunen, L. and Zwicky, A. M. (eds), *Natural Language Parsing: Psychological, Computational and Theoretical Perspectives*, CUP.
- Klavans, J. L.** (1979) *On Clitics As Words*, *CLS* 15.
- Klavans, J. L.** (1985) *The Independence of Syntax and Phonology in Cliticisation*, *Language*, Vol. 61, no. 1,
- Knight, K** (1989) *Unification: A Multidisciplinary Survey*, *ACN Computing Surveys*, Vol. 21, No. 1.
- Kunze, J.** (1987) *Some Remarks on Case Relations*, *EACL3* pp. 302 - 305.
- Lambek, J.** (1961) *On the Calculus of Syntactic Types*, in *Structure of Language and Its Mathematical Aspects: Proceedings of the Symposia in Applied Mathematics XII*, pp. 166 - 178, Providence, Rhode Island.

- Manzin, M-R.** (1983) *Restructuring and Reanalysis*, PhD Thesis, MIT.
- McDermott, D. and Doyle, J.** (1980) *Non-Monotonic Logic I*, Artificial Intelligence 13, 41 - 72.
- Mel'ciuk, L.** (1988) *Dependency Syntax: Theory and Practice*, Albany, New York. SUNY.
- Minsky, M.** (1975) *A Framework for Representing Knowledge*, in Winston, P. H. (ed), *The Psychology of Computer Vision*, New York: McGraw Hill.
- Moore, R.** (1983) *Semantic Considerations in Non-Monotonic Logic*, SRI Technical Report.
- Moortgat, M.** (1987) *Lambek Theorem Proving*, in Klein, E. and van Benthem, J. (eds), *Categories, Polymorphism and Unification*, pp. 169 - 200, Edinburgh and Amsterdam.
- Panini,** (c.forth century B.C.)
- Piaget, J** (1926) *The Language and Thought of the Child*, RKP.
- Piaget, J** (1953) *The Origins of Intelligence in Children*, RKP.
- Piaget, J** (1964) *Six Etudes de Psychologie*, Editions Gonthier S.A. Generalève English Translation *Six Psychological Studies* (1980), The Harvester Press.
- Piaget, J** (1970) *Genetic Epistemology*, COLUP
- Piaget, J** (1971) *Structuralism*, RKP.
- Piattelli-Palmarini, M. (ed)** (1979) *Language and Learning: The Debate between Jean Piaget and Noam Chomsky*, HUP.
- Pollard, C.** (1985) *Lectures on Head-Driven Phrase Structure Grammar*, Unpublished Lecture Notes, CSLI, Stanford University.
- Pollard, C. and Sag, I.** (1987) *Information-Based Syntax and Semantics*, CSLI No. 13.
- Quillian, M. R.** (1968) *Semantic Memory*, in Minsky, M. (ed), *Semantic Information Processing*, MIT Press.
- Reiter, R.** (1980) *A Logic For Default Reasoning*, AI Vol 13, pp. 81 - 132.
- Reiter, R. and Criuscolo, G.** (1983) *Some Representational Issues In Default Reasoning*, in Cercone, N. (ed), *Computational Linguistics*, Pergamon.
- Rizzi, L.** (1978) *A Restructuring Rule in Italian Syntax*, reprinted in Rizzi, L. (1982) *Issues in Italian Syntax*, Foris, Dordrecht.
- Robinson, J.** (1970) *Dependency Structures and Transformational Rules*, Language 46.
- Robinson, J.** (1965) *logic prog, .*
- de Saussure. F.** (1915) *Course de Linguistique General*, Edited by Bally, C. and Sechehaye, A. Geneva.

