Dependency, licensing and the nature of grammatical relations^{*}

CHRISTIAN KREPS

Abstract

Word Grammar (Hudson 1984, 1990), in common with other dependency-based theories of syntax, seeks to express syntactic knowledge in terms of direct relationships between words. Dependency theories have traditionally recognised a large class of primitive grammatical relations (GRs), including Subject, Object and Adjunct, amongst others. In this paper I call into question the status of these GRs, and argue instead that it is possible, and preferable, for a dependency theory to be based on just one type of syntactic relation — licensing. I suggest that residual properties of individual GRs can then be derived in a principled way from the simple structures composed of licensing relations.

1 Introduction

Word Grammar (WG) (Hudson 1984, 1990) belongs to the tradition of dependency theory, which finds its modern roots in the work of Tesnière (1959). In common with other versions of dependency grammar, WG seeks to express syntactic knowledge in terms of direct relations between words; these relations, or dependencies, are theoretical primitives and are not derived from any form of alternative structure. However, WG, like other theories, has traditionally recognised a large class of dependencies, known collectively as Grammatical Relations (GRs). This set of primitive GRs includes some well-known relations such as subject, object, indirect object and adjunct as well as more theory-particular relations such as visitor, extraposee and x-complement. It is has generally been taken for granted that any relational theory of syntax will have to recognise a similar set of primitive GRs.

^{*}I am grateful to Dick Hudson for his guidance and encouragement throughout my work. Thanks also to the participants of the Word Grammar Interest Group for their comments on some of the ideas raised in this paper.

In this paper I will argue that the subclassification of the dependency relation, and the consequent recognition of a class of primitive GRs, is both unnecessary and undesirable. Instead I will suggest that it is possible, and indeed preferable, for a dependency theory to recognise just one primitive syntactic relation and to derive residual properties of GRs such as subject and object from the monostratal syntactic structure involving this single relation. Although this is broadly equivalent to the approach of phrase structure grammar, as far as I am aware no such proposal has ever previously been advanced for a theory of dependency.

A simplified grammar utilising just one relation will be argued to enjoy two main advantages over more traditional versions of dependency theory. Firstly, I believe that the abolition of primitive GRs is desirable for reasons of elegance and overall parsimony; recognition of distinct GRs can lead to a proliferation of relations which, in many cases, are poorly-defined and open to charges of unconstrainedness. Secondly, I will argue that a 'monorelational' theory of syntax, supplemented with a system of derived semantic relations is actually capable of providing a more elegant and unified account of a broader range of data than currently offered by other theories. Specifically I will draw attention to facts relating to the interpretation of subjects in non-finite complement clauses.

Section 2 will offer a brief review of WG and its relation to phrase structure grammar, while section 3 outlines various problems associated with WG's recognition of distinct GRs. In section 4 I will explore an alternative theory of dependency based on just one syntactic relation, and I will briefly examine how this relation may also play a part in the morphological structure of words. Section 5 will then supplement this outline theory of syntactic dependency with a system of derived semantic relations. It should be pointed out here that this paper represents a collocation of various ideas explored more fully in my forthcoming thesis. Inevitably some discussion and argumentation have at times been sacrificed here for the sake of brevity.

2 Outline of Word Grammar

2.1 Dependency and constituency

One way of expressing a syntactic relationship between two words is by means of constituency. In its simplest form this allows two words, X and Y, to be linked by their shared participation in a phrasal constituent:

(1) $[_{XP} X Y] [_{YP} X Y]$

These phrasal constituents can then be embedded within one other in various ways, giving rise to the well-known types of branching configurational structure illustrated by the labelled brackets in (2):

Geometrical properties of this configurational structure can then serve as the basis for defining and distinguishing different syntactic relations between words. For example, complements have traditionally been described as sisters of a head X, whereas subjects are defined as the sister of an intermediate X' projection¹.

Endocentric constituent structure of the type illustrated in (1) and (2) is an integral part of many current syntactic theories, notably Principles and Parameters theory (Chomsky 1981, 1986) and its recent Minimalist extensions (Chomsky 1992, 1995)². Much of the machinery of these theories has been geared around phrase structure and the putative constituency obtaining between any two words that enter into a syntactic relationship. Indeed, similar patterns are even taken to extend beyond word-word relations; not only do we encounter phrases headed by 'functional' elements, such as C and AGR, we also find a similar system of phrase structure generalised to the internal morphological structure of words (Selkirk 1982).

This, however, is not the only way in which words can be brought together. A possibly simpler approach might be to take the syntactic relationships which link words together as basic, and not as being derived from or mediated by abstract configurational structure. In this way word X and word Y could be linked directly by a simple relation, R, without participating in any form of larger phrasal constituent. The relation R itself will be a primitive entity of the grammar, and any relationship between X and Y will be expressed solely in terms of R:

¹Phrase structure can also be understood in terms of set membership; the participation of X and Y in a phrasal constituent such as XP is equivalent to their membership of a set $\{X | Y\}$. This is a conception of Phrase Structure explored recently by Chomsky (1993).

²Constituency also plays an important part in other theories of syntax, such as GPSG (Gazdar et al. 1985) and HPSG (Pollard and Sag 1987, 1994).

$$(3) \qquad \qquad \boxed{R - } R - \\ X \qquad \qquad Y^3$$

This, in a nutshell, is the view of syntax taken by Dependency Grammar, notably Word Grammar (WG), (Hudson 1984, 1990); WG syntactic representations consist solely of word strings with individual pairs of words linked together by binary relations known as dependencies. (4) below shows an example, with each dependency represented by an arrow:

The arrows serve to express both the location and the direction of dependencies. Thus we see a direct syntactic relation between 'Ted' and 'crashed', for example, but not between, 'Ted' and 'Bentley'. Arrows point uniformly from head words to their dependents. Thus the arrow pointing from 'crashed' to 'Ted' expresses the fact that the later is a dependent of the former.

This element of asymmetry or inequality is a fundamental characteristic common to all conceptions of dependency, both linguistic and non-linguistic. Thus beyond the realm of language, we can describe Gibraltar as a dependency of the United Kingdom, but not vice versa. Dependency, no less in its linguistic sense than in its more general usage, incorporates the idea of one element in the relationship being 'more important' and in some sense controlling the other. Within a syntactic dependency relation this asymmetry is instantiated between a head word and a dependent word. The head is described as the governing element to which the dependent is subordinated, in a sense to be discussed later.

Since dependencies do not have to be expressed in terms of phrase structure, WG need not recognise any grammatical entity larger than the word, and consequently all syntactic knowledge can be expressed exclusively in terms of words and the dependencies which serve to link them (although an exception is made for co-ordination (Hudson 1988a)).

³The question naturally arises as to whether (1) and (3) are substantially different; how is the participation of two words in a phrasal constituent different from their being linked by a relation R? One obvious difference is that a phrase such as XP can itself enter into a syntactic relation with another element, whereas a relation R cannot. See Hudson (1984 ch. 3) for a discussion of further differences.

Generally, then, a well-formed sentence will be uniquely associated with a single wellformed network of dependency relations; WG is a monostratal theory and thus seeks no recourse to underlying structure or processes of syntactic derivation. A network of dependencies is judged to be well-formed if it obeys certain constraints, two of the most important of which are known colloquially as 'no tangling' and 'no dangling' (Hudson 1994). The former constraint, also known as projectivity, disallows relations to intersect one another, thus imposing a degree of locality on dependencies. This serves to rule out examples such as (5) below:



Here the relation between 'his' and 'Bentley' is insufficiently local in that another word, 'yesterday', intervenes which bears no relation to either of them. The second constraint, 'no dangling', simply requires that every word in a single well-formed structure be linked to at least one other word; no word can remain unconnected to another, hence the ungrammaticality of (6) where the word 'lychee' dangles:

The 'no-dangling' constraint guarantees that all words in a structure will depend on another word in the same structure. The one exception to this is the root word where chains of dependency originate and which itself depends on nothing. Each structure will have one root word, signalled by an unconnected downward-pointing arrow. In most cases the root word will be the matrix finite verb, as with 'crashed' in (4). I will say more about this later.

