Word Grammar and Construction Grammar[1]

Richard Hudson

Construction Grammar (CG) as defined by Adele Goldberg's recent work and Word Grammar (WG) are very similar to each other, and together form a bridge between cognitive and generative approaches to language structure. For example, like other cognitive theories CG and WG claim that language is usage-based, but like generative theories they recognise a separate level of syntax. However, I also recognise four general differences between CG and WG, all of which can be illustrated from the analysis of ditransitives. First, sentence structure consists of dependency structure in WG, whereas CG is somewhat agnostic about sentence structure. Second, CG appears to assume a rather simple kind of semantic structure whereas WG offers a rich network analysis of the entire semantic frame. Third, CG appears to recognise a single level of "form" which includes phonology, morphology and syntax, whereas these are related hierarchically in WG. The fourth difference is that WG spells out a theory of processing somewhat more fully than CG does. In each case, I show that the extra structure and theory available in WG would improve the CG analysis. Finally, I suggest that the notion of "construction" is problematic.

1. Overview

Word Grammar (**WG**) has so much in common with Construction Grammar (**CG**) that similarities are a lot easier to find than differences. These similarities are partly a case of independent parallel development, but I have always been a great admirer of the work of the CG group, especially Charles Fillmore (Fillmore and others 1988, Fillmore 1986, Fillmore 1982, Kay and Fillmore 1999), George Lakoff (Lakoff 1987, Lakoff 1977) and Adele Goldberg (Goldberg 1995, Goldberg and Bencini 2005, Goldberg 2002, Goldberg 2006, 1995; Goldberg and Bencini 2005). In fact, on an autobiographical note, I was so impressed by the early CG work that I visited Berkeley in 1987 to find out more. I borrowed a number of important ideas from the CG group covering such topics as prototype effects, lexical semantic analysis and multiple inheritance. This article, which continues the discussion started in Holmes and Hudson 2005, is an attempt to repay this debt by offering four very general ideas that (in my opinion) would make CG even better at very little cost:

- that syntactic structure consists of dependencies between words.
- that semantic frames are part of the analysis.
- that the levels of phonology, morphology, syntax and semantics are autonomous.
- that the cognitive context should be enriched.

Section 2 reviews what strike me as the most important similarities between the two theories. In general terms, I shall suggest that CG and WG share more ideas with each other than with other theories, but that both of them straddle the division between cognitive linguistics and generative linguistics. On the one hand, they are “cognitive” in their commitment to embedding a theory of language in a more general theory of cognition from which most (or maybe even all) of the properties of language may be derived; they share this commitment (and many other ideas) with Cognitive Grammar (Langacker 2000) and a range of other “cognitive” theories (Croft and Cruse 2004; Evans and Green 2006). But on the other hand, they are “generative”, like the Minimalist Program (Chomsky 1995) and Lexical-Functional Grammar (Bresnan 2001), in recognising an independent level of syntax; and like Head-driven Phrase Structure Grammar (Pollard and Sag 1994), they both allow complex structures to be inherited from simpler structures, some of which are very specific. This view of CG and WG as a bridge between competing theories of language structure is shown in Figure 1. Moreover, CG and WG also bridge the division between linguistics, conceived narrowly as the study of language structure, and psycholinguistics, the study of language use and learning. I would like to think that each of these bridges includes all the best bits of the other theories without too many of their weaknesses.

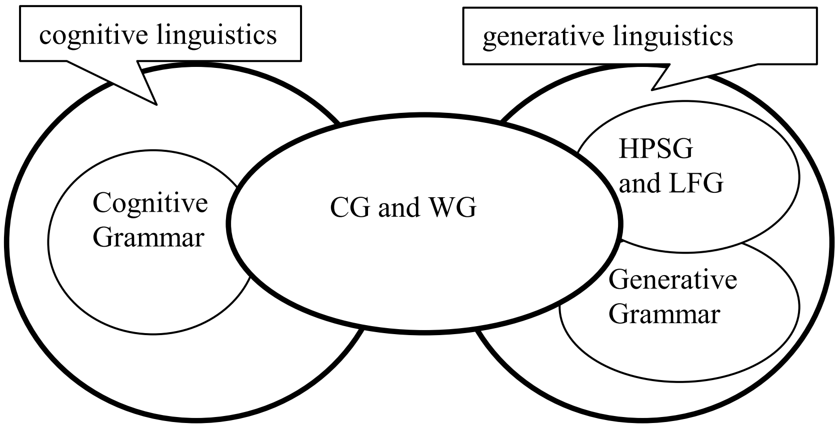

Figure 1 CG and WG are a bridge between cognitive and generative linguistics

Section 3 then introduces the first main area of disagreement between CG and WG: the nature of sentence structure. Is it based on phrases (CG) or on words (WG)? CG at least implicitly follows the American tradition of phrase structure (with the addition of functional labels), while WG is a typical European theory based on dependency structure. The two theories make different predictions and have different success rates in explaining various phenomena. I shall review the evidence which supports dependency analysis over phrase-structure analysis. In a sense this discussion will be about mere technicalities, but, the devil being in the detail, I shall draw some far-reaching conclusions.

Section 4 shows the value of taking the idea of "frame semantics" seriously. If every concept is defined by a "frame" of related concepts, then the framing concepts should be included in the analysis. The discussion applies this principle to the ditransitive construction and its meaning, and shows the benefits of including the framing semantic structures in the total analysis. It turns out that various meanings of this construction do share a common semantic structure, although their variation produces a family-resemblances cluster of meanings.

Section 5 then argues that form is a great deal more independent of meaning than is allowed if grammar pairs every form with a meaning (as in CG). In this section I argue that language has a traditonal multi-level architecture rather than the simple form-meaning structure of CG, and that dependency structures leave no separate role for constructions – in other words, dependencies applied to individual words are constructions (Gisborne, this volume), and every "constructional" fact can be stated in terms of single words and dependencies; moreover, as Gisborne observes, any category used invoked in classifying dependencies must be "framed" by a network of dependency types for just the same reasons that support "frames" in semantics. In short, the correspondences between syntax and meaning can be analysed better without assuming distinct constructions than with them. The WG view also includes the claim that there is one more distinct level than is sometimes recognised, an extra level of morphology between syntax and phonology.

Finally, section 6 is about the "cognitive context" of language – how conceptual knowledge is organised, how it is used and how it is learned. This is a fundamental question for theories such as CG and WG, both of which rest on the assumption that a theory of language structure must be embedded in a more general theory of cognition. The cognitive theories are encouragingly similar, but I shall pick out some important differences and suggest directions for further research.

A small challenge in comparing theories is the range of variation that can usually be found within any named theory. WG has its fair share of diversity (illustrated richly in Sugayama and Hudson 2006), but for simplicity I assume here my own most recent views on all issues, and since I have just finished a book about the theory (*Language Networks. The New Word Grammar*, Hudson 2007a), it is this book that defines WG. For CG, on the other hand, my definition of the theory will be based on the work of Adele Goldberg, the author whose work is most familiar to me (Goldberg 1995, Goldberg 1998, Goldberg and Bencini 2005, Goldberg 2002); and in particular I shall use her most recent book, *Constructions at Work. The Nature of Generalization in Language* (Goldberg 2006).

2. Similarities between CG and WG

As I explained in section 1, there are more distinctive features that unite CG and WG than that divide them. To borrow a useful term from Goldberg, both theories are "constructionist" (Goldberg 2006:3). Goldberg distinguishes two senses for this word.

On the one hand, theories are constructionist if they "emphasize the role of grammatical constructions" as "learned pairings of form with semantic or discourse function" (Goldberg 2006:5). (This approach assumes that constructions must be part of the internalised grammar, and not merely immanent in observable data – i.e. we are dealing with g-constructions rather than u-constructions, in the terminology of Rosta, this volume.) In other words, a constructionist theory claims that grammar is organised in such a way that each "form" that is stored can be paired directly with a structure which shows the form's meaning. Of course, it is uncontroversial to claim that formal structures (e.g. syntactic structures) can be mapped onto structures of meaning; but what is controversial is the claim that we store some formal structures which are quite **specific** as well as the more general patterns that they contain. For example, in both theories the grammar includes a stored entry for the ditransitive construction which includes some information which might be derived from more general constructions. This emphasis on specific syntactic patterns contrasts both theories with Chomsky's Minimalist Program, whose stated aim is to explain all syntactic patterns as the effects of independent principles: "The language-particular rules reduce to choice of values for … parameters. The notion of grammatical construction is eliminated, and with it, construction-particular rules."(Chomsky 1995:170)

However, another distinctive characteristic of CG constructions is the nature of the "form" that they pair with function. For phrasal constructions it is a **syntactic structure**. The recognition of syntax as a level of structure distinct from both semantics and phonology is a property that both CG and WG share not only with the Minimalist Program but also with Lexical-functional Grammar (Bresnan 2001). In contrast, at least early versions of the other main cognitive theory, Cognitive Grammar, deny the existence of a separate syntactic structure; instead, "only semantic, phonological and bipolar symbolic units are posited … Syntactic units are bipolar, with semantic and phonological poles." (Langacker 1990:102) In other words, syntactic units are merely realization relations between meanings and sounds. Somewhat similarly, Head-driven Phrase Structure Grammar merges syntactic and semantic structures into a single "synsem" structure, so the units of grammar are again bipolar "signs", without the independent syntactic structures of both CG and WG (Pollard and Sag 1994:3). The distinct syntactic structures of both CG and WG thus align these theories with the Minimalist Program and LFG, in contrast with both early Cognitive Grammar and HPSG.

The other sense that Goldberg gives to *constructionist* is that "languages are **learned** – that they are constructed on the basis of the input together with general cognitive, pragmatic and processing constraints" (Goldberg 2006:3). The crucial word here is *learned*, which stresses the major role of experience rather than genetics. Both CG and WG are "usage-based" theories, explaining knowledge as the residue of countless encounters with specific tokens of language. Every item of vocabulary is a generalization across a range of highly contextualized tokens, and every grammatical generalization is similarly based on the characteristics of a range of these stored vocabulary items. To take the ditransitive construction as an example again, it is learned by induction across a stored collection of verbs that take two objects, and each of these verb-types in turn is induced from a collection of tokens. This view of learning comes from Cognitive Grammar (Langacker 1987, Langacker 2000, Langacker 1990) and is one of the most important contributions of that theory because of its radical consequences for our view of language. In place of the static and "purely synchronic" idealization of most other theories, we have a constantly growing system of elements with different degrees of "entrenchment" in which synchrony and diachrony meet. Any theory that links language structure so closely to experience has to include a theory of how language is learned, and CG and WG both include such a theory (which I discuss briefly in section 4).

Perhaps the most important similarity, at least from the point of view of WG, is the shared assumption that the product of this learning – a person's knowledge of language – is a single unified **network**. "What makes a theory that allows constructions to exist a "construction-based theory" is the idea that the network of constructions captures our grammatical knowledge of language *in toto*, i.e. **it's constructions all the way down.**" (Goldberg 2006:18, author's emphasis) It should be noted that constructions include lexical items, so when Goldberg refers to "grammatical knowledge" she actually means "linguistic knowledge", i.e. our entire linguistic competence from specific lexical items to the broadest of grammatical and phonological generalizations. The sum total of linguistic knowledge is contained in a single network in which there is no formal distinction between lexical items and grammatical rules. The same assumption is fundamental to WG: "Language is a conceptual network." (Hudson 1984:1, quoted in Hudson 2007a:1). On the other hand, the networks envisaged in these two theoretical statements are rather different because they allow different kinds of nodes: just constructions in CG, but any kind of concepts (including constructions) in WG. I return to this difference in section 6.