- Schachter, P.** (1974) *Constraints on Clitic Order in Tagalog*, UCLA Papers in Syntax, 5. 97 - 118.
- Seligman, J. M.** (1990) *Perspectives: A Relativistic Approach To The Theory of Information*, PhD Tesis.
- Sgall, P. and Hajicova, E.** (1986) *The Ordering Principle*, PBML 45, pp. 1 - 26.
- Sgall, P. Hajicova, E. and Panevova, J.** (1986) *The Meaning of The Sentence In Its Semantic and Pragmatic Aspects*, Research Report, Prague.
- Shieber, S. M.** (1986) *An Introduction to Unification-Based Approaches to Grammar*, CSLI Lecture Note Series.
- Shieber, S. M.** (1987) *Evidence Against The Context-Freeness of Natural Language*, in Savitch et al. (eds), *The Formal Complexity of Natural Language*.
- Starosta, S.** (1982) *Case Relations, Perspective and Patient Centrality*, University of Hawaii Working Papers in Linguistics 14, pp. 1 - 33.
- Starosta, S.** (1988) *The Case For Lexicase*, Pinter.
- Steedman, M.** (1982) *On The Generality of the Nested Dependency Constraint and the Reason for an Exception in Dutch*, Linguistics 21, pp. 35-66.
- Steedman, M.** (1985) *Dependency and Coordination in the Grammar of Dutch and English*, Language Vol. 61 No. 3, pp. 523 - 568.
- Steedman, M. J.** (1986) *Combinatory Grammars and Parasitic Gaps*, Unpublished paper, Centre for Cognitive Science, University of Edinburgh.
- Steedman, M.** (1988) *Combinators and Grammars*, in Oehrle, R. Bach, E. and Wheeler, D. (eds), *Categorial Grammars and Natural Language Structures*, Reidel.
- Sterling, L. Shapiro, E.** (1986) *The Art of Prolog*, MIT Press.
- Tesnière, L.** (1959) *Elements de Syntaxe Structurale*, Paris, Klincksieck
- Touretzky, D. S.** (1986) *The Mathematics of Inheritance Systems*, New York: Morgan Kaufman.
- Turner, R.** (1984) *Logics For Artificial Intelligence*, ELLIS.
- Uzkoreit, H.** (1986) *Categorial Unification Grammars*, CSLI.
- Uzkoreit, H.** (1987) *Linear Precedence in Discontinuous Constituents: Complex Fronting in German*, in Huck, G. J. and Ojeda, A. E. (eds), *Syntax and Semantics Vol. 20: Discontinuous Constituents*.
- Valkonen, K. Jappinen, H. Lehtole, A. and Ylilammi, M.** (1987) *Declarative Model For Dependency Parsing - a view into blackboard methodology*, EACL3, pp. 218 - 225.
- Volino, M. S.** (1986) *On The Underlying Representations of Morphophonemic Alternants*, mimeo.
- Volino, M. S.** (1986) *An Aspect of Italian Syntax: Restructuring, The Projection Principle, and the Nature of Subcategorisation*, mimeo.

- Volino, M. S. and Calder, J.** (1990) *Inheritance and Defaults in Linguistics*, Internal DYANA report.
- Woods, W. A.** (1975) *What's In A Link: Foundations of Semantic Networks*, in Bobrow, D. G. and Collins, A. M. *Representation and Understanding*, Academic Press, New York.
- Zeevat, H. Klein, E. and Calder, J.** (1987) *An Introduction to Unification Categorical Grammar*, in Edinburgh Working Papers in Cognitive Science.
- Zubizaretta, M-L.** (1982) *On the Relationship of The Lexicon to Syntax* PhD Thesis, MIT.
- Zubizaretta, M-L.** (1985) *The Relation Between Morphophonology And Morphosyntax: The Case Of Romance Causatives*, *Linguistic Inquiry* Vol. 16, pp. 247 - 289.
- Zwicky, A.** (1977) *On Clitics*, Indiana University Linguistics Club.
- Zwicky, A.** (1985a) *Clitics and Particles*, *Language* Vol. 61, no.2.
- Zwicky, A.** (1985b) *Heads*, *Journal of Linguistics* 21, pp. 1 - 29.
- Zwicky, A. and Pullam, G.** (1983) *Cliticisation vs Inflection: English N'T.*, *Language*, Vol. 59, no. 3