Relations between syntactic heads and dependents can be discerned in phrase structure too where, in simple terms, a word X constitutes the head of its projection (XP). Returning to (2b) above, YP and ZP could be described as dependents of X; in some sense both YP and ZP are subordinated to X by virtue of the fact that they occur embedded within X's projection. In this way a degree of asymmetric dependency could

be said to obtain between any two elements which constitute a phrasal constituent⁴. Once again, though, asymmetry is here derived from geometrical attributes of the phrase structure configuration whereas in Dependency terms it is stated as a primitive property of syntactic relations.

In phrase structure theories facts pertaining to word order are expressed in terms of the head parameter (Chomsky 1981). This states a generalised ordering pattern between heads and their complements (ZP in (2b)), potentially subsuming a wide range of word order phenomena. Languages are predicted to be uniformly head-first or head-last. The head parameter is also recognised in WG where a similar generalisation can be expressed about the direction of dependencies. Thus while in some languages, like Welsh, nearly all dependencies are head-first, in others, notably Japanese and Turkish, they are usually head-last. Of course, some languages are better-behaved than others with respect to the head parameter. Thus in English, although heads generally precede their dependents (see (4) above), certain dependencies do not conform to this pattern, the most common exception being the subject which invariably precedes its head.

2.2 Dependencies and GRs

In example (4) the verb 'crashed' is shown to have three dependents, which we might informally label as subject — 'Ted', object — 'his Bentley', and adjunct — 'yesterday'. As I pointed out before, in WG these Grammatical Relations, or GRs, are recognised as syntactic primitives of the theory, and thus each dependency relation must be labelled according to which of these more specific GRs it instantiates (s=subject, o=object, a=adjunct, c=complement):



While 'crashed' constitutes the root of the sentence and doesn't depend on anything in (7), in (8) the same word depends on another verb, 'know', which is itself the root of the structure:

⁴Brody (1994) advances a proposal along these lines according to which dependencies are taken to exist alongside constituent structure.



Here the embedded clause, headed by the verb 'crashed', bears the object relation to the matrix verb 'know' in just the same way as 'his Bentley', headed by 'his', is the object of 'crashed'.

Each GR is described as a distinct subtype of the dependency relation, and as such, each is associated with its own particular set of properties. Thus GRs may differ from one another with respect to their distribution, direction or associated morphological marking, amongst other things. For example, while certain verbs may have a single object dependent, which will usually occur immediately after its head, any verb can potentially have any number of adjuncts, the position of which is somewhat less restricted. The question of how many separate GRs need to be recognised in an adequate theory of grammar is a controversial one. Most relational theories recognise a class of GRs which includes at least subject, object and indirect object along with various types of adjunct⁵. These relations will be familiar since they play a part, at least informally, in most other theories. The only point of contention concerns their primacy or otherwise.

In addition to these fairly well-known GRs, WG also recognises a number of more theory-particular relations including, for example, 'extraposee' and 'visitor', instantiations of which mediate between a displaced dependent and its non-local head (Hudson 1988b). I will not discuss these relations further here except to point out that my arguments against GRs later apply equally to them. Another GR particular to WG is the x-complement. Informally, x-complements could be described as complement verbs which share their subject with their head, as illustrated in (9):



⁵See, for example, Perlmutter (1983) and Blake (1990).

Although both these representations are somewhat simplified, they serve to illustrate how the x-complement relation (labelled x) allows the subject to be shared between more than one verb. WG is unusual in this respect since most other dependency theories impose maximum of one head per word (Robinson 1970). However, the sharing analysis sanctioned by the x-complement relation is valuable in that it yields a unified analysis of so-called 'raising' constructions (9a) and 'control phenomena' (9b). This is particularly useful since WG recognises neither syntactic movement nor empty categories of any kind and consequently analyses involving subject raising or PRO are inadmissible. Note, though, that the x-complement relation will have to be embellished somewhat if it is to account for cases of 'object control' (Hudson 1990 ch. 10). So too a subject can also be shared between an adjunct and its head, a potentially problematic situation for WG since adjuncts cannot be classed as x-complements. I will return to these issues in section 5.

3 Some problems with GRs

It has generally been assumed that all relational theories of grammar will have to recognise a set of distinct GRs. This is not, however, a logical necessity, and later I will argue that it is possible for a theory of dependency to be based on just one syntactic relation which does not have to be subclassified into more specific GRs. A grammar which utilises only one type of dependency relation would evidently be maximally simple and economical, minimalist even. For one thing, problematic questions as to how many GRs have to be recognised would evidently not arise. More importantly, such a theory, if viable, would be less vulnerable to charges of unconstrainedness, one of the criticisms most frequently levelled at dependency grammar.

Essentially the problem of constrainedness arises from the fact that dependency theories, including WG, do not incorporate any inherent limitation on the nature or location of syntactic relationships. One result of this is that theoretically a dependency relation could be postulated anywhere. Say, for example, that we come across a previously unknown phenomenon in a language which seems to suggest that a direct syntactic relation may exist between the subject and the indirect object of a verb. Although unlikely, there is nothing to prevent us from suggesting a relation of some sort between them; there is no theoretical constraint forcing us to seek an alternative, possibly more principled account of the facts. Ultimately, this can lead to a data-driven approach to language which makes no real predictions and, at its worst excess, hardly qualifies as a theory at all. To recast the issue in more Chomskyan terms, some theories of dependency

could be said to attain observational adequacy too easily at the expense of explanatory adequacy (Chomsky 1965).

The absence of inherent constraints on the location of dependency relations is matched by an equivalent freedom concerning the nature of specific GRs which can be recognised. As I said before, the class of GRs is an open one which can be expanded as and when required. Which GRs are recognised in a particular theory will be determined more by the data, or even the personal preference of the linguist, than by inherent principles of dependency. To take just one example, WG recognises an x-complement relation, where a complement verb shares its subject with its head. There is nothing in the theory, however, to rule out the existence of a y-complement relation, which allows, say, for the sharing of an object or adjunct. The fact that such a relation does not exist is essentially because the data do not require it. Once again, then, it might be difficult to formulate coherent predictions about language within such a theory when another GR could be just around the corner.

Here, then, we see a potential advantage of phrase structure which, though more complex, does impose a degree of inherent discipline on syntactic relations, a discipline ultimately arising from the geometry of the constituent configuration. Of course some might question the wisdom of shifting the governing criteria of a linguistic theory from language data to issues of geometry. Certainly in many ways the confines of constituency are too rigid and constraining for natural language, which is why phrase structure theories have to be supplemented in one way or another, either by a transformational component (Chomsky 1957), or a highly enriched feature system (Gazdar et al. 1985, Pollard and Sag 1994). Both of these addenda bring with them their own areas of weakness and complexity.

The recognition of a class of primitive GRs has another unwelcome consequence in that it entails an otherwise unnecessary increase in the size of the grammar. As I pointed out, each GR is distinct and has its own specific properties. Not all of these properties can be derived from more general principles of dependency, and thus each GR will be associated with a range of specific facts which, presumably, will have to be stored somewhere in the grammar. The obvious implication of this is that the grammar will have to expand in proportion to the number of GRs that are recognised. This is particularly relevant in the case of WG where facts about language are stored as individual rules which take the form of propositions. According to this declarative conception of grammar, the total set of

propositions relating to a language will constitute the grammar of that language⁶. It is not hard to see, then, how the set of propositions corresponding to a grammar will have to be enlarged in order to accommodate facts pertaining to each GR that is recognised by the theory.