Why do I think the network idea is so important? (In contrast, Goldberg merely takes it for granted; in fact, neither of her books even includes the word *network* in its index.) Because this brings language structure very clearly into the realm of long-term memory, which most cognitive psychologists think of as a network (Reisberg 1997:257). If “knowledge of language is knowledge” (Goldberg 1995:5) – a beautiful formulation with which I agree totally – then knowledge of language must have the same organisation as other kinds of knowledge; and if other kinds of knowledge (e.g. about birthday parties or kinship) are organised as networks, then the same must be true of language. This conclusion may seem innocuous, but it actually excludes any contrast between “rules” or “principles” and stored knowledge, thereby immediately ruling out any theory that invokes extra principles, rules or constraints which are not expressed either in network terms or in terms of processing or learning. If language really is a single unified network, as described by WG and CG, then most of the theories that dominate linguistics are fundamentally wrong: not only all of Chomsky’s theories (Chomsky 1965, Chomsky and Lasnik 1993, Chomsky 1995, Chomsky 1957) but also Head-driven Phrase Structure Grammar (Pollard and Sag 1994), Lexical Functional Grammar (Bresnan 2001) and others. Clearly, the network idea raises some fundamental (and difficult) questions for us all.

These similarities between CG and WG are all so fundamental that they deserve a great deal more discussion, and it would be easy to extend the list by pointing out other similarities. For example, I share the “commitment in principle to account for the entirety of each language” (Kay and Fillmore 1999), including its non-canonical constructions such as *What about a drink?*; I have even given this commitment the memorably awful name “poly-constructionism” (Hudson 1990:5). The main point, I think, is to establish that CG and WG start from very similar basic assumptions about the way language works and how it fits into the human mind, and have the same ultimate goal. Consequently it should be possible for ideas to flow relatively smoothly between the theories, and as far as I can see, the ideas that I outline below are fully compatible with the basic aims and assumptions of CG.

3. Syntactic structure consists of dependencies

As we have seen, “it’s constructions all the way down” in CG. Constructions are defined as “learned pairings of form with semantic or discourse function, including morphemes or words, idioms, partially lexically filled and fully general phrasal patterns" (Goldberg 2006:5). The discussion in this section will focus on the phrasal patterns, whether idiomatic, partially lexically filled or fully general. As Goldberg points out (ibid), some linguists reserve the term "construction" for phrasal patterns, and call single words or morphemes "signs", so this section is about constructions in this narrower sense. The question is how to represent "patterns" in multi-word sentence structure.

Ever since I first used the name "Word Grammar" (Hudson 1984), I have argued that sentence structure consists of nothing but **dependencies between individual words** – hence the name of the theory. At one time I thought differently. The first grammatical theory that I adopted was what at that time was still called Systemic Grammar (Halliday 1961, Halliday 1985; Hudson 1971), which, under the influence of post-Bloomfieldian grammarians in the USA, assumed a hierarchical part-whole analysis of sentences. At that time, like most other linguists, I was unaware that there was an alternative, the dependency-grammar tradition of Europe, which is still taken for granted in the school-teaching of many European countries and which arguably dominated linguistics until the twentieth century (Covington 1984, Gaifman 1965, Heringer 1993, Kunze 1975, Mel'cuk 1988, Owens 1988, Percival 1990, Tesnière 1959, Venneman 1977). In contrast, most American linguists still follow Bloomfield into phrase structure (Bloomfield 1933, Percival 1976), and in this respect CG appears to be a typical American theory. In CG, a sentence seems to have a hierarchical phrase structure, whereas in WG there is no phrase structure but there are direct dependency links between individual words. A simple example is shown in Figure 2. As we shall see in later diagrams, the arcs are also labelled to distinguish subjects, objects and so on. The essential point to notice about this diagram is that there is no separate node for the clause (or sentence), nor for the noun phrases *his students* and *good marks*. In each case, the head word (*gave, his, marks*) carries all the properties of the phrase that it heads, so a separate node would be redundant.

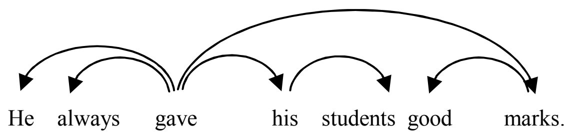

Figure 2: In WG, words are linked directly, not via phrases

However, this difference may not go very deep, because CG is almost agnostic on the details of sentence structure. In fact, there is not a single example of a full sentence structure in either of Goldberg's books. The only diagrammed structures are for single constructions, which are presented as collections of structure like the one for the ditransitive construction in Figure 3 (Goldberg 2006:20). The ditransitive construction consists of this entire combination of elements.

Sem: intend-CAUSE-RECEIVE (agt rec$_{\text{(secondary topic)}}$ theme)

verb ( )

Syn: Subj Obj1 Obj2

Figure 3: The ditransitive construction in CG

The CG analysis translates easily into the WG notation of Figure 4, where for simplicity I replace "intend-CAUSE-RECEIVE" by its synonym 'giving' and merge the two syntactic layers.

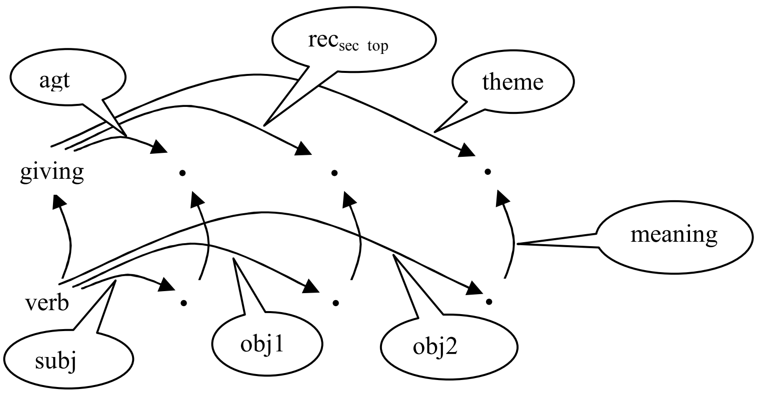

Figure 4: The CG ditransitive construction in WG notation

The most obvious difference between Figure 4 and Figure 3 is the use of arrows in place of the brackets. For example, instead of "verb (Subj Obj1 Obj2)" we now have a separate arrow from "verb" to each of its arguments. This notation has the advantage of clearly distinguishing relations and their classification (e.g. as Subj, Obj1 or Obj2) from non-relational nodes or "categories" such as "verb". This clear distinction in notation between relations and nodes reflects the very different statuses of the things labelled "verb" (a category) and "subj" (a relation), and would actually serve CG better by avoiding the uncertainty over relations and categories that Goldberg mentions (Goldberg 2006:21, fn 2). In Figure 4 the difference is exaggerated by the use of callouts, but WG diagrams normally label arrows directly. Moreover, the separation of the label from the node or arrow that it labels allows us, where necessary, to leave an arrow or a node unlabelled.

The main question in this section is the choice between phrase structure and dependency structure in syntax, so we should concentrate on the syntactic parts of these two diagrams. The use of labelled relations such as Subj in the CG diagram is already a gesture in the direction of dependency structure in contrast with phrase structure as found in classical versions of phrase structure in the Chomskyan tradition. On the other hand, labelled relations are combined with phrase structure in other theories such as Lexical Functional Grammar (Bresnan 2001), Systemic Functional Grammar (Halliday 1985), Relational Grammar (Blake 1990) and Functional Grammar (Dik 1989), so labelled relations do not in themselves indicate dependency structure. The crucial question is whether the dependents are phrases (for phrase structure) or single words (for dependency structure). So far as I know, the CG literature does not address this question, and phrase structure is by and large taken for granted. I should like to suggest that dependency structure fits the assumptions of CG better.

Before I turn to the specifics of CG, we can review some general advantages of dependency structure. Compare the two structures in Figure 5 for the sentence *Cows eat green grass*, where the phrase-structure diagram is adapted to CG by the addition of relation labels ("s", "o" and "h" for "subject", "object" and "head" respectively; "a" stands for "adjunct") and by the omission of a VP label to allow the subject and object to be sisters (as seems to be intended in the formulae).

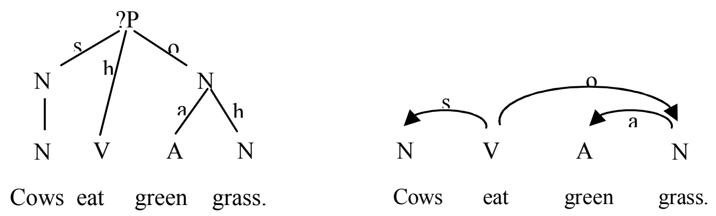

Figure 5: Phrase structure and dependency structure compared.

The most obvious difference lies in the number of nodes and links: just one node per word on the right, compared with an extra three on the left, and just three links on the right compared with six on the left. Those extra nodes and links require justification in terms of generalizations that they allow which would otherwise not be possible. It is much harder to find such generalizations than might be thought. Here are some possibilities to consider.

First, maybe the phrasal nodes help with classification: NP is different from N, so it may be important for the grammar to distinguish them. The standard assumption is that this is indeed so, because nouns combine with adjectives to form NPs, whereas it is NPs that combine with verbs to form clauses; and this distinction requires the "unary branching" above *cows* to show that this is in the intersection of the two classes: both N and NP. But this distinction is easy to make in terms of dependencies: a noun allows an adjective as its dependent and a verb as its parent (the word on which it depends). Phrase structure simply mirrors these dependencies by adding an extra node to hold together the head and all its dependents. The dependency arrow from *grass* to *green* achieves exactly the same effect as the phrasal NP node, so the latter is a more complicated solution with three nodes and two links instead of two nodes and one link. Moreover, CG seems to follow the general trend in phrase structure of requiring phrases to be endocentric, with one word picked out as the phrase's head. What this means is that the classification of the phrase is entirely predictable from that of the head, so the difference between "phrase" and "word" is the only possible contribution of the phrase. Since this distinction can easily be read off the dependency relations, phrase nodes are redundant.

The second possible role for phrasal nodes is in handling word order: phrases hold all the dependents of a word together, but this can be done just as easily without phrases. Simply assuming phrases does not in fact achieve this effect, but only when combined with the extra theoretical assumption that the words inside a phrase must form a continuous string. This can be expressed in various ways, e.g. in terms of brackets or non-intersecting lines in structure diagrams. But the same is true of dependency relations. These too can be combined with an equivalent extra theoretical assumption about the dependents of a word forming a continuous string, which may also be related in various ways to structure diagrams. For example, if we split the phrase *green grass* to give **Cows green eat grass*, each of the diagrams in Figure 6 contains two intersecting links, so we might simply ban intersecting links. There happen to be better solutions in WG (Hudson 2007a:131ff), but the point is that it is just as easy to keep phrases together in dependency structure as it is in phrase structure. And of course, the well-known exceptions such as raising, extraction and extraposition, where phrases are allowed to be discontinuous, can be accommodated at least as easily in dependency structure as in phrase structure (Hudson 1990:113ff, 354ff).

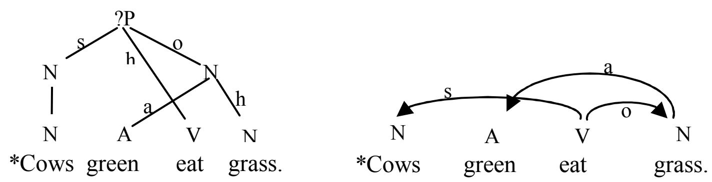

Figure 6: A discontinuous phrase in phrase structure and dependency structure.

