Constituency and dependency

RICHARD A. HUDSON

Abstract

I argue that dependency is necessary in syntax, but constituency is not. That is, there is no need to recognise a hierarchical structure with clauses and phrases grouping words together, in addition to the dependency relations which need to be shown between words. Consequently there is no need for 'daughter-dependency' rules in a so-called 'Daughter-dependency grammar', but the position of sister-dependency rules is assured, and in fact identified with the subcategorisation role of pan-lexical entries for words.

1. Background

The history of syntactic analysis and theory in this century is dominated by the notions 'constituency' and 'dependency'. Constituency is the part–whole relation shown in the American structuralist and transformational tradition by means of trees or boxes, while dependency is the part–part relation of a 'modifier' to its 'head', which is more typically shown by the European tradition of 'stemmas'. The two diagrams in (1) show, respectively, constituency without dependency and dependency without constituency:

(1) a.

This sentence raises some surprisingly difficult problems

b.

This sentence raises some surprisingly difficult problems

Already in Jespersen’s *Philosophy of Grammar* (1924) both dependency and constituency were described (though not by these names, of course) as characteristics of syntactic sentence-structure. On the one hand, he distinguished between words being used as ‘primaries’, ‘secondaries’ and ‘tertiaries’ according to the dependency relations between them (e.g. in *extremely hot weather, extremely* is a tertiary, *hot* a secondary and *weather* a primary; *op cit*: 96). But on the other hand he recognised that one term of such a dependency relation might be a group of words acting as a single unit. For example, he analyses (1924:102) the sentence *He met the kind old Archbishop of Canterbury* as containing a word-group (what we would now call a noun-phrase, though he does not use this term), *the kind old Archbishop of Canterbury*, as a primary in relation to *met*. (In contrast with virtually the whole of the dependency tradition that followed him, Jespersen treats the verb as a modifier of the subject and object, rather than vice versa.) However, it is not at all clear why Jespersen makes this claim, as he then goes on to analyse *the kind old Archbishop of Canterbury* in terms of its own internal structure, and treats *Archbishop* as a primary with the other words depending on it as secondaries and tertiaries. Why then do we need to recognise the whole group as such, rather than treating the word *Archbishop* as the primary which is object of *met*, with the other words depending on it? Surprising as this may seem, the question has never been given a satisfactory answer, except by those who believe that Jespersen was wrong to add the notion of constituency to that of dependency. This is the view I shall advocate in the present paper.

The question is not whether in some sense a group of words may behave syntactically like a single word. Ever since the early days of Immediate Constituent analysis in America (first formalised in Wells, 1947) one of the standard tests for the unity of a group of words has been to try replacing it by a single word (e.g. *the kind old Archbishop of Canterbury* may be replaced by *him* without changing any other aspect of the sentence’s structure). Even the father of dependency theory, Lucien Tesnière, recognised that dependency relations may bind words together into groups, and gave a name (*noeud*, ‘knot’) to such groups (1959:14). What has to be decided, rather, is whether there is any need to refer in a grammar to constituency relations *in addition to* dependency relations. After all, it can easily be seen that the dependency diagram in (1b) above shows how the words fall into groups just as the constituency diagram does, namely by bringing together words under a single node of the diagram (e.g. both *surprisingly* and *difficult* are attached to the same node, namely that dominating *difficult*). This fact has been recognised for some time (see in particular Robinson, 1970) as a general property of dependency diagrams: any dependency diagram may be converted into a
constituency diagram by a mechanical procedure which involves trivial reorganisation of the nodes and branches (and some adjustment of labelling, which is not in question at the moment). Notwithstanding this equivalence of dependency diagrams to constituency diagrams, it might be that there are advantages in recognising constituent-structure as well as dependency, so that the grammar could refer directly to groups of words as wholes, rather than just to the head word in such a group. For instance, there might be rules of grammar which must refer to the notions ‘clause’ or ‘noun-phrase’, and which cannot be replaced by rules referring instead to the head-words of such groups (verbs and nouns, respectively), leaving the other dependency rules to keep the modifiers in step with the head. I shall consider a number of rules in section 4 which look like strong candidates for such rules, but I shall show that they can in fact all be formulated satisfactorily without reference to constituency. This conclusion raises serious problems for one aspect of Daughter-dependency grammar, of course, since the theory as described in Hudson (1976a) recognises constituency as well as dependency.