Naturally, a dependency grammar based on just one relation would only need to store properties of this single relation. This would entail a significant simplification and reduction in the size of the grammar. In fact, as I will argue later, apart from very general facts such as a head parameter setting, a 'monorelational' syntax can essentially reduce to a lexicon specifying the combinatorial properties of words, a property reminiscent of Categorial Grammars (Ades and Steedman 1982, Wood 1993). In this way it is possible to increase the proportion of grammatical knowledge stored as lexical information, very much in line with WG's own intentions (Hudson 1990 ch. 1). The question of the grammar's size, of course, is not just a matter of theoretical elegance and parsimony; the issue will inevitably have a direct bearing on questions of learnability and parsing.

There are other problems associated more particularly with WG's treatment of GRs and lexical valency which I will not go into here for the sake of brevity. It is important to remember, however, that none of the problems raised in this section are inherent to dependency theory itself, but for the most part stem from the recognition of distinct GRs. In section 4 I will explore an alternative version of dependency grammar which utilises just one type of relation. In this way I hope it will be possible to avoid the difficulties outlined above.

4 Beyond grammatical relations

4.1 Introduction

What follows is an exploration of a dependency theory that makes use of only one structural relation, based on the notion of licensing. I will argue that this relation may offer a coherent and principled account of syntax driven, for the most part, by the valency specification of individual words. Many of the suggestions I will make differ in significant

⁶An account of linguistic knowledge along these lines may initially appear to be excessively stipulative. However, WG incorporates a sophisticated mechanism of inheritance and overriding which allows propositions to be stated economically at exactly the right level of generality (for a full account of this system see Fraser and Hudson 1992).

ways from standard assumptions of WG, nevertheless the fundamental syntactic tenets of WG will be retained; the proposed account is a monostratal theory of dependency and does not recognise syntactic derivation, empty categories or constituency of any kind above the level of the word.

4.2 Dependency as licensing

It is perhaps surprising that one question that has been very seldom raised is what the dependency relation might really be. What does it mean to say that one word depends on another? What, in other words, is the precise content of the relation between X and Y in (10)?

(10) $\xrightarrow{X Y}$

While most grammarians would agree that the concept of dependency embodies an element of asymmetry or inequality, this is clearly insufficient as a basis for syntactic relations; a degree of inequality can be discerned between virtually any two words and consequently the concept is too nebulous to be of serious use. Other than this, however, it is difficult to discern any precise, widely-accepted content to dependency from the relevant literature⁷. While grammarians have sometimes cited properties which characterise dependency relations (Hudson 1990 ch. 6), more often than not these properties are artefactual, and offer little or no insight into the actual content of the relation itself. One possibility, however, raised recently by Hudson (pc) is that dependency might amount to some form of contingency relation, whereby the occurrence of a dependent is sanctioned by the presence of the head. This is an idea which I will explore further.

In my opinion this absence of a widely-accepted and precise content to the dependency relation has had a detrimental effect on the reputation of WG and other theories. A common perception is that dependency is little more than a notational variant of other

⁷Indicative of this, perhaps, is a recent proposal by Rosta (1994) according to which dependency structure amounts to little more than a well-formedness constraint imposed upon a basic system of (more contentful) grammatical relations.

representational systems. This view, though erroneous⁸, is not entirely unjustified, given that grammarians have generally failed to invest the dependency relation with sufficient autonomous content to set it apart from other systems. Moreover, this lack of definable content to the relation is also partly responsible for the excessive power and consequent unconstrainedness sometimes associated with dependency grammars, discussed in the previous section; if it is possible to postulate the existence of a dependency anywhere, this is surely because the relation itself is too 'invisible'.

What is required, then, is a syntactic relation based on a coherent and meaningful content, over and above the rather nebulous element of asymmetry. A relation with this sort of independent meaning could then serve as the basis of a more principled, constrained theory of syntax which imposed inherent limitations on the number and location of dependencies. For one thing a better-defined content would make the dependency relation more 'visible', and thus its supposed existence in a given context would be more open to scrutiny. Hopefully this content would also serve to constrain the number of more specific GRs that can be recognised.

One possible starting point is Hudson's idea of dependency as a contingency relation. According to this view the dependency relation essentially boils down to the contingency of the dependent's existence on the presence of the head. Thus with reference to (10) above, the relation between X and Y expresses the fact that X's presence sanctions the existence of Y. This rather abstract idea of existential contingency can be translated directly into the simpler, more user-friendly concept of licensing. Quite simply we can interpret the relation in (10) as expressing the fact that the head X licenses the dependent Y. This can then be generalised to all cases, and each syntactic dependency will thus have to embody a licensing relation between the two participating words; if one word does not licence the occurrence of another, then there can be no syntactic relationship between these words (note, though, that this says nothing as to whether a semantic relation exists between them). From now on I will use the terms 'licensing' and 'dependency' interchangeably. So too the terms 'head' and 'licenser' will be used synonymously as will 'dependent' and 'licensee'.

The concept of licensing invests the dependency relation with a specific and intuitively plausible content, and one which effectively captures, I think, a requisite degree of asymmetry. A dependent is only present in a structure by virtue of a licensing head; this leaves no ambiguity as to the direction of the asymmetry between them. So too licensing

⁸Hudson (1984 ch. 3) outlines various arguments against the supposed equivalence between dependency and constituency. See also footnote 3.

also has the advantage of combining this with a relatively well-defined content which is common to many theories. For example, Case-assignment in Principles and Parameters theory is often described as a licensing procedure by which the overt occurrence of a NP is sanctioned (Chomsky 1986). So too, I suspect that the same concept is also broadly equivalent to much of what is both implicitly and explicitly assumed in Word Grammar. To take just one example, a grammatical proposition stating that word X has a complement is in some sense equivalent to saying that X licenses a dependent.

One thing that should be apparent about the licensing relation is that it does not seem to be amenable to the type of subclassification that characterises more conventional versions of dependency. Either X licenses Y or it doesn't; there is little scope here for variation or the recognition of distinct subtypes. As I suggested before, this restriction arises from the more specific content of licensing-based dependency. The fact that X licenses, say, a 'subject' rather than an 'object' would thus have nothing to do with the licensing relation itself, and it would appear to make little sense to try and differentiate between distinct, labelled dependencies. If this is true then evidently the very real distinctions between subjects, objects and adjuncts will have to be stated in some other way.

4.3 Licensing structure

Licensing is a criterion by which the occurrence of a word is sanctioned. Assuming that all words in a well-formed structure are subject to the same sanctioning requirement, we can then infer that all words will need a licenser; any word without a head should simply not appear. This very general requirement can be represented schematically as in (11), where W is a variable over all words, and the downward pointing arrow represents its dependence on something else.

(11)

The generalisation expressed in (11) is true of all words and thus it does not have to be stated as a property of specific lexical items⁹. Nevertheless, for reasons of clarity I will continue to represent individual words' requirement of a licenser.

Given that licensing constitutes the central dynamic underpinning our system of dependency, we can derive in the simplest possible way the 'no-dangling' requirement of WG. Any word that remains unconnected to another word will evidently not be licensed by a head, and thus should not appear; the ungrammaticality of (6) above, for example, can be attributed to the fact that 'lychee' bears no syntactic relation to anything else and thus is not licensed. Here too we can discern a degree of testability which distinguishes the licensing relation from other versions of dependency. To discover the head of X we need only find the word which sanctions X's presence, or, to put it another way, the word which, if removed from the structure, would render X's presence infelicitous. There will, inevitably, be some cases of ambiguity, but in general a dependency will be more visible and easier to define than before.