To judge by the two defences of phrase structure just considered, the usual arguments for phrase structure are woefully inadequate because none of them considers the dependency alternative. One defence that does consider it explicitly is the argument that dependency structure is merely a notational variant of phrase structure, with just the same weak generative capacity (Gaifman 1965, Robinson 1970). It is certainly true that there are types of phrase-structure grammar that can systematically be matched by dependency grammars with the same weak generative capacity (in the technical sense of the range of symbol strings that they generate, disregarding the structures assigned to these strings). However, the grammars we have these days are so much more sophisticated than anything considered in the 1960s that such comparisons are meaningless, and especially so if we consider strong generative capacity (i.e. the range of structures assigned). To take a rather spectacular example of non-equivalence, dependency grammars allow mutual dependency unless this is explicitly ruled out, whereas this is fundamentally impossible in phrase structure because nothing can be part of itself. This is not a mere hypothetical possibility in WG, which recognises syntactic structures where two words are interdependent (Hudson 2007a:142).

The general argument for dependency structure thus rests on the following claims:

- It is simpler than phrase structure in terms of the nodes and links in a structure diagram.
- It allows the same generalizations about classification and word order as phrase structure.
- But it is not a mere notational variant, as it allows analyses (e.g. mutual dependency) which are not possible in phrase structure.

There is a great deal more general evidence in favour of dependency analysis for which there is no room here; this evidence comes from areas as diverse as computational linguistics and child language (Hudson 2007b:118ff). It is true that dependency analysis is vulnerable to a few theoretical challenges, the strongest of which rests on the absence of any obvious equivalent of the c-command relation that has played such a dominant role in Chomskyan linguistics (ibid:122). However, CG does not use this notion so it is irrelevant to the choice between phrase structure and dependency structure as the theoretical basis for CG syntax.

I now present some benefits of dependency structure which are more directly relevant to CG. Firstly, dependency structures are "flat", and in particular they have a single layer of structure for each verb and its dependents, so dependency structure yields one structure per clause (defined as a phrase headed by a verb), which seems to be what CG analyses need; for example, the CG formula for the ditransitive construction includes the subject as well as the two objects, with the implication that these are sisters, as they necessarily are in a dependency analysis.

Secondly, dependency structures in a sentence are lexically specific because each of their nodes is a specific word (though of course they may be represented in a much more general way in a grammar). This allows the grammar, where necessary, to refer directly to a particular word or lexical item as a dependent of another particular item. For instance, it can say that the verb LONG requires the preposition *for* as its dependent; this statement refers to two words which are directly related in a dependency structure such as the one in Figure 7. In phrase structure, on the other hand, the same link (indicated by the arc below the words) is only indirect; to be precise, it relates *longed* to its niece via three links. This structure obscures the fact that only the head of the phrase *for rain* is relevant to *longed*, so CG would in principle allows the combination LONG RAIN to be stored, whereas WG only allows LONG FOR.

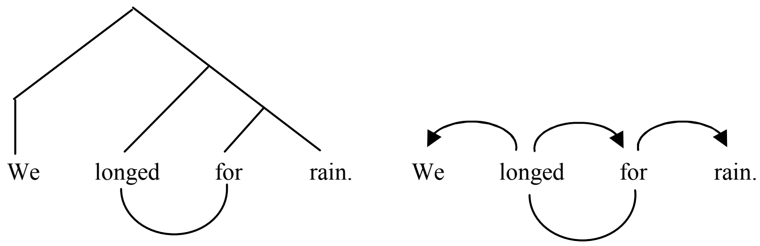

Figure 7: Lexical selection of for *by LONG in phrase structure and dependency structure*

The third advantage of adopting dependency structure in CG concerns semantics. In dependency structure, a word's dependent modifies that word's meaning – e.g. *big dog* refers to a kind of dog, and *eats ice-cream* to a kind of eating – so the phrase's head carries the meaning of the entire phrase. Of course, the modified word is distinct from its basic counterpart, so the *dog* of *big dog* is represented by a different node from the lexeme DOG, and similarly for *eats* in *eats ice-cream.* In each case, the specific word is a special sub-case of the general lexeme, with distinct syntax and distinct semantics.

This facility of locating phrasal differences on the head word is helpful when handling idioms and other stored phrases; for example, the meaning of KICK THE BUCKET is carried by what we can call $\text{KICK}_{\text{bucket}}$ (a special sub-case of the general lexeme KICK which has *the bucket* as its object and means 'die'). Similarly, to the extent that GIVE in "composite predicates" such as GIVE X A KICKING has a special meaning (Trousdale, this volume) this can be attributed to the special sub-case of KICK which combines with deverbal objects. In neither case is there any need to postulate a phrasal construction to carry the stored meaning. Indeed, dependency structure is a better basis than phrase structure for analysing idioms because it forces the prediction that the fixed elements in an idiom will always include the head word; so there could be no idiom such as "Verb *the bird the worm*", in which any verb could be combined with two lexically specified objects (Holmes and Hudson 2005).

Such idiomatic dependency structures can be combined incomplex memorized structures such as the one for the famous *What's X doing Y?* construction (as in *What's your car doing in my parking space?*), for which Figure 8 is the analysis proposed in Holmes and Hudson 2005.

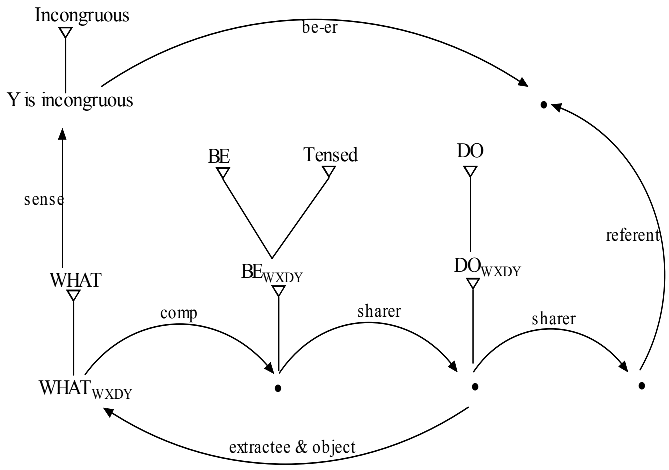

Figure 8. A WG analysis of the What's X doing Y? *construction.*

The fourth, and most important, advantage of dependency structure involves the intimate pairing of syntactic and semantic relations in dependencies. Every dependency is defined in part by its meaning, so it is just right for mapping onto semantic roles (as in a CG construction). In WG, relations are concepts and, like any other kind of concept, they are learned inductively from bundles of features which tend to co-occur – in other words, which correlate positively with one another. When the relations concerned are between pairs of words, the correlated features tend to include both semantic relations between the words' meanings (semantic roles) as well as syntactic relations between the words themselves such as word order or agreement. Indeed, in some European versions of dependency grammar, dependencies are exclusively semantic and have nothing to do with such superficial syntactic considerations as word order (Bröker 1998, Heringer 1993, Kunze 1975, Tesnière 1959, Sgall and others 1986), though

I know of no theoretical justification for this restriction. The essential requirement for recognising a dependency between two words is that there should be at least two relations between them which tend to co-occur, and it makes no difference whether these relations are deeply semantic or superficially syntactic. However, dependencies do typically combine meaning with one or more syntactic relation such as word order, which makes them particularly suitable for pairing with meaning in a construction. Indeed, the semantic relations which are part of the definition of the dependency relation duplicate the pairing achieved by a construction, so in a sense the construction is already part of the dependency and a separate construction is redundant. For instance, the definition of the syntactic dependency "indirect object" includes a pairing with a semantic role such as "recipient", just as in the CG conception of a construction.

Given a network analysis in which syntactic structures consisting of meaningful dependencies between words are mapped onto semantic structures, what extra role is left for constructions? One possible answer is suggested by Gisborne (this volume): all the work done in CG by phrasal constructions can be left to individual dependencies, linking individual words. However, if this is right (as I believe it is), no work is left for phrasal constructions in the CG sense. At this point we might consider the experimental evidence for the psychological reality of constructions (Goldberg 2006:120-5). One particularly impressive type of evidence involves "structural priming", where a syntactic pattern is made easier to retrieve ("primed") by earlier occurrences of the same pattern, providing a convenient psychological test for what counts as "the same pattern" in syntax. What emerged from the experiments is that structural priming requires similarity of meaning as well as of syntax (Goldberg and Bencini 2005); for example, one experiment used stimuli like the following:

(1) *The 747 was landed by the pilot.*
(2) *The 747 might land near the airport control tower.*

The significance of these examples is that they have similar "surface syntax" if we simply consider the words and phrases and their categories; but their very different semantic structure was enough to prevent either from priming the other. In contrast, one passive sentence such as 0 does prime another even if the lexical content of the two sentences is completely different.

How we interpret these results depends on what assumptions we make about syntactic structure. If it consists of a phrase structure which is so superficial that 0 and 0 have the same syntax, the experiments do indeed show that syntax alone is relevant only to the extent that it is paired with semantics – in other words, that we think in constructions. But this would be a very odd conclusion, because it rests on the assumption that we assign a syntactic structure which is inadequate. Why should we do anything so silly? Oddly, this assumption is not even part of CG itself, because the syntactic pole of a construction is defined in terms of abstract and semantically relevant grammatical functions such as Obj1 and Obj2. Surely sentences 0 and 0 would have different syntactic structures even in CG, let alone in WG where syntactic structure consists of dependencies and a superficial structure is impossible? What the experiments show, therefore, is that we think either in the functionally-labelled phrase structure of CG or in the dependencies of WG. Even given the assumptions of CG, therefore, they do not show that we think in constructions.

In conclusion, then, I believe that syntactic structure does in fact consist of dependencies, each defined in terms of a mixture of semantic and syntactic features, so the same should be true for the syntactic pole of a CG construction. But accepting dependency structure also means that the constructions of CG are too large, because they presuppose phrases when in fact there is nothing in syntax longer than a word. Applying this conclusion to the ditransitive construction, the crucial elements are the indirect-object dependency and the verb, so these are all that remains of the CG "construction".

4. Semantic frames are part of the analysis

The idea of semantic frames is very familiar in CG but semantic frames are rarely spelled out formally in the way that, say, argument structures are. The main insight of Frame Semantics is that word meanings must be studied in relation to "a structured background of experience, beliefs or practices, constituting a kind of conceptual prerequisite for understanding the meaning" (Fillmore and Atkins 1992:75). For example, the meaning of *Monday* is best explained in terms of the cycle of days in a week, and that of *elbow* in terms of the structure of an arm. In contrast with old-fashioned semantic fields, these frames consist of concepts – i.e. meanings – rather than words; so the definition of Elbow (the meaning of the word *elbow*) relates it to the concepts Arm, Hand and so on. Crucially, the relations among these concepts must "go far beyond anything envisioned by current theories of thematic roles or deep cases," so "we need frame-specific semantic role categories" (Fillmore and Atkins 1992:84). For instance, the relation between Monday and Tuesday might be "Day-after", a relation which only applies to the days of the week.