Before attacking the question whether constituency is necessary, however, we have to face the other extreme possibility: that dependency is not necessary. Clearly, if we could show that there was no need to show dependency relations directly in a grammar, then the arguments against constituency would lose their force. Dependency representations give information over and above that given in a constituency representation (namely, the direction of dependencies: surprisingly depends on difficult, rather than vice versa), but even if this extra information turns out to be unnecessary, the basic information about groupings of words clearly has to be given in order for the grammar to work. (For example, the rule which inverts subjects and verbs in English has to make sure that the verb is followed not only by the subject noun but also by all the latter’s modifiers.) In this case, a direct representation of constituency will turn out to be necessary, since there will be no need for a dependency representation from which constituency may be derived. Consequently, the first job is to show that the extra information given in a dependency diagram is indeed necessary, and this is what I shall try to do in the next section, after which I shall discuss the notion ‘dependency’ itself. (To anticipate the discussion in section 3 I shall argue that A depends on B if A contributes to the semantic structure of B; in some cases this will mean that A and B are interdependent, since each contributes to the other’s semantic structure.)
2. Is dependency necessary?

In answering this question, I shall assume that it refers to what is necessary for a generative grammar to be satisfactory. Here I take 'generative grammar' in the sense of a grammar which allows us to decide in a mechanical way whether or not some string of words constitutes a well-formed sentence of the language concerned, but I shall not of course assume that the grammar contains transformations. I shall assume that the input to the grammar is an arbitrary structural representation of the string of words, and the output is a verdict on whether or not such a representation is permitted by the grammar. A generative grammar will be 'satisfactory' to the extent that it says what we want it to say about the language concerned. Thus what I have to show is that there are some facts about languages which most linguists would presumably like to be able to reflect in their grammars, but which can only be reflected if the grammar refers explicitly to dependency relations.

To simplify terminology, I shall call the two terms of any dependency relation the 'head' and 'modifier' respectively, corresponding to Tesnière's 'régissant' and 'subordonné' and the terms 'Regens' and 'Dependens' used in much German literature on dependency theory (e.g. Korhonen, 1977). There does not appear to be a better pair of terms in English, but these terms have the disadvantage of being traditionally used to refer to one particular type of construction, namely phrases like noun-phrases and adjective-phrases, in which the head is a noun or adjective, as the case may be, and the modifiers are adjectives, adverbs and so on. I shall generalise these dependency relations by using 'head ' to refer also to prepositions and verbs, and by making the terms strictly relative. Thus, in This sentence raises surprisingly difficult problems, difficult is head in relation to surprisingly, but modifier in relation to problems, and the latter is head in relation to difficult, but modifier relative to raises. The following arguments all require the user of a grammar (assumed to be a mindless automaton) to be able to recognise dependency relations in an abstract structural representation, and to be able to decide which of the elements concerned is head and which is modifier. I assume that a representation which shows pure constituency (such as (1a)) would not give the user enough information to do so, and would therefore not provide the basis for a satisfactory grammar.

A. Cross-category sequence generalisations

This argument was first offered by Tesnière himself (1959:22), but has since been developed by others, including Venneman (1975) and Heine (1975). It is obviously closely related to the discussion of word-order
universals in Greenberg (1963). Very simply, there seems to be a strong tendency for languages to be consistent in their positioning of modifiers relative to their heads. In some languages (which Tesnière calls ‘centrifugal’), modifiers tend to follow their heads: objects follow verbs, prepositional objects follow the preposition, adjectives follow nouns, genitives follow nouns, and so on. In other languages, modifiers tend to precede their heads, and Tesnière calls such languages ‘centripetal’. According to Tesnière, there are very few languages, if any, which do not fit clearly into one or the other of these two categories, though there are many languages which are not pure examples — in which most, but not all, modifiers precede their heads, for example. Tesnière’s evidence for this claim may be suspect — for instance, in his table (1959:32) classifying various language families according to the centrifugal/centripetal contrast he refers to such gross categories as ‘languages of America’! However, Greenberg’s data show a clear tendency among the 30 languages that he studied for languages with prepositions to put objects after verbs, and for those with postpositions to put objects before verbs. In any case, we need not go into the question whether there really is a world-wide tendency for all modifiers to appear on the same side of their heads within a given language. Even if there are an appreciable number of languages in which pre- and post-head modifiers are about equally balanced, there still remain the established cases of languages in which one order is dominant, and any linguist working on such a language would surely want to be able to express this generalisation in his grammar, while allowing for the exceptions in special statements. If dependency relations are shown explicitly in his sentence-structures, there is no problem; for instance, a grammar of Welsh may contain the rule ‘modifiers follow heads’, to apply in all cases where no other rule requires an alternative order; and vice versa for Japanese. If dependency relations are not shown explicitly, the rule cannot be stated, and each particular construction must be covered by a separate sequence statement: ‘objects follow the verb’, ‘prepositional objects follow the preposition’, ‘possessives follow their sister noun’, and so on. If a language does not fit clearly into either of Tesnière’s types, this simply means that the general notion of dependency is not relevant to linearisation in that language, and it might well turn out that some other principle — such as ‘communicative dynamism’ or syntactic class — plays a dominant role instead.