Whereas all words need a licenser, some, but not all words will themselves license dependents. This class of licensing words, whose very presence explicitly sanctions the occurrence of another, will include transitive verbs and prepositions as well as complementisers and determiners. In this way the property of licensing another word could be said to be equivalent to selecting a complement. Since, however, only some words are classed as licensers, this property cannot be captured by any universal generalisation of the type in (11). Instead a word's capacity to license another will have to be stated as an individual combinatorial requirement, analogous to a subcategorisation frame. Thus the preposition 'with', for example, licences another word (a property represented by the upward pointing arrow) whereas the noun 'gin' doesn't. (Both these words, of course, will have to be licensed themselves):

(12) **X N** with gin

Generally a word's requirement that it be licensed will be satisfied by another word which is capable of licensing a dependent. Conversely the capacity of these same words to license others will be fulfilled by the presence of other words which, by default, will need a head. In this way a well-formed syntactic representation will consist of one or

⁹In more overtly WG terminology (11) corresponds to a general proposition stating that all words must have a head (Hudson 1990 ch. 10).

more chains of dependency with each word being licensed by another. Evidently there must be a point from which these chains originate, since they cannot be of indefinite length. Each structure will thus have to contain one word which, while licensing one or more dependents, will not itself have a head — the equivalent of the root word in WG. While this would appear to contradict the generalisation expressed in (11), this principle made no claim as to what is responsible for licensing. Although words will generally be licensed syntactically by other words, let's assume that in any structure one (and only one) word must be licensed non-syntactically, perhaps by a desire on the part of the speaker to communicate. This 'root word' will usually be a finite verb, but there is no reason why any word can't be subject to this type of non-syntactic licensing:



There is nothing to suggest that these structures are not grammatical (in the formal sense), although syntactic theory has seldom paid much attention to them.

Consider now the example in (14):

(14) Ted drinks meths.

What might be the licensing structure for such a sentence? It's tempting perhaps to think that the verb 'drinks' licenses two arguments, here 'Ted' and 'meths'. This would produce a similar structure to the WG-type analysis shown in (4) and (7) where direct syntactic relationships link the verb and its two arguments. Semantically there is almost certainly a relation between these two arguments and the verb, 'drinks' being a two-place predicate. However, in terms of syntax the situation is less clear. There are good reasons to think that the verb does license one argument, its object 'meths'. If non-finite, though, the same verb will not license a subject, even though there is still a semantic requirement for one. Thus we find examples such as those in (15), where the semantic properties of 'drinks' are clearly not reflected by its syntactic argument structure (analyses involving empty categories are, of course, inadmissible):

- (15) a. Ted was too drunk to drink meths.
 - b. To drink meths might be dangerous.

What we can infer from examples such as these is that the finiteness of the verb, rather than the verb itself, seems to be involved in the licensing of subjects, and generally only a finite verb will license the arguments it requires semantically.

How, then, might we capture the difference between finite and non-finite verbs? Finiteness in some form or other appears to play a part in most languages (Klein 1994), and the question of how it should be integrated into lexical and syntactic structure has arisen in many theories of grammar. One option is to see finiteness (or some equivalent) as an abstract element which in languages like English is amalgamated with the verb by a process of grammatical derivation. This was the traditional view of Principles and Parameters theory (Pollock 1989), although the lexicalist nature of the more recent Minimalist Programme entails that the combination of verbs and finiteness must be accomplished prior to any grammatical derivation (Chomsky 1992). This view is essentially not so very different from that of WG, where finiteness is described as a primitive feature of verbs.

An alternative, and I think more satisfactory, approach might be to regard finiteness as a separate element rather than a feature of the verb. In this way a finite verb would consist of two amalgamated components, the verb itself and a finite element, which I shall refer to as FIN. These two components will together constitute a single word, illustrated in (16) below by their inclusion within square brackets:

(16) $[FIN V]^{10}$

According to this convention a finite verb such as 'drinks' would be represented as (17):

(17) [FIN drink]

Assuming this analysis of finite verbs to be plausible, there is no reason to think that FIN, as a separate element, isn't free to specify its own syntactic requirements in just the same way as the verb does. Thus both elements of a finite transitive verb could license separate dependents; the verb itself need only license its object, while FIN could be responsible for licensing the subject.

Returning now to the example in (14) we can say that it is the FIN element rather than the verb 'drink' which licenses the subject 'Ted'.

¹⁰The ordering of these two elements is irrelevant for our purposes, although I will continue to represent them with FIN preceding the verb.

This analysis has the effect of removing any direct syntactic relation between the verb 'drink' and its subject 'Ted', a potentially desirable result for a number of reasons. Firstly, there is no longer any need to draw a distinction between the syntactic properties of finite and non-finite verbs; instead this supposed distinction will be a matter of whether or not the verb is fused with a FIN element. Furthermore, we are now in a position to offer a simple account for examples such as those in (15) where non-finite verbs are unaccompanied by a subject; given that there is no FIN element associated with these verbs, there will be nothing to license a subject:



Conversely, we can now explain why a finite verb must be accompanied by a subject, even if the verb has no semantic need of one:

Pleonastic elements such as 'it' in (20) appear solely to satisfy the syntactic valency of FIN, and have nothing at all to do with the verb. Consider now (21):

1

Here once again the subject 'John' is licensed by the syntactic properties of FIN, though the nature of its semantic relation to the verb 'seem' is unclear. There is, however, a clear semantic relation between 'John' and 'like', even though the absence of finiteness in the latter verb means that no subject is licensed. Informally then we could say that the finiteness of 'seem' licenses a dependent which semantically belongs elsewhere in the

sentence. I will say more about structures like this later. For the moment, though, we can perhaps see here the beginnings of a possible account of 'raising' phenomena.

Another advantage of the [FIN V] analysis of finite verbs is that it derives a nonconfigurational equivalent of the syntactic VP; a verb will be more closely bound up with its object than with its subject by virtue of the fact that it directly licenses the former but not the latter. We are therefore in a position to account for the well-known asymmetries between subjects and objects, more specifically the so-called verb-object bonding phenomena (Tomlin 1986, Speas 1990). So too case and word order distinctions between subjects and objects can now be derived from the fact that the two are licensed by different elements. This whole approach is, of course, not dissimilar from that of Principles and Parameters theory, where the I° head of the derived [I + V] constituent is responsible for assigning case to (hence licensing) the subject. The differences between these analyses, however, are as important as the similarities. For one thing FIN, unlike I, does not correspond to a structurally-defined position and is entirely absent in non-finite constructions. Another crucial difference, of course, is that the [FIN V] word is neither the product of movement nor is itself dependent on movement for 'feature-checking'. Instead [FIN V] represents the lexicalised fusion of two elements which enter into a wordinternal relation. I will return to this issue in the following section.

What I am suggesting, then, is a simple system of syntax which involves just one structural relation. This relation will remain constant irrespective of the two words which participate in it. A determiner or a preposition will require a dependent in exactly the same way as does a verb or FIN. Similarly all words will have exactly the same need for a head, irrespective of what that head actually is in a given structure. The absence of distinct GRs within the theory means that apart from very general properties of the dependency relation, such as its direction, virtually the entire syntax can be stated in terms of lexical entries. Since there is only one syntactic relation, all that needs to be stored is how individual words participate in this relation. Essentially the grammar will thus reduce to a lexicon listing words and their combinatorial properties; individual syntactic

structures will merely be a reflex of words' propensity for licensing others¹¹. In this way the structure in (22) will be one product of the grammar/lexicon fragment in (23):



Of course the fragment in (23) could serve as the basis for other structures such as 'we know Ted'. Note, however, that these entries only express argument structure and say nothing about the licensing of adjuncts, which will be the subject of section 4.5. So too the entries in (23) are assumed to be entirely non-directional; by convention the arrows which represent a word's requirement of a head and/or dependent point from left to right. This has nothing to do with the actual linear order in which these elements occur in a given structure; as I pointed out before, the fact that heads often precede their dependents in English is determined by principles of word order which are entirely independent of individual words' licensing properties.