These ideas have been developed more fully in WG than in CG (or, for that matter, in any other theory of language structure that I am aware of). On the one hand, the principle of "re-cycling" guarantees that the meaning of one word should be defined where possible in terms of the meanings of other words, rather than in terms of some kind of universal basic vocabulary of concepts and relations (Hudson and Holmes 2000). For example, the paper just cited offers an analysis of the concepts expressed by lexemes such as BICYCLE and CYCLE, in which very specific concepts such as Pedal are used in the analysis of several other concepts (e.g. as parts of Bicycle and also as the moving element in Pedalling). Re-cycling means that every concept in the network is defined solely by its relations to other concepts, so the analysis of word meaning cannot be separated from the analysis of the speaker's complete conceptual structure. Of course this means that a complete analysis is impossible, but this difficulty is amply outweighed by the fact that partial analyses are relatively easy. All that is needed is careful thought and sensitivity to semantic facts, rather than a rigid universal framework of categories and relations. The result is analyses which can be richly articulated and very precise (Gisborne 1996, Holmes 2005).

On the other hand, the idea of frame-specific semantic role-categories is very easy to accommodate thanks to the way in which WG assumes that relations are learned. As I said in section 3, "relations are concepts and, like any other kind of concept, they are learned inductively from bundles of features which tend to co-occur – in other words, which correlate with one another." Like other concepts, relations may have any degree of generality from the most specific relation between two tokens of experience to the most general; for example, the string of letters that I am writing now have mutual relations which are a particular case of the relations between letters in a string, which in turn exemplify relations between objects in a string, which involve the most general relations such as Before or After. Each of these relations has a place in the network and most (possibly all) may be learned from experience.

WG even offers a fairly specific theory of how new concepts (including relations) are learned (Hudson 2007a:52-9; for similar theories applied to syntax, see Ninio 2006, Chang and others 2006, and Tomasello and Bates 2001:163-290). According to this theory, we are continually enriching our mental networks by adding new links, each representing a new relation, and we do this in two ways: either by observing and remembering a token of experience, or by inducing a more general relation from a collection of existing links whose properties correlate. These two kinds of growth are opposite sides of the same coin of "exemplar-based" learning which Goldberg discusses in some detail, and which combines the storage of individual exemplars with the drawing of generalizations (Goldberg 2006:44-65). The resulting potentially vast collection of links is held together by the same kind of inheritance hierarchy as is assumed in CG (Goldberg 2006:13), in which specific concepts "inherit" properties from more general ones. On the basis of the usage that we experience, plus the inductions we draw, we construct a hierarchy of relations in which each lower relation "isa" at least one other relation whose properties it inherits by default; so for example my mental network for the relations between letters in front of me isa Inter-letter relation which isa Inter-string-member relation which isa Before. It is a question of fact (and research) how these relations are distributed across languages and even across speakers of the same language; but the main point is that new relations are easily created either by the language learner or by the analytical linguist. Moreover, the usual caveat about inheritance applies: inheritance does not preclude storage. Indeed, if generalizations are built by induction from stored exemplars, we can be sure that some stored exemplars are stored with properties that they could, in principle, inherit from a super-category, so WG rejects what Hoffmann (this volume) calls "complete inheritance models" in which storage is minimized.

The ideas of re-cycling and relation-creation are important in WG because they affect the analyses that are produced. CG analyses typically invoke specific relations without defining them, but WG analyses tend to define them by spelling out the relevant inheritance hierarchy – i.e. the "semantic frame". This difference between the theories may show nothing more profound than different interests and priorities of the researchers, but it may also have something to do with notation. The network notation of WG encourages the analyst to explore semantic frames, whereas this would be much harder in the notation of CG, which strikes me as rather rigid and cumbersome. For example, I have used WG networks to analyse the semantic frame of commercial transactions which defines the relations needed in the semantics of the verbs BUY, SELL, CHARGE, SPEND, PAY and COST (Hudson 2008). In this analysis, the meanings of these verbs are related to the meanings of TRADE, CONSUME, USE, GIVE and GET, and the analysis invokes a range of different relations ranging from very general relations such as Part and Time to frame-specific ones like

Giver and Gift. But crucially, the more specific relations are not only invoked, but they are also defined by the frames in which they are introduced and from which specific instances inherit properties. In short, the frame is part of the analysis.

As in previous sections, we can use the English ditransitive construction to show the benefits of these general ideas. Figure 4 has already presented a direct translation of this construction from CG into WG notation, but the purpose at that point was simply to introduce the WG notation. WG actually allows a much deeper analysis in which the semantic roles are "unpacked" into a structural analysis. Details of this analysis can be found in Holmes and Hudson 2005, which also gives WG analyses for a number of other constructions including the *What's X doing Y?* construction. (A great deal more discussion of constructions can be found in Holmes 2005.) The following discussion extracts the most relevant points.

Ditransitive constructions are interesting and challenging because a single syntactic pattern (the "double-object" pattern discussed in detail in Hudson 1992) expresses two different meanings, called "to" or "for" according to whether they can be paraphrased as in 0 or 0.

(3) *She gave her friend a present.* = She gave a present to her friend.
(4) *She found her friend a present.* = She found a present for her friend.

A handful of ditransitives do not allow either kind of paraphrase; for example, the verbs ASK and ENVY are hard to paraphrase in this way:

(5) *She asked her friend a question.* = She asked a question *to/*for/?of her friend.
(6) *She envied her friend his wealth*. = She envied his wealth *to/*for her friend.

We shall return below to these awkward cases. A second contrast is one based on lexical selection or its absence. For example, GIVE clearly selects its indirect object, but KICK equally clearly does not; in terms of typicality, GIVE is a typical ditransitive verb, whereas KICK is a typical two-participant transitive verb. And yet, KICK can in fact be used with an indirect object:

(7) *She kicked her friend the ball*. = She kicked the ball to her friend.

This contrast cuts across the one between the "to" and "for" meanings, so unselected indirect objects may be paraphrased either by *to,* as in 0, or by *for*:

(8) *She baked her friend a cake*. = She baked a cake for her friend.

Table 1 presents examples of the four intersecting cases; my evidence for claiming that the indirect object is lexically selected by both GIVE and FIND is that this pattern is listed for these two lexemes in at least two modern dictionaries (Anon 2003, Sinclair 1987), neither of which lists it for either KICK or BAKE. (I recognise that it is very hard to apply to particular cases, but I argue below that this does not matter for the general point.) The challenge, then, is to produce an analysis which allows the full range of possibilities, and I shall suggest that the solution is to include the semantic frame in the analysis.

	Lexically selected	Unselected
'to'	She gave her friend a present.	She kicked her friend the ball.
'for'	She found her friend a present.	She baked her friend a cake.

Table 1: Four kinds of ditransitive constructions

We start with the plausible assumption that the lexeme GIVE provides the model for other ditransitive verbs. Once we have a semantic structure for this verb, we shall see that all the other ditransitives are sufficiently similar in their semantic structures to explain why they all use the same syntactic structure. What exactly does GIVE mean? If I give you a pound, then:

- Before the giving, the pound belonged to me.
- After the giving, the pound is yours.
- The giving is a deliberate action.
- The change of ownership is its intended result.

In other words, the semantic frame for Giving (the sense of GIVE) includes ownership, action and causation. In prose, the analysis that I offer is as follows, where I use the verb *isa* for the classification relation (as in "Dick isa linguist"):

- Giving isa **Doing** – an event which has a **do-er** and a **purpose**.
- Giving isa **Making** – an event which has a **patient,** a **result** and a **source** related as follows:
    - the result and source are two complementary states,
    - the source is replaced at the time of the making by the result,
    - the patient is the "**er**" (e.g. sitt-er, sleep-er, be-er) of both states.
- The result of Giving isa **Belonging**, and likewise for the source. (I explain below why this state is called "belonging" rather than "having".)

The claim is that all the highlighted relations and event-types are part of the semantic structure of GIVE, so they are all available for explaining the uses of the ditransitive construction.

The WG notation allows us to express the analysis in a way which combines formal precision with psychological plausibility; after all, if language really is a cognitive network, what better notation for it than a collection of nodes and links? To avoid overload, I shall introduce the analysis in stages. We start with the very simple relation between Giving and Doing (the typical purposeful action), which explains why Giving has an "er" (i.e. a giver) and a purpose: even if these relations were not stored for Giving, as they surely are, they could be inherited from Doing. This part of the frame is shown in Figure 9. The most important general point in this diagram is that each of the relations that Giving has are classified by an "isa" link which allows it to inherit whatever properties may be associated with the more general relation at the other end of the isa link (the end where the small triangle is). The diagram does not try to show what these properties might be, but they would certainly include the cluster of relations that define "purpose" in terms of Wanting and Controlling by the "do-er".

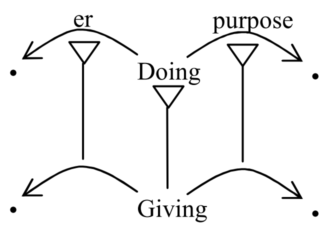

Figure 9: Giving isa Doing

The properties of Making are more complex because they involve a change from one state to another; for example, if I make a cake, the cake changes from not existing to existing, and if I make it cool, it changes from not cool to cool. This is what Figure 10 shows. As before, the links to Giving can all be inherited from Making, but for simplicity I have only included one isa link from a relation of Giving to one of Making. In prose, the diagram says that Making has a source, a patient and a result, such that the patient is the "er" of both the source and the result, and the result isa the source. Thanks to default inheritance, this isa link allows one or more inheritable properties of the source to be overridden, but it does not allow the source and result to be completely unrelated states – e.g. a change from being poor to being tall. For simplicity, I have omitted the time relations whereby the Making is after the time of the source and before that of the result.

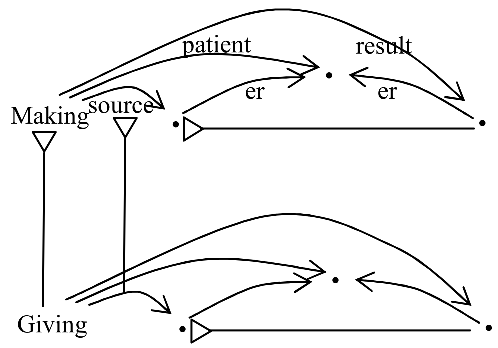

Figure 10: Giving isa Making

Finally we have the fact that giving has to do with ownership, which in this analysis is expressed in terms of a state of belonging. In giving, the gift passes from the giver to a second person, the receiver, so both the gift and receiver are affected; so which of them is the patient? It might have been the receiver, but in view of the general structure for Making, it has to be the gift because this is the participant which is er of both the source and result states – which is why this must be "belonging" rather than the more obvious "having". This analysis is due to Holmes (Holmes 2005:139-45), who suggests that what I am calling "belonging" is actually the primitive spatial relation 'at', the meaning of the preposition *at*. One of the attractions of this analysis is to explain the alternation with *to*, which generally refers to an 'at' which is a result. This change of ownership is shown in Figure 11, which also defines a new relationship (recipient, the second owner) and identifies the first owner with the giver.