B. Asymmetry of sequence rules

In general, the position of a modifier is determined relative to that of its head, rather than vice versa. For example, linguists commonly treat the
verb as the 'pivot' of a sentence and state the positions of other elements relative to it, which fits our generalisation if we take the verb as the head of the whole sentence, as in standard dependency theory. Similarly, there is a very simple rule in English that any item (except the finite main verb) may be put at the start of the sentence, provided that all its dependents are moved with it. Thus, we find sentences like (2a) but not (2b), and like (3a) but not (3b).

(2) a. Forget her birthday he may, but ...
   b. *Forget he may her birthday, but ...

(3) a. On that bike you’d go faster.
   b. *On you’d go faster that bike.

(It appears that this restriction does not apply to German in the same way, in view of sentences like (4) in which the verb is topicalised without its dependent objects:

(4) Vergessen werde ich Dich nicht. ‘I'll never forget you’.)

Similarly, if a noun-phrase is object of a preposition, it always occupies the position appropriate for the preposition (its head), rather than the one that it would have had without the preposition. For example, in English indirect objects without a preposition must come immediately next to the verb, but if they occur with to or for then they come in the normal position for prepositional phrases (after the direct object, unless there is a disparity of 'weight'). Similarly, a noun-phrase which would be subject of an active verb, and occur in the normal subject position, is relegated to the end of the clause if by is added to it as a result of passivisation. A more complex case involves gapping (discussed more fully in Hudson 1976b), in sentences like (5):

(5) John invited Mary, and Bill, Sue.

However sentences like this are to be treated (and I prefer to treat them as involving displacement of Bill and Sue, rather than deletion of invited), the general constraint is that the 'remnant' items (viz. Bill and Sue) must be accompanied by all their modifiers. Putting it another way, the remnant may be turned into a complete clause by borrowing extra heads or extra co-modifiers from the first clause, but not by borrowing extra modifiers. Thus, (6a) is good because the borrowed items (will prepare) are items on which the remnant depend, whereas in (6b) the borrowed items are will, on which present depends, but also the paper, which is a modifier of present and therefore is not borrowable, but ought to be located next to present as per usual.
Constituency and dependency

(6) a. John will prepare the paper, and Bill, the handout.
b. *John will prepare the paper, and Bill, present.

In a pure constituency approach it is hard to generalise across such a wide variety of constructions in order to explain why the constraints are as they are, though of course it is possible to formulate ad hoc individual constraints for each construction. However, given explicit representations of dependency it is both easy and natural to bring all such cases together under a single rule or convention: always position a modifier relative to its head (using whatever rules are available for doing so, and at least keeping it as close as possible to its head), unless other factors intervene.

C. Cross-category generalisations of morphological determination

Wherever the inflectional form of a word is determined by the properties of another word, the two words concerned are always in a modifier–head relation (that is, one depends directly on the other), and the determination goes from head to modifier, which means that the form of the modifier is determined by the properties of the head, rather than vice-versa. For example, adjectives show concord with their head-nouns, and in languages with noun-cases the object of a preposition has its case governed by the preposition. In either of these examples it may seem that what is excluded is in any case inconceivable, but this is not so: it is easy to imagine a language in which nouns take different forms according to which class of adjectives modifies them (including an ‘unmarked’ form for use when there is no modifying adjective), and it is also possible to imagine a language in which prepositions had different forms for use, say, with animate and inanimate nouns. So far as I know, neither of these possibilities is realised in any natural language, though the descriptive problems are considerable; for instance, it is well known that prepositions are to some extent selected by the semantic properties of their object noun (one travels on a train but in a car, for example). A moment’s thought is enough to show that the rule as formulated above raises even more serious descriptive problems in connection with other constructions. For example, in some languages a predicative adjective shows concord with the subject (cf. Il est grand ‘He is tall’ versus Elle est grande ‘She is tall’); are we to say that in such languages, and perhaps in all other languages as well, a predicative adjective is directly dependent on the subject (rather than on the verb, as it would be if it were a direct object)? Or perhaps we can say that it depends directly on the subject as well as the verb? This is in fact an analysis which offers many attractions, for other reasons, and I shall
return to it briefly below. Similar problems seem to be raised by another obvious apparent counter-example to the claim that modifiers never determine the form of a head, namely the subject-finite verb relation in many languages, including English. In the dependency-theory tradition (though not in Jespersen, it will be remembered), the verb is treated as head and the subject as one of its modifiers, but it has been suggested (e.g. by Werner, 1975; see the survey in Bauer, 1976) that the subject might be treated as head and the verb as its modifier, and I believe this is an analysis which deserves to be considered very seriously, as it appears to have some important attractions. If then these apparent counter-examples can be disposed of, we can formulate a very useful general strategy for concord and government rules: when the form of a word is determined by the properties of another word, it is never necessary to specify what the latter is, since it will always be the first word’s head. Needless to say, without the notion of dependency, and direct representations of dependency in structural diagrams, there would be no hope of formulating such an attractively general rule.