The almost complete reduction of the syntax to lexical properties of words has desirable implications for parsing. When it comes to processing a sentence, all the parser has to do is recognise a word, access its combinatorial requirements, and search for other words in the vicinity which satisfy these requirements. A simple locality constraint upon this searching process during the parsing operation can then derive the principle of projectivity or 'no-tangling' (see (5) above), thus further reducing what needs to be stored in the grammar (Kreps forthcoming). I will not pursue these issues here, though (see Fraser 1993 for more about dependency and parsing).

¹¹This view of syntax is, of course, reminiscent of Categorial Grammar (CG) (Ades and Steedman 1982, Oehrle *et al.* 1988). Indeed, CG has itself been described as a variant of dependency theory (Wood 1993 ch. 1), although I do not have space to discuss the various parallels here. It should be pointed out, however, that important differences remain between the monorelational dependency theory that I am exploring here and traditional versions of CG. For one thing, I assume that a word's combinatorial properties remain distinct from its category. Thus both 'wreck' and 'die' will be classed as verbs in spite of their different valency specifications. Furthermore, as will become clear in section 5, the principles of syntactic licensing discussed so far are fundamentally incompatible with semantic structure, thus refuting any version of the 'rule-to-rule hypothesis' favoured by most versions of CG.

4.4 Words and elements

Above I suggested that finite verbs could be analysed as single words consisting of two separate components, the verb itself and a finiteness element FIN. While dependency theories such as WG have no need of any constituents larger than the word, constituency of some sort will obviously have to be recognised at a sub-lexical level. Words are undoubtedly composed of smaller units such as morphemes and phonemes, and all theories must have a means of expressing this fact. The potentially problematic aspect of the analysis of finite verbs in (16) and (17), though, is that both elements of the word were shown to specify their own distinct syntactic properties; both FIN and the verb license a dependent. This analysis would appear to fly in the face of the earlier claim that words were the only units of syntactic relevance for a WG-type dependency theory. How, then, can sub-lexical elements such as FIN enter into a syntactic relation which is supposed to be the preserve of words?

Another important question raised by the same analysis concerns the nature of the relation between the two elements of a finite verb; is the relation between FIN and the verb morphological, and, if so, how does this morphological relation differ from the syntactic dependency relation examined so far?

The simplest answer to these questions would be to suggest that there is actually no difference between the syntactic relations which link words and the 'morphological' relations which link elements within words; the same licensing-based version of dependency could serve to link both words and elements. In this way the two elements comprising a finite verb would enter into a licensing relationship with one another while at the same time licensing their own dependent words:

I assume that FIN licenses the verb rather than vice versa, and thus FIN will constitute the root element of a matrix finite structure. There are many reasons supporting this analysis, though I don't have the space to argue the point properly here.

Although WG does not share this analysis of finite verbs, in certain circumstances it does allow sub-lexical components to participate in dependency relations, notably in the case of clitic constructions (Hudson 1984, Volino 1990) and gerundives (Hudson 1990). The latter are said to be composed of two elements, a verb and an ING 'clitic'. These two

elements enter into a dependency relation, yielding a word with the internal structure shown in (25):

(25)

In essence, then, all I am suggesting is that this kind of analysis could be extended to finite verbs.

Hudson takes the rather unusual view that both the bracketed constituent in (25) and the two elements which constitute it should be classed as words. In this way a single word such as 'walking' can be made up of two (or more) smaller words. This analysis raises some awkward questions for WG and might represent the thin end of a large and dangerous wedge; in what sense is a word the largest unit of the syntax if the same word can be a part of a larger word? Why, moreover, can't two words such as 'drink' and 'meths' in (24) themselves constitute another 'word'? I will avoid questions such as these by examining an alternative approach to the issue whereby elements, rather than words, are recognised as the basic units of syntax.

Let's assume, then, that our basic unit of syntax is the element which, for our purposes, could be considered broadly equivalent to the morpheme. Licensing dependencies will, therefore, hold exclusively between elements:

Words, on the other hand, are lexical rather than syntactic units; they represent those elements (or combinations of elements) which have been stored in the lexicon. Words are thus derivative components which are only of relevance to syntax indirectly, via the elements which go to make them up. More often than not there will, of course, be a oneto-one correspondence between elements and words, in cases where a lexical item is composed of a single element. This will be true, for example, of nouns, prepositions, determiners and 'bare', non-finite verbs such as 'kiss' and 'crash'. In these cases licensing

relations could be said to exist between words, but again, only indirectly, by virtue of the fact that these words happen to be elements:

(27) a. $[E1]_w [E2]_w [E3]_w$ b. $[in]_w [the]_w [box]_w$

In other cases, however, a word could be composed of two or more elements which enter into a licensing relation.

(28) a.
$$\longrightarrow$$
 b.
[E1 E2]_w *[E1 E2]_w

As far as English is concerned, this sort of analysis would not only apply to finite verbs and gerundives, but might also extend to plural nouns and genitives as well:

In these cases, then, a word could be said to correspond to the fusion of two elements which enter into a syntactic relation. The important point is, though, that the licensing relation will remain the same regardless of whether or not the two participating elements are fused together as a single word. Thus the three structures in (30) below are all structurally equivalent, differing only in the combinations of elements which have been lexicalised.

This raises a number of interesting possibilities. For one thing, it is possible that licensing structure actually remains fairly constant universally, with languages differing according to which elements, or combinations of elements, they store as words. Thus, for example, while FIN and the verb are fused in English the same two elements may occur as two separate words in another language:

This appears to be true of languages like Lao (Hoshino and Marcus 1981) and Fijian (Schütz 1985), for example, where finiteness and tense are routinely expressed as distinct particles entirely separate from the verb. Similarly, while (29) represents a genitive noun in English consisting of a nominal and a genitive element, in a language like French the genitive element occurs as a separate word (as indeed it may in English too — 'the Bentley of Ted').

The other possibility, of course, is that separate single-element words which enter into a licensing relation in English may occur as one fused word in another language. Indeed, the prediction is that fusion could occur between any head and its dependent. Without wishing to go into detail here, possible 'word-internal' analogues of virtually all the licensing dependencies examined so far may be discerned in certain languages. For example, it could be argued that 'pro-drop' languages with rich verbal inflection allow the fusion of FIN with a verb and a subject pronoun. Languages like Albanian and Swedish with definite marking on nouns might be displaying fusion between a definite determiner and its nominal dependent, while the rich system of nominal inflection in Finnish and Hungarian could be the result of fusion between what in other languages are prepositions and their nominal dependents. Object clitic constructions could be analysed as the fusion of a verb and an object pronoun, although another possibility is that languages might allow the fusion of a verb and its full nominal dependent. This could be what lies behind some of the incorporation constructions discussed in Baker (1988).

Evidently much work remains to be done in exploring some of these ideas more fully. I hope here to have offered an insight into how the licensing relation offers the possibility of breaking down the sometimes rather artificial distinction between syntax and morphology. Any distinction between the two will essentially be a matter of where languages superimpose words on elements. Much of what I have suggested echoes phrase structure theories, where configurational structure is taken to extend beyond syntax to the internal morphological structure of words (Selkirk 1982). In essence, though, what I am suggesting is that there might not be any real distinction between syntax and morphology at all.

4.5 Arguments and adjuncts

Examples of the licensing relation discussed so far have involved the mutual participation of the head and the dependent; a dependency has been seen to arise from the simultaneous requirement of a head to license something and the (universal) requirement of a dependent to be licensed. This is true, for example, of all the syntactic relations in (22) above, as illustrated by the accompanying grammar/lexicon fragment in (23). I will refer to this type of relation as mutually-sponsored licensing in that the licensing relation between the head and dependent is a product of the syntactic requirements of both.