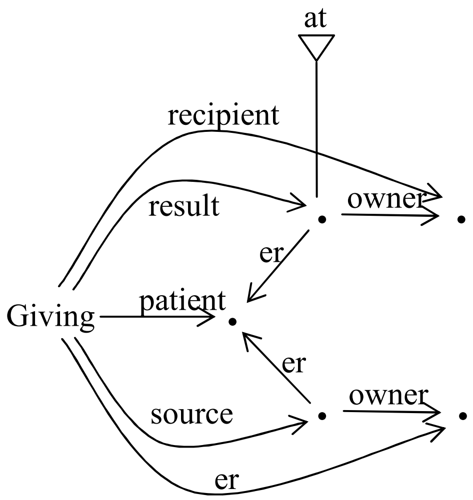

Figure 11: Giving causes a change of ownership

This completes the semantic analysis of Giving, which I assume is also the area of general cognition in which we handle the idea of giving even when we are not talking about it – e.g. in thinking about giving someone a present. All that remains is to show how this cognitive structure can be put into words. Since the concept Giving is the sense of the verb GIVE, the question is how the various semantic relations of Giving map onto the syntactic dependents of GIVE. We must allow two possible mappings: the ditransitive and the prepositional. In order to accommodate both possibilities we distinguish two sub-lexemes of GIVE, each of which isa GIVE: $\text{GIVE}_{\text{ditransitive}}$ and $\text{GIVE}_{\text{prepositional}}$. (These must be sub-lexemes of GIVE rather than distinct lexemes because they have the same sense and the same mappings for the subject and direct object, so the only differences lie in the syntactic expression of the recipient; sub-lexemes are an important part of any WG analysis, and no doubt GIVE has many other sub-lexemes, such as the one for mono-transitive uses as in *He gave a pound*.) I shall take the relations Subject, Object and Indirect-object for granted, but they too can be defined in much the same way as the semantic relations we have considered. These mapping relations are shown in Figure 12 (where all but one vertical arrows link a word to its referent, the exception being the one between GIVE and its sense; the distinction between senses and referents is familiar, but not relevant here).

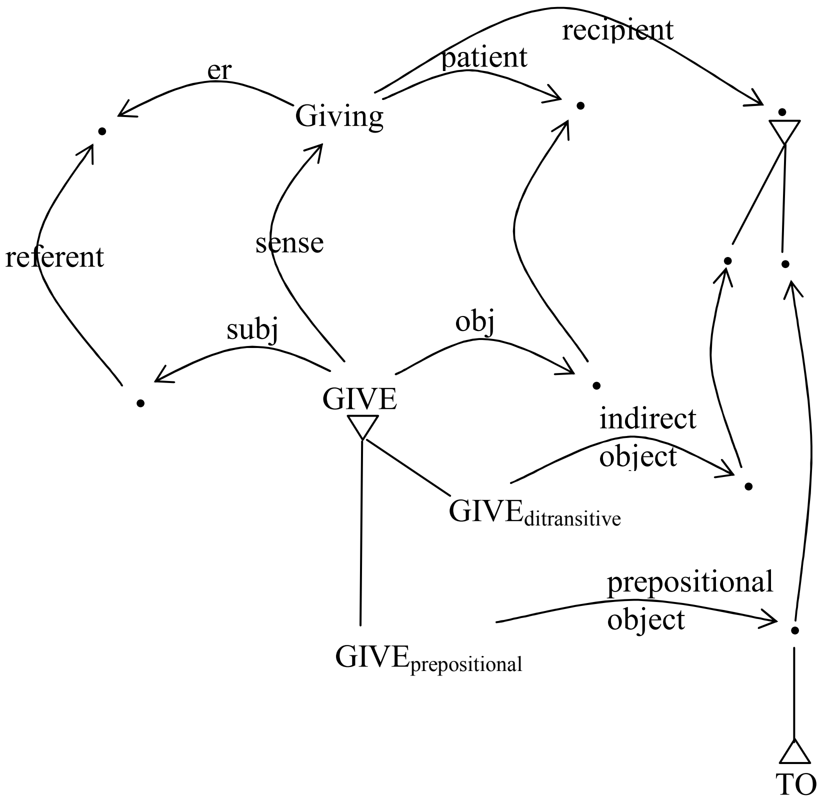

Figure 12: The two syntactic options for Giving

Now that we have a complete analysis of Giving and the verb GIVE, we can return to our investigation of the ditransitive construction. I can now show how the rich semantic frame that we have built for GIVE helps us to generalise from this example of a lexically selected “to” pattern to the other three patterns in Table 1. These generalisations explain why they all have the same syntax in spite of different meanings and different selection.

We start with the "for" ditransitives such as 0 *She found her friend a present*. What exactly is the difference in meaning between finding and giving such that one is paraphrased by *to* and the other by *for*? And equally importantly, what do they have in common such that they both use the indirect object? The relevant part of the semantic frame is the network in Figure 11, which defines the relation in giving between the source state and the result. What "giving X to Y" and "finding X for Y" have in common is that in both cases, the outcome is that Y has X – or in our terms, X belongs to Y. In short, they both have the same result structure. Putting this in functional terms, if I want to express the idea that applying some action to X puts X into your possession, an obvious model is the structure for Giving which maps X onto the direct object and you onto the indirect object. On the other hand, the structure for Giving cannot be inherited lock, stock and barrel because in finding, in contrast with giving, the finder does not start by owning X. Rather, the "source" (the start-state) of Finding is undefined, in contrast with Giving where it is the giver that owns the patient. This is why finding is a less typical example of the ditransitive construction than giving.

What about the verbs such as ENVY and ASK which have no prepositional paraphrase? Is their syntactic similarity to GIVE merely arbitrary, or is there some semantic similarity which motivates it? ENVY is easy because I can only envy you what belongs to you; so I can envy you your success, but I don't think I can envy you your enemy's demise. But of course your success is not the result of my envying, so extending the indirect object to ENVY is really stretching the model to its limits. What ENVY and GIVE have in common is merely the fact that the direct object belongs to the indirect object; but the convenience of the syntax presumably outweighs the semantic anomaly. ASK works the other way round: if I ask you a question, the question is mine initially, but becomes yours in the sense that it is your responsibility to answer it. (This is even clearer with SET; if I set you a question, it becomes your problem, not mine.) The conclusion, then, is that the use of indirect objects with ENVY and ASK is motivated by partial similarities between their semantic structures and that of GIVE. However, my reason for discussing these examples is that these similarities are visible only if we can explore the whole of the semantic frame. If the only semantic structure available for each verb is one layer deep, so to speak, the verbs have no particular semantic similarities and their syntactic similarities are inexplicable.

The second challenge from Table 1 is to explain why indirect objects are available even when there is no lexical selection, as may well be the case with verbs such as BAKE and KICK. For these verbs, the ditransitive construction is quite marginal in contrast with the central role it plays for GIVE; indeed, there are good reasons for thinking that our first encounter as children with this construction is precisely with the verb GIVE (Goldberg 2006:77). Lexical selection seems to be a matter of degree, with verbs spread across a spectrum ranging from completely lexicalized (GIVE) to completely unlexicalized neologisms such as FAX or TEXT as in *I texted him the news* (Holmes 2005:258). It is very hard, and perhaps even impossible, to know for sure whether or not any given verb lexically selects an indirect object (Croft 1998), so I am merely guessing that BAKE and KICK do not; but nothing hangs on this choice, as the main point is simply that the ditransitive construction may be applied to verbs that we have not previously heard in this pattern.

What kinds of verb will accept this innovative extension of the ditransitive pattern? Levin 1993 lists ten classes such as “verbs of future having”, “verbs of throwing” and “verbs of transfer of a message”, but it may be possible to explain this apparently random collection if we define the extension correctly. In that case, all we need to say is that the extension is possible for any verb that is syntactically and semantically compatible with it. The verb must have compatible syntax which allows a direct object, because the indirect object is by definition one which co-occurs with a direct object and which combines its referent semantically with the latter’s. But equally importantly, the verb’s semantic frame must be compatible with the construction’s meaning; and to be somewhat more precise, it must define an action applied to a patient which puts the patient into the ownership of the recipient. Baking and kicking pass this test: baking brings the object it into existence and kicking moves it; but some verbs do not pass. For example, opening a door does not change the door in such a way that it could belong to someone, so we cannot open someone the door, although we can open the door for someone (Holmes 2005:52); moreover, we probably can open someone a can of beans, because this puts the beans into their possession. It is unlikely that this kind of information could be stored lexically, so it must follow from more general principles; but if the general principles involve semantic compatibility, then they must involve the entire semantic frame of the verbs concerned rather than a simple semantic structure such as the argument roles that CG envisages. These ideas need to be developed more fully and formalised, but I believe that they promise a much better explanation than one based on an arbitrary list of verb classes.

This analysis of the ditransitive construction raises a serious question about the logic of inheritance: How can ordinary transitive verbs like

BAKE and KICK inherit the properties of a ditransitive without being classified as ditransitive? I return to this question in section 6. Meanwhile, I hope to have shown that the semantic structures of CG are not rich enough to express some important generalisations, and that the analysis needs to include the entire semantic frame. This is already the practice in WG, and I see no reason in principle why the same should not be true in CG.

5. Levels of organisation are autonomous

In CG, a construction is defined as a "conventionalized pairing of form and function" (Goldberg 2006:3), and it is constructions, rather than the forms and functions themselves, which are the basic organizational units of the grammar; in Goldberg's words quoted in section 2, "it's constructions all the way down." Taking this claim literally, formal patterns and meaning patterns are not themselves part of the network of grammatical knowledge. Instead, each network node is a complex of information consisting of a formal structure, a meaning structure and the mapping between the two – something very much like Figure 3, the ditransitive construction in CG. This is a very different view of knowledge structure from the one in WG, where the corresponding slogan is that "it's networks all the way down" (Hudson 2007a:232); in this respect, WG follows Stratificational Grammar (Lamb 1966, Lamb 1998). In a WG network, the nodes are merely nodes where arcs meet, and can never be "boxes" full of information which is not itself part of the network.

One particular manifestation of this difference is in the role of inheritance hierarchies, which apply in CG only to constructions (Goldberg 2006:13-14). This is similar to the first version of Head-driven Phrase Structure Grammar, where attribute-value matrices are "typed" (i.e. related in an inheritance hierarchy), but the individual attributes and values are not (Pollard and Sag 1994:20). In contrast, more recently HPSG has moved towards WG by allowing the individual elements to be typed as well (Sag 1997). This is an important improvement because it allows generalisations which (at least when taken literally) the principles of CG do not seem to allow – generalisations across grammatical functions or across semantic functions; or indeed, generalisations across classes of words or phrases or across semantic classes.

To return to our example of the ditransitive construction, the CG analysis invokes three grammatical functions labeled "Subj", "Obj1" and "Obj2", and three argument roles labelled "agt", "$\text{rec}_{\text{secondarytopic}}$" and "theme" (Goldberg 2006:20). As far as the argument roles are concerned, these are not "drawn from a universal or limited set" (Goldberg 2006:20), because they are "defined in terms of the semantic requirements of particular constructions" (ibid:39). By implication, a role such as "agt" is defined afresh for each construction where it is mentioned; but how? How can we define a specific role such as the "donor" role without referring to other roles and ideas which go beyond the specific construction in question? The WG answer is that we cannot because (as explained in section 4) each semantic relation is defined by the total semantic frame, including the general event types which help to define the action concerned. As far as CG is concerned, given the importance of frame semantics in CG, we might expect the same reply there too; and yet participant roles seem to be left entirely undefined. Of course there is a limit to what can be covered in a single book, but in this case the problem goes deeper, and touches on the basic assumption that grammar "is constructions all the way down". If constructions are the only elements that may be related to one another, it follows that this is not possible for the elements of a construction such as the argument roles and consequently these roles are inherently undefinable.

Similar questions arise for the grammatical functions which comprise the formal pole of a phrasal construction, such as "Subj", "Obj1" and "Obj2" in Figure 3. Each of these labels summarises a bundle of properties, some formal and others semantic – details of word order, agreement, reflexivization, relativization, agenthood, animacy, topicality, and a number of other properties which tend strongly to correlate both cross-linguistically and inside each language (Keenan 1976, Dowty 1991); for example, as Rosta argues (this volume), lexical subjects tend to be more agentive than any other participant. Somehow, somewhere, the analysis of a language needs to show how the relevant properties are inherited by the grammatical function labels. In CG the only mechanism for inheriting such information is to treat it as a property of an entire construction, because it is only constructions that can inherit; in the present case, this presumably means treating the properties of subjects and objects as properties of more schematic "clause" constructions, with (presumably) a different clause type for each grammatical function. In contrast, WG allows each relation to inherit its individual properties directly from more schematic relations, without invoking a parallel series of clause types. The question is whether constructions improve the analysis, and at present we have no evidence that they do. It is simpler to inherit the properties of grammatical relations directly from more general relations than to inherit them from phrasal constructions where they are combined with semantic relations.

In short, by treating meaning and form as inseparable sides of the same coin, constructions prevent generalizations from being made about one without the other, so constructions are too large in terms of levels. (Similarly, section 3 concluded that constructions were too large in terms of length, because phrases should be replaced by single words and dependencies.) This objection rests on the traditional assumption that each "level of analysis" (which I prefer to call "level of organization") has its own organising principles and generalisations which need to be stated in addition to whatever correspondences there may be between levels. Even if there were only two levels ("form" and meaning), this would still be true because the levels are organised quite differently – e.g. linear order is crucial to form whereas meaning is organised as an unordered network. Each level has its own inheritance hierarchies and its own rules for building structures (Lamb 1966, 1998); and each level is mapped onto its neighbours by "correspondence rules" (Jackendoff 1997) or relations of "meaning" or "realization" (the WG terminology). This is the WG view of how language is organised (Hudson 2007a:72-81), but CG seems to reject it by its exclusive focus on constructions as the only units of language structure. However important cross-level correspondences are, it is just as important to be able to analyse the combinatorial patterns found in form and meaning, and where they are autonomous, to treat them independently.

Another undesirable consequence of taking constructions as basic is the implication that there is only a single level of "form", in contrast with a more traditional view in which linguistic structure can be factored into a series of levels including phonology, morphology and syntax. Constructions are "pairings of form with semantic or discourse function, including morphemes or words, idioms, partially lexically filled and fully general phrasal patterns" (Goldberg 2006:5). In this view, morphemes, words and phrases exist on the same level and differ only in size, just as in classical American structuralism where morphemes were the basic constituents of a single level of "grammar" (Matthews 1993:82).

The objections to this view are well-known thanks to the long debate dating back at least to the classic defence of the European "Word-and-Paradigm" model in which words and morphemes exist on different levels (Robins 2001), and to my mind the objections are overwhelming. For one thing, morphs need not be paired with a meaning; for example, there is no obvious meaning (however broadly we interpret this term) to be paired with the morphs that can be isolated in *perceive* and *deceive*, or with inflectional and derivational suffixes such as *–ing* and *–er*; and conversely, there is no obvious morph to be paired with the individual meaning contrasts in words like Latin *amo*, where the suffix *–o* indicates present tense, indicative mood and a first-person singular subject. Many morphologists (e.g. Stump 2001, Aronoff 1994) adopt the "Word-and-Paradigm" solution which separates grammatical structure into two levels: morphological and syntactic. Morphological structure (which I call simply "form" - Hudson 2007a:74-8) consists of morphs which realize words either singly (e.g. *cat*) or in combination (*cats*), so morphs pair phonological structures with words. In contrast, syntactic structures consist of words which realize semantic structures both individually and in combinations, so what words pair is morphs and meanings. In this analysis, words and word-combinations can reasonably be described as "constructions" because they pair a "form" with a meaning; but morphological units have no meaning, so they do not qualify as constructions. This is not a mere quibble about the meaning of the term "construction", but a rejection of the fundamental claim that there is nothing in language except constructions. Apart from the patterns which can be paired directly with a meaning, there are many which cannot, so to identify grammar with constructions is to ignore not only morphology, but also phonology.

The traditional multi-level analysis is supported by evidence from psycholinguistics. I argue in section 6 that a network carries activation, so evidence for activation patterns throws light on the internal organisation of the network. If there are separate sub-networks for phonology, morphology, syntax and semantics, these should show up in the psycholinguistic evidence; and this is precisely what we find. There are two main kinds of evidence for detailed activation patterns, both of which arise because activation spreads blindly, spilling over from active nodes onto their neighbours. Under normal processing conditions the spilt activation is mopped up more or less efficiently, but its effects can be seen under two special conditions: experiments on "priming", and speech errors. In both cases, we can conclude that because node A's activation has spread onto node B, A and B must be near neighbours in the network. The evidence does not, of course, tell us what the relation between A and B is, but it does reveal some kind of fairly direct relation. This kind of evidence is important for any discussion of the architecture of the language network, so I now review the two kinds of evidence.

Priming experiments take place in front of a computer screen which presents a stimulus word A and then another word B, which requires some kind of processing (such as a decision about whether or not it is an English word). The crucial question is how long the experimental subject takes to process word B (when measured in milliseconds – the effects are very slight). If A is related to B (e.g. *doctor*, then *nurse*), the delay is measurably shorter than if the two words are unrelated, so A is said to "prime" B; and if A primes B, we can be sure the two are closely linked in the mental network. But of course there are all sorts of different kinds of links between words, so it us up to the linguist to classify the links. As predicted by WG, we find priming at each of the levels of organisation:

- purely phonological priming (e.g. of *worse* by *nurse*; see Brooks and MacWhinney 2000, James and Burke 2000, Norris and others 2002, Spinelli and others 2001),
- purely morphological priming (hard to illustrate in English, but demonstrated for Hebrew – see Frost and others 2000),
- purely syntactic priming (e.g. by one passive priming another; see the discussion of 0 and 0 in section 3),
- purely semantic priming (e.g. of *doctor* by *nurse*; see Beer and Diehl 2001, Moss and others 1995, Perea and Rosa 2002, Smith and others 2001).

These separate types of priming are as expected in WG, given the network links between elements on the separate levels of phonology, morphology, syntax and semantics; but CG predicts only one kind of priming: from one entire construction to another. It remains to be seen whether there is any evidence for constructional priming which could be separated from syntactic priming.

Moreover, the evidence from priming is supported by observational evidence from speech errors, where a node which is topologically near to the target happens to get an unintended boost of activation which pushes it to the head of the "production queue" in the speaker's mind. For example, when Dr Spooner famously said "Young man, you have tasted the whole worm" (instead of "you have wasted the whole term"), we may assume that the activation of the planned *term* spread, via its first phoneme /t/, to combine with that of *wasted* to select *tasted* instead. Speech errors confirm the existence of activation at individual levels:

- purely phonological, e.g. *orgasms* for: *organisms* (Aitchison 1994:38)
- purely morphological, e.g. *slicely thinned* (for: *thinly sliced*) (Levelt and others 1999)
- purely syntactic, e.g. *I'm making the kettle on* (for: *making some tea* + *putting the kettle on*) (Harley 1995:355)

- purely semantic, e.g. *fork* (for: *spoon)* (Harley 1995:352)

Once again, these results are not predicted if every network node pairs a "form" with a meaning.

My conclusion, therefore, is that the notion "construction" is not needed in grammar. Moreover, it is not merely redundant: it causes problems in grammatical analysis because it implies boundaries that are impossible to justify. If grammar consists of constructions, each construction ought to have boundaries; but where are they? We cannot define a construction as a unique pairing of form with meaning because there are polysemous or even homonymous constructions such as the word FOR in 0 and 0 (Goldberg 2006:38).

(9) *The statue stood for three hours.*
(10) *He exchanged the socks for a belt.*

But how can a single construction allow two different meanings if meaning is an essential pole of the pairing? And if this is possible, how is a construction different from a mere form? No doubt synonymy raises similar questions: if two formal patterns share the same meaning, do they therefore belong to the same construction? Such questions are familiar from the literature on "lexical items", and as in that debate, the only reasonable conclusion may be that we are asking the wrong question because both "construction" and "lexical item" are inventions rather than discoveries.

6. The cognitive context

As I mentioned earlier, both CG and WG are part of the general "cognitive linguistics" movement. In my opinion, the most important tenet of this movement is the one quoted in section 1: "Knowledge of language is knowledge" (Goldberg 1995:5). For most of the twentieth century, the dominant model of language structure was structuralism, which emphasised not only the internal structure of language but also its external separateness. In the early years this isolationism was good because it insulated linguistics from some unhelpful ideas from pseudo-psychology, but now that psychology has grown up it is time to build on its theories. Of course this is precisely what psycholinguistics does in its investigations of language learning and use, but if linguistic knowledge is controlled by the same principles as the rest of knowledge, this should also have consequences for linguistics, the theory of language structure. In particular, whatever psychologists think about the structure of general knowledge should generalise to the structure of language and to theories of language structure such as CG and WG. Moreover, even matters of language use (and learning) are relevant to these theories because the structures proposed for language must be compatible with whatever we know about use and learning. I think these general principles are shared by CG and WG.

However, it is easier to accept the principles than to translate them into a fully worked-out theory of language structure, use and learning in the light of the best current psychological work. This kind of theory is what I mean by the "cognitive context" for linguistic theory. My impression is that cross-disciplinary work in this area is weak, as psychologists and linguists generally ignore one another's theories and live in different cultural worlds. We all have a lot of bridge-building to do. I admire the bridges that Goldberg has already built towards experimental psycholinguistics (Goldberg 2006), but WG includes some bridges that are still missing in CG and which start from some of the most elementary and uncontroversial ideas of cognitive psychology:

- spreading activation
- universal inheritance hierarchies
- best-fit binding

(Apart from spreading activation, these names need the explanations that I shall provide below.) I believe that all these ideas could be incorporated into CG with very little change to the rest of the theory – indeed, I shall suggest that some of them are already implied by the existing tenets of the theory.

6.1 Spreading activation

Spreading activation is certainly one of the implicit ideas of CG, although the term does not appear in the index of either of Goldberg's books. As I mentioned in section 5, Goldberg rightly sees a link between language structure and priming, the phenomenon whereby one token makes a later token easier to process. As Goldberg argues, priming results illuminate the structure of the language network by giving a measure of "topological distance" – the distance between two nodes measured in terms of the number of intervening links – so if sentence A does prime sentence C but B does not, this must be because A is topologically closer to C than B is.

One consequence of priming for language structure is that nodes must be associated with an activity level; this idea is already accepted at least in principle in CG in the sense that constructions are expected to have a degree of "entrenchment" or "token frequency" (Goldberg 2006:93) which reflects the number of times they have been heard. The main determinants of activation level are frequency and recency (including the very short-term recency of priming experiments), so CG already has a place for activation on a single node. Unfortunately, the WG theory of activation is not much more advanced than the CG one, but it has two possible contributions to make.

One is the idea from section 3 that "relations are concepts". This means that most relations are carried by network nodes (rather than by links, as one might expect); for example, a syntactic dependency link such as "subject" passes through an ordinary network node which itself has two elementary links (called "argument" and "value") to the verb and noun nodes. As I shall explain more fully in section 6.2, relational nodes are also related by inheritance hierarchies so they have "isa" links to one another; so in our example, the "subject" node has isa links upwards to "dependency" and downwards to particular kinds of subject, such as inverted subjects. As I explained in section 4, this kind of analysis solves the problem of defining relations as needed not only in semantics but more generally. However, it also means that relations can carry activation in just the same way as non-relational nodes. For example, when listening, the relation "meaning" will be highly active so the focus of attention will be on meaning, but in a discussion of etymology the "etymology" link (for those of us who have one) will be much more active than usual.

The other possible contribution of WG is that the locus of activation is the individual node rather than something larger which we might call a construction. This suggestion repeats the negative conclusion from other sections that constructions cannot be distinguished from their parts, and that the nodes of a cognitive network are in fact individual words, morphs, syntactic relations and so on, rather than constructions. The conclusion may be premature, but at least the issue deserves more debate.