It is natural to ask, though, whether all instances of the licensing relation have to be mutually-sponsored in this way. Evidently it is unlikely that a dependency could exist if neither the head nor the dependent required it (any such relation would be totally unmotivated and untestable), and I will discount this possibility here. However, there is no logical reason why a dependency could not arise from the lexical requirements of just one of the two participating elements. For example a licensing relation could be sponsored solely by the head. This would involve cases where a head X licenses a dependent Y which happens to be licensed already by another head Z, resulting in the type of double-headed structure encountered in WG. In Kreps (forthcoming) I will suggest that this sort of head-sponsored licensing does indeed occur in certain circumstances, although I shall have nothing more to say about it here. Instead I would like to concentrate briefly on the other possible permutation of the licensing relation, where a dependency is specified solely by the dependent and not the head. Unsurprisingly, I will refer to this as dependent-sponsored licensing.

Cases of dependent-sponsored licensing will arise when a syntactic relation is the product only of the dependent's need for a licenser. In this case the head could be said to license the dependent 'passively' without actually specifying any syntactic requirement to do so. As an analogy we could speak of a whale 'licensing' the presence of barnacles upon its body without necessarily requiring — or even wanting — them there. The relationship between whales and barnacles is driven solely by the latters' need of somewhere to sit. Essentially, then, the distinction between mutually-sponsored licensing and dependent-sponsored licensing boils down to whether a head actually requires a licensee or merely allows one.

The relevance of dependent-sponsored licensing and mutually-sponsored licensing is that together they allow us to derive in a simple and principled way the crucial distinction between arguments and adjuncts. If we had to isolate a single defining characteristic of arguments it would surely be that their presence is required by combinatorial properties of their head. If, then, a head specifically licenses its argument and that same argument requires a licenser, it is easy to see that a head-argument relation will inevitably be mutually-sponsored. The presence of adjuncts, on the other hand, is not specified by a head. For the most part they can occur freely and in any combination, their presence basically determined by the speaker's wish to include them. All that is required of an adjunct is that it occurs with something which it can modify. In this way the mere presence of another word in a structure could suffice to license an adjunct 'passively'. Syntactically, then, the relation between an adjunct and its head will be driven solely by the former's need to be licensed, a clear example of dependent-sponsored licensing.

The broken arrow line in (32) below represents a dependent-sponsored licensing relation between a head and an adjunct:

(32)

A structure such as this will be the product of the grammar fragment in (33)

FIN, 'wreck' and 'his' all specifically require licensees. Whatever fulfils this licensing capacity of these words in a given structure will be classed as their argument. Thus in (32) 'Bentley' is the argument of 'his' which in turn is the argument of 'wreck' which is itself an argument of FIN. All these relations are mutually-sponsored. However in (32) 'wreck' is shown to license two elements while only specifying one dependent in (33). Evidently, then, one of these licensing relations has nothing to do with the head, and will have to be sponsored solely by the dependent. This dependent will be classed as an adjunct.

Notice, however, that this definition of adjuncts is contextual rather than absolute. There is no inherent syntactic property of 'yesterday' which determines that it is an adjunct; it happens to be an adjunct in (32) simply because there is no active valency specification available to license it as an argument. In other cases, however, 'yesterday' can be a syntactic argument:

¹²Following the discussion in the previous section, I assume that FIN specifies two dependents, one of which will be the verb with which it is fused.



In exactly the same way a phrase headed by a preposition can be either a syntactic argument or an adjunct, depending on the licensing context:

The only difference between (35a) and (35b) arises from the different selectional properties of the matrix verbs; 'see' licenses one argument while 'put' licenses two. Here then we can see a possible advantage of deriving 'adjuncthood' contextually. However, this approach raises certain important questions. Why, for example, can't 'yesterday' be the argument of any head (*Ted wrecked his yesterday)? Similarly referring back to the example in (6), why can't the word 'lychee' be 'passively' licensed like an adjunct rather than being left to dangle?



Syntactically, I suspect that there may be nothing wrong with (36) at all. Certainly the licensing properties of all the constituent words are satisfied. Instead the problem seems to be a matter of interpretation. Specifically the word 'lychee' appears to bear no meaningful relation to the rest of the sentence in (36), although 'yesterday' does. Very briefly, I believe that the difference between 'lychee' and 'yesterday' is that the latter specifies its own semantic relation to its head; in a sense 'yesterday' is similar to a

semantic predicate which takes its head (whatever it modifies) as an argument¹³. 'Lychee', on the other hand, specifies no semantic relation to anything else. Instead it relies on another word, its head, to assign it some sort of semantic role. Assuming, then, that semantic roles cannot be assigned by heads to their adjuncts¹⁴, we can now begin to see why 'lychee' is an unsuitable adjunct in (36). Similarly, I assume that the capacity of words like 'yesterday' to specify a semantic relation to their own head makes them unsuitable arguments in most cases since the resulting structures would yield a clash of semantic relations. The prediction is that words like 'yesterday' should only occur as arguments of heads which do not assign any semantic role. The copula in (34) is a plausible candidate.

These are issues which deserve a more thorough investigation than I can provide here. The important point for now is that the basic distinction between adjuncts and arguments can be derived from formal properties of the licensing relation. Simply, an argument can be defined as anything that participates in a mutually-sponsored licensing relation with its head whereas an adjunct can be defined as anything whose syntactic relation with its head arises as a product only of its own valency properties. I assume that this differentiation can then account for the well-attested differences between the behaviour of arguments and adjuncts, such as extraction asymmetries (Kreps: forthcoming). Although this syntactic definition of arguments and adjuncts is contextual, it does seem that inherent semantic properties of words will determine their general suitability as one or the other. One advantage of this approach, of course, is that it is no longer necessary to recognise a separate adjunct relation as a primitive entity of the grammar; mutually-sponsored and dependent-sponsored licensing do not in any sense constitute distinct GRs. Instead they represent two of the three logically possible permutations of the binary licensing relation. A relation between X and Y can be required by X, Y or both X and Y.

As a final note here, it is worth pointing out that the definition of argument suggested above is purely syntactic, and does not necessarily have anything to do with semantic argument structure. A transitive verb such as 'wreck', for example, licenses only one syntactic argument, as illustrated by its entry in (33). Semantically, however, the verb is a two-place predicate which requires two arguments. This discrepancy between syntactic and semantic argument structure will be the topic of the next and final section.

¹³It is crucial here to understand the difference between a syntactic and a semantic argument; the latter has nothing to do with syntactic licensing structure. See below.

¹⁴A reasonable assumption given that heads don't even know that their adjunct dependents exist.

5 Syntactic and semantic relations

In section 3 I outlined a model of dependency syntax based solely on the licensing relation. I suggested that instances of this syntactic relation arise from words' combinatorial properties, their individual requirements to license and be licensed by other words. However, as I pointed out in the previous section, a word's licensing properties do not always correspond with its semantic requirements. In this last section I will explore the possibility of deriving a simple system of semantic relations on the basis of the syntactic licensing structure described so far.

Consider the licensing structure below:

What this representation fails to illustrate is the fact that 'Ted' is interpreted as the subject both of 'want' and 'drink'. These verbs are two place predicates and, semantically at least, they require two arguments. In spite of this, however, they license just one dependent each, their respective objects 'to (drink something)' and 'something'. Neither 'want' nor 'drink' appears to be involved in licensing 'Ted' and thus there can be no syntactic relation between them. Here we see an example of the discrepancy between syntactic and semantic argument structure which I referred to in 4.5. What is required, then, is some means supplementing a syntactic structure such as (37) with a system of semantic relations which would, for example, express the fact that 'Ted' is linked to 'want' and 'drink'.