6.2 Universal inheritance hierarchies

The second idea from WG is that everything is part of an inheritance hierarchy. Of course the notion of inheritance is already well established in CG, but it only applies to constructions. In elementary cognitive psychology, and even more so in Artificial Intelligence, inheritance is seen as the basic mechanism of generalisation in all domains (Luger and Stubblefield 1993:35), so it is strange to find it restricted as in CG. However, the idea of applying it more generally seems to lie behind CG theorizing: "Inheritance hierarchies have long been found useful for representing all types of generalizations" (Goldberg 2006:13). The idea of limiting inheritance to constructions is linked to the general isolation of the internal elements of constructions that I have noted in other sections. For example, there seems to be no provision for an inheritance hierarchy for either words or event-types, although both are standardly classified in hierarchies (e.g. BE isa auxiliary verb isa verb isa word, Giving isa Causing isa Action isa Event). Once again, it would be easy to change CG to accommodate the extra hierarchies, but as with the previous WG idea, the result would be to make constructions less distinctive because they would no longer be the only "typed" entities.

Pushing this idea further, however, WG not only generalizes inheritance hierarchies to all kinds of entity, but because "relations are concepts" (section 3) to relations (except the basic Isa relation itself and a handful of other primitive relations). For example, it is possible to treat the syntactic relation Indirect-object as a subcase of Object, which in turn is a subcase of Complement; and in semantics the Seller role might be a subcase of Exchanger which in turn is a subcase of Agent. One attraction of treating relations in this way is that if a relation is a particular kind of concept, it can be learned in just the same way as other concepts; and this being so, relational concepts can be learned in the same numbers as other concepts. In other words, we no longer need to feel bound to restrict the inventory of relations to a small universal set. This approach to relations seems to be exactly what is needed to accommodate the CG view of semantic roles, whether argument roles or the more specific participant roles:

> … because they are defined in terms of the semantic requirements of particular constructions, argument roles … are more specific and numerous than traditional thematic roles …. Participant roles may be highly specific and

are often unique to a particular verb's meaning; they therefore naturally capture traditional selectional restrictions. (Goldberg 2006:39)

6.3 Best-fit binding

The third WG idea concerns the treatment of tokens of experience – what CG calls "exemplars" (Goldberg 2006:47). The two theories agree about the importance of tokens in the grammar, as a vast collection of half-remembered experiences from which more general categories are induced (Hudson 2007a:54-5); so CG and WG seem to incorporate more or less the same theory of learning (which I sketched briefly in section 4). However, CG appears to say very little about processing, whereas WG has at least the beginnings of a general theory of how we understand our incoming experiences and how we plan our behaviour – i.e., in relation to language, about how we understand and how we speak (Hudson 2007a:41-52). Any theory of language structure ought to be paired with a theory of processing because the two kinds of theory are interdependent: language structure must be usable, and processes must be applicable to the structures that are actually found. For example, if processing involves spreading activation, then the language structure that we assume must be one that allows activation to spread freely across network links (as I suggested above).

To give a concrete example of how a theory of processing might help CG, let us consider again the ditransitive construction and the question that I raised briefly at the end of section 4 about how the process of inheritance applies the stored construction to particular cases. Every verb is itself a construction, so it has a place in the inheritance hierarchy of constructions which also includes the ditransitive construction itself (Goldberg 2006:14); and presumably each construction defines a different class of verb (e.g. the ditransitive construction defines the verb-class Ditransitive), so the hierarchies of verbs and constructions are linked. Consequently, it is easy to see how inheritance applies the general construction to a verb such as GIVE which isa Ditransitive; but what about a verb such as KICK and BAKE? The point of the examples is that they are not stored under Ditransitive, so how can they inherit from this construction? In general, A inherits from B only if A isa B (or more generally if there is a chain of isa links from A to B); so how come a verb which has no such isa relation to Ditransitive can be used in the ditransitive construction? Clearly the problem is soluble, because we can, in fact, extend the ditransitive pattern to verbs which are not stored as such. But before an analysis such as Goldberg's can be evaluated it needs to be paired with a general theory of processing which explains how generalisations can go beyond the inheritance hierarchy.

We might look to psycholinguistics for a suitable theory. There is no shortage of theories which assign a central place to spreading activation (for a survey, see Hudson 2007b:41-2). However, spreading activation only explains part of the process: selection of a target among the millions of stored nodes; for example, it is activation spreading from each of the letter-characters that you read that guides your mind to the target word. What these theories do not explain is what we do with this target once we have found it, by applying it to the current task. This seems to involve two very different processes: binding the target node to the current token, and then enriching the token from the target's properties. For example, when you read a token of the word *cat*, you first identify the target as the lexeme CAT, but that is not the end of the matter; you then have to record mentally that your token is indeed an example of CAT, in order finally to work out what the token means, what word class it belongs to, and so on. The same is true when working in the opposite direction, from the meaning 'cat' to an utterance that carries this meaning: in this case, the target is once again a token which inherits its unknown characteristics (especially its pronunciation) from the word selected by spreading activation. In other words, both hearers and speakers have to combine spreading activation with binding and default inheritance; but so far as I know, no existing psycholinguistic theory does combines these elements.

One theory of processing that does combine them is the WG one (Hudson 2007a:41-52). Suppose I am speaking, and my target is to utter some word W to describe a situation in which Mary made a cake, which we can call S. (We consider the more complicated meaning that requires a ditransitive below.) In short, all I know about this word token is that it is a word and that it can refer to S, so the immediate task is to use my stored knowledge to enrich W: what word (or words) are needed in order to refer to S? There is no single word which fits S sufficiently precisely, so we have to find a more general one whose meaning can be made sufficiently precise by adding dependents (bearing in mind that these too may be modified recursively by their own dependents). Eventually I shall work out that I need to say *Mary baked a cake*, but all I know at the start is that I need a word which means that Mary baked a cake.

The first step, therefore, is to add a new node W to my mental network, but I can go a little further without any further thought. I know that W is going to be modelled on some stored word (or words), so I can introduce a dummy node M for this model word along with W, and of course I already know that M isa Word. So right from the first step, W is attached to the main network via M and Word: W isa M isa Word. In order to enrich W, I need to identify M with one or more stored nodes, inherit whatever properties these nodes make available, and then recursively enrich any other nodes that may have been added by inheritance (including the various dependent words).

The mechanism for enriching W in this way has three parts:

- activation which spreads through the grammar network from two highly active nodes, S (the situation to be described) and W, and which converges on various nodes in between.
- a "binding" mechanism which binds W to the best available stored word-types.
- default inheritance which inherits properties for W from these word-types.

The activation spreads from S to the concept Baking, and from there to the word-type BAKE that expresses this concepts; and at the same time, it spreads from W to M. Highly active dummy nodes such as M trigger a process of binding which binds it (by an "identity" link) to any stored nodes which satisfy two criteria:

- they are compatible with existing isa links of M, i.e. in this case each one isa Word.
- they are the most active such nodes, which guarantees that they provide the best global fit with the target properties.

One attraction of this approach is that it allows multiple inheritance[2], so I can identify M not only with BAKE, but also with Past-tense (which will express the time relations of S); each of these models provides a different kind of information about the target word W. Finally, once the identity of M has been established as the past tense of BAKE (in WG notation: "BAKE:past"), default inheritance applies to enrich W's properties from BAKE and Past-tense.

Another attraction is that it accommodates some degree of deviation, because the search is for the best available global fit, not for a perfect fit. (The "best-fit" principle is due to Winograd 1976.) This flexibility is helpful in explaining "loose" use, as when we might use BAKE for a process which is not exactly baking but which (in the current context) is nearer to baking than to anything else. This possibility opens the door to the full range of meaning-extensions such as metaphor and metonymy discussed in cognitive linguistics. Similarly, it allows us to ignore misspellings, mispronunciations and slips of the tongue provided that the right node receives the most activation from all parts of our entire cognitive network.

Now let us change the scenario to one in which the situation S is one in which Mary makes the cake for John. In this case, the activation spreads from S not only to BAKE and Past-tense, but also to the words suitable for doing something for someone else: not only to FOR, but also to the ditransitive verb-type. Leaving aside the choice between these alternatives, let us assume that one of them has more activation than the other, and has as much activation as BAKE and Past-tense. In that case, the binding process can attach M to Ditransitive as well, thereby allowing W to inherit an indirect object even though BAKE does not in itself allow one.

The general point of this discussion is that the notion of "generation" is a process which has to be spelled out before we know what a given grammar does or does not generate; and in a cognitively-oriented theory such as CG or WG, the best way to define generation is in terms of human processing. Consequently, we cannot explore language structure without at the same time considering language use, so we need to develop both kinds of theory in parallel. My account of ditransitives stands or falls by the theory of processing that I have just sketched, in contrast with the much more comfortable theories in which competence can be studied without reference to performance. On the other hand, I believe that other theories are in fact vulnerable in precisely the same way, so it is better to recognise the problem than to pretend we can ignore it. The cognitive context has to be part of any theory of language structure.

7. Conclusion

The similarities between CG and WG make the differences all the more surprising and worth discussion. I suspect that many, and perhaps most, of the differences are not matters of disagreement so much as of different research interests. For example, Goldberg has developed the theory of learning more fully than I have, whereas I have spent more time thinking about processing. Similarly, I am primarily (or at least historically) a morpho-syntactician whereas Fillmore and Goldberg are more interested in lexical semantics and argument structure, so it is understandable that CG takes a very conservative, or even agnostic, position on matters of sentence structure. Given these differences of focus, we might expect the two theories to be complementary rather than to conflict. It is easy to imagine a new theory which "isa" both of them and which inherits all their properties with very little need for conflict resolution.

However, I should finish by highlighting an issue that has come up several times in this article and where the two theories may be harder to reconcile. Do constructions exist? If all they are is mapping relations between some kind of form and some kind of function, then they are the same as the realisation relations or correspondences that every theory accepts. But the main claim of CG is that they are more than that.

One more precise claim is that they are very much more specific and numerous than the very general interacting structures of the Minimalist Program, but this is actually a claim about formal structures being specific and numerous, tied with the uncontroversial claim that any form may be mapped to a function. One could easily accept (as I do) that we store a lot of detailed morphological or syntactic patterns without thereby being committed to any particular view of how these are mapped to meaning.

A much more controversial claim about constructions is that they are the basic units of language – for example, that only constructions are organised in an inheritance hierarchy. However, I have pointed out that the formal and semantic elements of different constructions also have relations to one another, so in what sense do only constructions exist? I wonder what difference it would make if all the boundaries around constructions were removed. I suspect that the result would be hard to distinguish from WG, but I must leave the answer to the CG experts.

[1] I should like to thank the following for helpful comments: Jasper Holmes, Haitao Liu and two anonymous reviewers.

[2] Aarts (this volume) argues against multiple inheritance when applied to two lexical classes (noun and verb), but if inheritance is used at all, it is hard to avoid multiple inheritance when dealing with complex word-class definitions which combine a lexeme with an inflection (e.g. BUY, past – i.e. the past tense of BUY). The only way to avoid multiple inheritance would be to make one of these categories into a sub-case of the other, but this is not possible.

References

Aitchison, Jean.
1994 *Words in the Mind. An Introduction to the Mental Lexicon. Second Edition.* Oxford: Blackwell.
Anon
2003 *Longman Dictionary of Contemporary English*. Harlow: Pearson Education.
Aronoff, Mark
1994 *Morphology by Itself. Stems and inflectional classes.* Cambridge, MA: MIT Press.
Beer, Anton and Virginia Diehl
2001 The role of short-term memory in semantic priming. *Journal of General Psychology* **128**. 329-350.
Blake, Barry
1990 *Relational Grammar*. London: Croom Helm.
Bloomfield, Leonard
1933 *Language*. New York: Holt, Rinehart and Winston.
Bresnan, Joan
2001 *Lexical-Functional Syntax*. Oxford: Blackwell.
Bröker, Norbert
1998 A projection architecture for dependency grammar and how it compares to LFG. In *Proceedings of the LFG 98 Conference*. Butt, M. & Holloway King, T. (eds.), Stanford: CSLI.
Brooks, Patricia and Brian MacWhinney
2000 Phonological priming in children's picture naming. *Journal of Child Language* **27**. 335-366.
Chang, Franklin, Gary Dell and Kathryn Bock
2006 Becoming Syntactic. *Psychological Review* **113**. 234-272.
Chomsky, Noam
1957 *Syntactic Structures*. The Hague: Mouton.
1965 *Aspects of the Theory of Syntax.* Cambridge, MA: MIT Press.
1995 *The Minimalist Program*. Cambridge, MA: MIT Press.
Chomsky, Noam and Howard Lasnik
1993 The Theory of Principles and Parameters. In *Syntax: An International Handbook of Contemporary Research*. Jacobs, J., Stechow, A. v., Sternefeld, W., & Venneman, T. (eds.), Berlin: Walter de Gruyter. 506-569.
Covington, Michael
1984 *Syntactic Theory in the high middle ages.* Cambridge: Cambridge University Press.
Croft, William

1998 Linguistic evidence and mental representations (introspective linguistic data). *Cognitive Linguistics* **9**. 151-173.
Croft, William and Alan Cruse
2004 *Cognitive Linguistics*. Cambridge University Press.
Dik, Simon
1989 *The Theory of Functional Grammar I: The structure of the clause.* Dordrecht: Foris.
Dowty, David
1991 Thematic proto-roles and argument selection. *Language* **67**. 547-619.
Evans, Vyvyan and Melanie Green
2006 *Cognitive Linguistics. An introduction.* Edinburgh: Edinburgh University Press.
Fillmore, Charles
1982 Frame semantics. In *Linguistics in the Morning Calm*. Anon (ed.), Seoul: Hanshin; Linguistic Society of Korea. 111-138.
1986 Varieties of conditional sentences. *Eastern States Conference on Linguistics* **3**. 163-182.
Fillmore, Charles and Sue Atkins
1992 Towards a frame-based lexicon: the semantics of RISK and its neighbours. In *Frames, Fields and Contrasts. New essays in semantic and lexical organisation,* Lehrer, A. & Kittay, E. (eds.)*,* 75-102. Hillsdale, NJ: Erlbaum.
Fillmore, Charles, Paul Kay, and Mary O'Connor
1988 Regularity and idiomaticity in grammatical constructions: the case of let alone. *Language* **64**. 501-538.
Frost, Ram, Avital Deutsch, Orna Gilboa, Michael Tannenbaum, and William Marslen-Wilson
2000 Morphological priming: Dissociation of phonological, semantic, and morphological factors. *Memory & Cognition* **28**. 1277-1288.
Gaifman, Haim
1965 Dependency systems and phrase-structure systems. *Information and Control* **8**. 304-337.
Gisborne, Nikolas
1996 *English Perception Verbs*. PhD dissertation, UCL, London.
Goldberg, Adele
1995 *Constructions. A Construction Grammar Approach to Argument Structure.* Chicago: University of Chicago Press.
1998 Patterns of experience in patterns of language. In *The new psychology*, Tomasello, M. (ed.), Tomasello, M. (ed.), 203-219. London: Lawrence Erlbaum.

2002 Words by Default: the Persian Complex Predicate Construction. In *Linguistic Mismatch: Scope and Theory*, Francis, E. & Michaelis, L. (eds.). Berlin: Mouton.

2006 *Constructions at work. The nature of generalization in language.* Oxford: Oxford University Press.

Goldberg, Adele and Giulia Bencini

2005 Support from Language Processing for a Constructional Approach to Grammar. In *Language In Use: Cognitive And Discourse Perspectives On Language And Language Learning. Georgetown University Round Table on Languages and Linguistics,* Tyler, Andrea (ed.). Georgetown University Press.

Halliday, Michael

1961 Categories of the theory of grammar. *Word* **17**. 241-292.

1985 *An Introduction to Functional Grammar*. London: Arnold.

Harley, Trevor

1995 *The Psychology of Language*. Hove: Psychology Press.

Heringer, Hans J.

1993 Dependency syntax - basic ideas and the classical model. In *Syntax - An International Handbook of Contemporary Research, volume 1*. Jacobs, Joachim, Arnim von Stechow, Wolfgang Sternefeld, & Theo Venneman (eds.), 298-316. Berlin: Walter de Gruyter.

Holmes, Jasper

2005 *Lexical Properties of English Verbs*. PhD dissertation, UCL, London.

Holmes, Jasper and Richard Hudson

2005 Constructions in Word Grammar. In *Construction Grammars. Cognitive grounding and theoretical extensions*. Östman, Jan-Ola & Miriam Fried (eds.), 243-272. Amsterdam: Benjamins. Amsterdam: Benjamins.

Hudson, Richard

1971 *English Complex Sentences. An introduction to systemic grammar.* Amsterdam: North Holland.

1984 *Word Grammar.* Oxford: Blackwell.

1990 *English Word Grammar.* Oxford: Blackwell.

1992 So-called 'double-objects' and grammatical relations. *Language* **68**. 251-276.

2007 *Language Networks: the new Word Grammar*. Oxford: Oxford University Press.

2008 Buying and selling in Word Grammar. In *Critical Concepts in Lexicology*. Hanks, Patrick (ed.). London: Routledge.

Hudson, Richard and Jasper Holmes

2000 Re-cycling in the Encyclopedia. . In *The Lexicon/Encyclopedia Interface,* Peeters, Bert (ed.), 259-290. Amsterdam: Elsevier.

Jackendoff, Ray
1997 *The Architecture of the Language Faculty*. Cambridge, MA: MIT Press.

James, Lori and Deborah Burke
2000 Phonological Priming Effects on Word Retrieval and Tip-of-the-Tongue Experiences in Young and Older Adults. *Jornal of Experimental Psychology -Learning, Memory and Cognition* **26**. 1378-1391.

Kay, Paul and Charles Fillmore
1999 Grammatical constructions and linguistic generalizations: The what's X doing Y? Construction. *Language* **75**. 1-33.

Keenan, Edward
1976 Towards a universal definition of 'subject'. In *Subject and Topic*, Li, C. (ed.), 303-333. New York: Academic Press.

Kunze, Jürgen
1975 *Abhängigkeitsgrammatik.* Berlin: Akademie-Verlag.

Lakoff, George
1977 Linguistic gestalts. *Papers From the Regional Meeting of the Chicago Linguistics Society* **13**. 236-287.
1987 *Women, Fire and Dangerous Things*. Chicago: University of Chicago Press.

Lamb, Sydney
1966 *Outline of Stratificational Grammar*. Washington, DC: Georgetown University Press.
1998 *Pathways of the Brain. The neurocognitive basis of language.* Amsterdam: Benjamins.

Langacker, Ronald
1987 *Foundations of Cognitive Grammar I: Theoretical prerequisites*. Stanford: Stanford University Press.
1990 *Concept, Image and Symbol. The Cognitive Basis of Grammar.* Berlin: Mouton de Gruyter.
2000 A dynamic usage-based model. In *Usage-based Models of Language*. Barlow, M. & Kemmer, S. (eds.), 1-63. Stanford: CSLI.

Levelt, Willem, Ardi Roelofs, and Antje Meyer
1999 A theory of lexical access in speech production. *Behavioral and Brain Sciences* **22**. 1-45.

Levin, Beth
1993 *English Verb Classes and Alternations. A preliminary investigation.* Chicago: University of Chicago Press.

Luger, George and William Stubblefield

1993 *Artificial Intelligence. Structures and strategies for complex problem solving*. New York: Benjamin Cummings.

Matthews, Peter
1993 *Grammatical theory in the United States from Bloomfield to Chomsky*. Cambridge: Cambridge University Press.

Mel'cuk, Igor
1988 *Dependency syntax : theory and practice.* Albany: State University Press of New York.

Moss, Helen, Ruth Ostrin, Lorraine Tyler, and William Marslen-Wilson
1995 Accessing different types of lexical semantic information - evidence from priming. *Journal of Experimental Psychology-Learning Memory and Cognition* **21**. 863-883.

Ninio, Anat
2006 *Language and the learning curve: A new theory of syntactic development.* Oxford: Oxford University Press.

Norris, Dennis, James McQueen, and Ann Cutler
2002 Bias effects in facilitatory phonological priming. *Memory & Cognition* **30**. 399-411.

Owens, Jonathan
1988 *The Foundations of Grammar: an Introduction to Mediaeval Arabic Grammatical Theory.* Amsterdam: Benjamins.

Percival, Keith
1976 On the historical source of immediate constituent analysis. In *Notes from the Linguistic Underground,* McCawley, James (ed.), 229-42. London: Academic Press.
1990 Reflections on the History of Dependency Notions in Linguistics. *Historiographia Linguistica.* **17**. 29-47.

Perea, Manuel and Eva Rosa
2002 The effects of associative and semantic priming in the lexical decision task. *Psychological Research-Psychologische Forschung* **66**. 180-194.

Pollard, Carl and Ivan Sag
1994 *Head-Driven Phrase Structure Grammar*. Chicago: Chicago University Press.

Reisberg, Daniel
1997 *Cognition. Exploring the Science of the Mind.* New York: Norton.

Robins, Robert
2001 In Defence of WP. (Reprinted from TPHS, 1959). *Transactions of the Philological Society* **99**. 114-144.

Robinson, Jane
1970 Dependency structure and transformational rules. *Language* **46**. 259-285.

Sag, Ivan
1997 English relative clause constructions. *Journal of Linguistics* **33**. 431-483.
Sgall, Petr, Eva Hajicova, and Jarmila Panevova
1986 *The Meaning of the Sentence in its Semantic and Pragmatic Aspects*. Prague: Academia.
Sinclair, John
1987 *Collins Cobuild English Language Dictionary*. London: Collins.
Smith, Marilyn, Shlomo Bentin, and Thomas Spalek
2001 Attention constraints of semantic activation during visual word recognition. *Journal of Experimental Psychology-Learning Memory and Cognition* **27**. 1289-1298.
Spinelli, Elsa, Juan Segui, and Monique Radeau
2001 Phonological priming in spoken word recognition with bisyllabic targets. *Language and Cognitive Processes* **16**. 367-392.
Stump, Gregory
2001 *Inflectional Morphology: A Theory of Paradigm Structure*. Cambridge: Cambridge University Press.
Sugayama, Kensei and Richard Hudson
2006 *Word Grammar. New perspectives on a theory of language structure.* London: Continuum.
Tesnière, Lucien
1959 *Éléments de syntaxe structurale*. Paris: Klincksieck.
Tomasello, Michael and Elizabeth Bates
2001 *Language Development: The Essential Readings*. Oxford: Blackwell.
Venneman, Theo
1977 Konstituenz und Dependenz in einigen neueren Grammatiktheorien. *Sprachwissenshaften* **2**. 259-301.
Winograd, Terry
1976 Towards a procedural understanding of semantics. *Revue Internationale De Philosophie* **30**. 260-303.