In spite of my arguments against GRs such as subject, object and adjunct, I referred to 'Ted' as the 'subject' of 'want' and 'drink'. There is actually no harm in this, just so long as the term is used informally to describe a relation rather than being recognised as a primitive component of the theory; while 'Ted' may be described as the subject of 'drink', in purely formal syntactic terms it is the licensee of FIN. From here on I will continue to use the terms 'subject' and 'object' informally, largely as a way of distinguishing between two semantic arguments of a predicate. I could just as easily have used other terms such

as R1 and R2, or 'er' and 'ee', the semantic relations of WG to which my use of the terms 'subject' and 'object' is to some extent equivalent (Hudson 1990 ch. 7)¹⁵.

Whatever the possible discrepancies between syntactic and semantic structure, it is clear that there must be a principled and uniform relation between the two. If this were not true then there would be no sure way of deriving a sentence's meaning from its syntactic form. At the very least, a predicate's semantic arguments will have to be located within a syntactic structure. What I will explore here, then, is how a system of semantic relations may be derived in a simple and principled way from the syntactic licensing structure outlined in the previous section.

From the examples discussed so far it should be clear that if a predicate has more than one semantic argument then one of these — what I refer to informally as the object will correspond to a syntactic dependent of the predicate itself. In (37), for example the objects of 'want' and 'drink' are their respective licensees 'to (drink something)' and 'something'. I assume then that the object of X can be routinely defined according to the principle in (38):

(38) Object of X is a licensee of X.

In this way the object relations, represented beneath the structure, can be derived on the basis of (37) as follows:

Although all objects will be dependents according to (38), it is clearly not the case that all dependents will be objects. According to the assumptions outlined above, the terms 'subject' and 'object' are taken to make an informal distinction between different semantic arguments of a predicate. Although 'to' and FIN in (39) have syntactic licensees, it seems unlikely that these elements play any part in the semantic predicate-argument structure of the sentence, and there is consequently no need to recognise their dependents as objects.

¹⁵I assume that the arguments identified here as subject and object will have to be assigned some sort of semantic role, the precise nature of which will be determined by the predicate. Thus, for example, the subject of a verb such as 'faint' will be an experiencer, while the subject of 'run' will be an agent.

In the case of objects, then, (38) guarantees that there will be a close correspondence between semantic and syntactic structure. The subject relation evidently can't be derived in the same straightforward way, though, given that verbs do not even enter into a direct licensing relation with their subject arguments. The subject relation will thus have to be derived on the basis of some more complex aspect of syntactic structure. What I will assume here is that the subject of a predicate can be defined as its closest suitable argument.

(40) Subject of X is the closest suitable argument to X

The reference to closeness in (40) doesn't necessarily have anything to do with physical proximity, but is instead a structural notion referring to links in a dependency chain. The suitability requirement in (40) basically serves to exclude direct and indirect heads and dependents of X; obviously the nearest argument to a predicate will be its own licensee, if it has one. However, this does not qualify as a suitable argument since, according to (38), it will be defined as the predicate's object. Evidently then, to find a predicate's subject we will have to look in the other direction, up the licensing chain to the predicate's head, and beyond.

The first head up the dependency chain will evidently be the predicate's own licenser. If this head has a dependent, other than the predicate itself of course, then it will be the closest suitable argument to the original predicate and will thus qualify as its subject. If, however, the predicate's immediate head does not have another dependent then it will be necessary to move one step up the licensing chain to the next head to see if this has any suitable dependent, and so on¹⁶. Consider the schematic representation in (41):

(41)



Say that F is a predicate requiring a subject and an object. G, the licensee of F, will qualify as it's object. In seeking a subject we must move up the licensing chain to find the nearest head above F with a suitable dependent. The first point up the chain is F's own licenser, E. E's only dependent, however, is F which, of course, can't be the subject of

¹⁶This somewhat cumbersome algorithm for defining a subject relation can be sharpened considerably when processing factors are taken into account (Kreps: forthcoming).

itself, and thus we must move one step further up the chain to E's licenser, C. C has two dependents, E and D. E doesn't qualify as a suitable argument for F because it is the head of F. C's other dependent, D, is, however, a suitable argument, and the closest one to F. D will thus qualify as F's subject. If D didn't exist, of course, then the search for a subject would have to go on up the dependency chain to C's licenser, B which happens to have a suitable argument, A. Given the existence of D in (41), however, A will not qualify as the subject of F since D is closer to F.

The inclusion of a relative term like 'closest' in the definition in (40) allows a degree of flexibility to be built into the derivation of subjects. This is a positive advantage in that it potentially allows for a uniform definition of subject to be maintained over a variety of structures. Moreover, unlike an object, which is defined in absolute terms, a single word can simultaneously qualify as the 'nearest suitable argument' (and hence the subject) of more than one predicate. Consider first the simple structure in (42):

The predicate 'want' has two arguments, one of them, its object, will be its licensee 'a' (I assume that the status of 'a' as object of 'want' is also passed down to its own licensee 'drink'). To find the subject of 'want' we move up the dependency chain to the element immediately above 'want', its head FIN. FIN has a second dependent 'Ted' which will qualify as the nearest suitable argument to 'want' and will thus be identified as its subject. As before, derived semantic relations are represented beneath the structure:

Returning now to (37) above, we can employ the same principle to derive the subject of both 'want' and 'drink'. The subject of 'want' will obviously be derived in the same way as it was in (43). In the case of the verb 'drink', however, its immediate head, 'to', does not have another dependent and thus we must move further up the dependency chain to find a suitable argument. The head of 'to', 'want', doesn't have a second dependent either, however its own head, FIN, does in the form of 'Ted'. FIN is thus the first head above 'drink' which licenses a suitable argument, and 'Ted' will qualify as the subject of both 'want' and 'drink':



Exactly the same analysis will apply to 'raising structures' involving verbs like 'seem':

It is unclear whether 'Ted' really bears any semantic relation to 'seems' or if it is only a syntactic dependent of the FIN element fused with 'seem'. Essentially, though, this is unimportant here; 'like' clearly does require a subject, and this will be defined in exactly the same way as before. The first head up the licensing chain with a suitable argument is again FIN and thus 'Ted' will qualify as the subject of 'like'. Apart from the question concerning the semantic relation between 'Ted' and 'seem', the structure in (45) is exactly parallel to that in (44), in spite of the fact that these two sentences often receive very different analyses in other theories. The flexible definition in (40) allows the subject to be derived in exactly the same way for simple matrix clauses (43), 'subject control' constructions (44), and examples of 'raising' (45). Moreover this unified analysis involves no syntactic movement or empty categories such as PRO, neither does it require any distinct syntactic relation analogous to WG's x-complement.

The principle in (40) may also be used to derive the subject relation in other constructions. (46) below is an example of an 'object control' structure:



The verb 'persuade' is semantically a three-place predicate which licenses two dependents. Let's assume, then, that 'persuade' will have two objects, one nominal and the other clausal, corresponding to its two dependents, and one subject, derived according to the principle in (40). A partial semantic structure for (46) is shown in (47):



Turning now to the predicate 'drink', as before its object will correspond to its dependent 'the (petrol)', in accordance with (38). Moving up the licensing chain to locate the subject, the first head which licenses a suitable argument is the verb 'persuade'. In this way the (other) object of 'persuade', 'Fred' will be defined as the subject of 'drink':



Although this semantic structure may seem complex, from an interpretational point of view it appears to be correct, and, moreover, it is entirely derived from the much simpler syntactic structure by the two principles in (38) and (40).

I assume that exactly the same analysis will apply in similar cases of 'object control' with verbs such as 'help' and 'ask'. Essentially the flexible definition of the subject relation allows the difference between 'subject control' (44) and 'object control' (48) to be reduced to a matter of whether or not the matrix verb licenses a second dependent apart from the complement clause. If it does then this second dependent will qualify as the closest argument to the non-finite verb in the embedded clause, and will be defined as its subject. If there is no second dependent, then the nearest suitable argument to the embedded verb will automatically be the licensee of the matrix FIN, as in the case of (44). There is, however, one notable and troublesome exception to this generalisation involving the verb 'promise'. The subject of a non-finite complement verb is routinely interpreted as the subject of 'promise', in spite of the presence of a second licensee of the verb:



I cannot pretend that I have a satisfactory answer to this problem. It is worth pointing out, though, that 'promise' is, as far as I know, a unique exception.

There is one final construction worth examining in connection with derived subjects, illustrated in (50):



This is an example of what is sometimes referred to as an ECM (Exceptional Case Marking) construction (Chomsky 1981). For our purposes, though, (50) differs from (49) only in terms of the semantic argument structure of the matrix verb. I assume that 'expect' here licenses two syntactic dependents. Unlike 'persuade', however, semantically 'expect' is a two-place predicate. The only thing expected in (50) is the state of affairs whereby Ted will get drunk. I assume, then, that 'expect' has a subject and only one object, which will be the clause headed by 'to'. The verb's other licensee 'Ted' will thus bear no semantic relation to it at all, another apparent mismatch between syntactic and semantic argument structure:



Clearly 'Ted' will have to be related semantically to something in (51). It so happens, of course, that 'drink' requires a subject and an object. The latter will be defined as its dependent 'it' and the subject will once again be defined as the closest suitable argument. Looking up the licensing chain from 'drink', the first head with a second argument is the

verb 'expect' whose hitherto roleless dependent 'Ted' will be identified as the subject of 'drink':



In this section I hope to have demonstrated how the syntactic licensing structure described in section 4 may serve as a basis for deriving semantic relations. In particular I have shown how the same flexible definition in (40) can correctly identify a predicate's subject in a variety of circumstances. The subject will be derived in exactly the same way for simple finite verbs (43), 'subject control' and 'object control' structures (44) and (48) as well as 'raising' and 'ECM' constructions (45) and (52)¹⁷. In all these cases the subject of a predicate will be its closest suitable argument. Moreover, the simple analysis which underpins all these examples has not involved any syntactic movement or the recognition of empty categories such as PRO. Neither has the analysis required any specific syntactic relation such as WG's x-complement. The sharing relation sanctioned by the x-complement in WG is now an automatic reflex of the licensing structure in combination with the principle of subject derivation in (40). Far from being a weakness of the theory, then, the mismatch between syntactic and semantic structure which the licensing relation forces can actually be seen as a strength.

6 Conclusions

In conclusion, I hope to have shown how it is possible for dependency theory to survive without having to recognise a class of distinct grammatical relations. Indeed, I believe that the proliferation of GRs that has characterised some grammars has had a bad overall

¹⁷In fact the same definition of the subject relation also allows us to account for the interpretation associated with adjunct clauses such as that in (53):

⁽⁵³⁾ Ted crashed his Bentley after drinking the aftershave.

However this requires certain ancillary assumptions concerning adjuncts which I do not have the space to describe here (Kreps: forthcoming).

effect on relational theories of syntax. Recasting the dependency relation in terms of licensing, however, demands a more parsimonious, constrained grammar, where no separate GRs are recognised and where dependencies only arise from well-defined combinatorial properties of words. In addition a unitary relation such as licensing allows certain generalisations to be captured between syntactic and morphological structure. I have shown how the simple, almost minimal syntactic structures involving the licensing relation can also serve as a basis upon which a system of semantic relations can be derived with surprising simplicity and uniformity.

Many of the assumptions made in this paper have parted company in fundamental ways from what is standardly taken for granted in WG, and some of my proposals may actually seem to have more in common with certain aspects of Principles and Parameters Theory or Categorial Grammar. I believe, however, that too much valuable work has been done within all of these traditions for any one of them to be totally wrong, and perhaps the best hope for any approach to language is to seek to combine some of the insights and strengths of 'competing' theories.

References

Ades, A and M. Steedman (1982). On the order of words. Linguistics and Philosophy 4, 517-558.

- Baker, M. (1988). *Incorporation: a theory of grammatical function changing*. Chicago University Press, Chicago.
- Blake, B. (1990). Relational Grammar. Routledge, London.
- Brody, M. (1994). Phrase Structure and Dependence. UCL Working Papers in Linguistics 6.
- Chomsky, N. (1957). Syntactic Structures. Mouton, Amsterdam.
- Chomsky, N. (1965). Aspects of the Theory of Syntax. MIT Press, Cambridge, Mass.
- Chomsky, N. (1981). Lectures on Government and Binding. Foris, Dordrecht.
- Chomsky, N. (1986). Knowledge of Language: Its Nature, Origin and Use. Praeger, New York.
- Chomsky, N. (1992). A minimalist program for linguistic theory. MIT Occasional papers in linguistics 1.
- Chomsky, N. (1993). Bare phrase structure. MIT Occasional Papers in Linguistics 5.
- Chomsky, N. (1995). The Minimalist Program. MIT Press, Cambridge, Mass.
- Fraser, N. (1993). Dependency Parsing. PhD Thesis, University of London
- Fraser, N. and R. Hudson (1992). Inheritance in Word Grammar. Computational Linguistics 18. 133-158.
- Gazdar, G., E. Klein, G. Pullum and I. Sag (1985). *Generalised Phrase Structure Grammar*. Blackwell, Oxford.
- Hoshino, T. and R. Marcus (1981). Lao for Beginners. Charles E. Tuttle.
- Hudson, R. (1984). Word Grammar. Blackwell, Oxford.
- Hudson, R. (1988a). Coordination and Grammatical Relations. Journal of Linguistics 24. 303-342.
- Hudson, R. (1988b). Extraction and Grammatical Relations. Lingua 76. 177-208.

Hudson, R. (1990). English Word Grammar. Blackwell, Oxford.

- Hudson, R. (1994). Discontinuous phrases in Dependency Grammar. UCL Working Papers in Linguistics 6.
- Klein, W. (1994). Time in Language. Routledge, London.
- Kreps, K. (forthcoming). PhD thesis, University of London.
- Oehrle, R., E. Bach and D. Wheeler (eds) (1988). *Categorial Grammars and Natural Language Structures*. Reidel, Dordrecht.
- Perlmutter, D. (ed) (1983). Studies in Relational Grammar 1. University of Chicago Press, Chicago.
- Pollard, C. and I. Sag (1987). *Information-Based Syntax and Semantics, Volume 1. Fundamentals*. CSLI, Stanford.
- Pollard, C. and I. Sag (1994). *Head-Driven Phrase Structure Grammar*. University of Chicago Press, Chicago.
- Pollock, J-Y. (1989). Verb movement, universal grammar and the structure of IP. *Linguistic Inquiry* 20. 365-424.
- Robinson, J. J. (1970). Dependency structures and transformational rules. Language 46. 259-285.
- Rosta, A. (1994). Dependency and grammatical relations. UCL Working Papers in Linguistics 6.
- Selkirk, E. (1982). The Syntax of Words. MIT Press, Cambridge, MA.
- Schütz, A. (1985). The Fijian language. University of Hawaii Press, Honolulu.
- Speas, M. (1990). Phrase Structure in Natural Language. Kluwer Academic Press, Dordrecht.
- Tesnière, L. (1959). Eléments de Syntaxe Structurale. Klincksieck, Paris.
- Tomlin, R. (1986). Basic Word Order: Functional Principles. Croom Helm, London.
- Volino, M. (1990). Word Grammar, Unification, and the Syntax of Italian Clitics. PhD Thesis, University of Edinburgh.
- Wood, M. M. (1993). Categorial Grammars. Routledge, London.