D. Cross-category generalisations of semantic relations

Modifiers supply fillers for the ‘slots’ in the semantic representations of heads, rather than vice versa. For instance, in very hot, we may assume that the semantic representation for hot includes a slot, or variable, labelled ‘degree’, and the meaning of very (something like ‘well above the average’) is put in this slot as part of the semantic representation of a sentence containing very hot. Similarly, the semantic representation of a verb is commonly seen as a frame specifying a variety of slots, some of which refer to factors such as manner and degree, with others referring to participants involved in the event or state, such as actors, goals and so on. In this view, the modifiers of a verb each contribute a filler for one of the verb’s semantic slots. Thus syntactic dependency relations may be seen as at least very neatly in step with semantic structure, and possibly as completely isomorphic with the latter. There has been considerable debate in the dependency literature on whether dependency is a syntactic or a semantic relation (see for instance the references in Korhonen, 1977), and it may be that the answer is ‘Both!’. However, before we can be sure that dependency is completely isomorphic with semantic structure, there are a number of descriptive questions to be answered. For example, in the previous paragraph I suggested that the facts of morphology might push us to treat the subject of a verb as its head, rather than vice versa, but vice versa is presumably what is required by the semantic facts, on the
Constituency and dependency

187

assumption that the subject supplies one of the fillers for the verb's semantic slots. The solution to this problem may turn out to be that the dependency needs to go in both directions — in other words, the relation between subject and verb is one of interdependence (recognised notably by Heringer, 1970:78) rather than of one-way dependency. Again there are other reasons for favouring this analysis, which must be discussed elsewhere. If this solution is adopted, and other apparent discrepancies between the requirements of semantics and of other factors can be similarly removed by postulating interdependence, then it will be possible to formulate a general convention for relating syntactic and semantic structures to one another: to find fillers for slots in a semantic frame, consider only modifiers of the word concerned as possible candidates. (It may be that this convention controls only the normal cases, and that it may be overridden by specific rules; for example, it is probable that the filler for the 'subject' slot of a non-finite verb is not always related to the verb by a direct dependency link, whether as head or modifier, and that special rules may need to be formulated for filling the subject slot of non-finite verbs; an example would be the subject of verbs like work in 'Super-equi' constructions like John finds it hard to work on linguistic theory.) A pure constituency approach would allow us to restrict the fillers of a word's semantic slots to the word's sister constituents, but it would not allow us to predict that sisters will be relevant only if they are modifiers, and never if they are in a one-way 'head' relation.

E. Cross-category generalisations about subcategorisation

This argument is similar to the preceding one, and may be outlined quite briefly. If a word is given a syntactic frame showing the kind of syntactic environment with which it is compatible, the items referred to in that environment will all be modifiers, rather than heads. To take an obvious example, verbs are subcategorised (by means of a frame, I assume) according to whether they take objects and what kinds of object they take, from a syntactic point of view (noun-phrases, that-clauses, and so on; or in case-languages like German, according to the case in which an object noun-phrase is required to appear); but one does not try to subcategorise nouns according to whether or not they 'take' a verb in the 'object' relation. In other words, only a word's modifiers may provide fillers for its syntactic frames, just as only they may fill its semantic frames. (Of course, this is not the same as saying that modifiers always fill slots specified in the individual word's syntactic frame. For example, there may not be any reason for giving individual adjectives a syntactic frame containing a slot
for 'degree adverb', on the assumption that this slot is available for any adjective, but the theory of 'Pan-lexicalism' (Hudson, 1979) allows us to specify frames for adjectives in general, and these frames can contain syntactic slot for degree adverbs, just as the frame for finite verbs in general may contain a slot for subject, making it unnecessary to refer to the latter in the entries for individual verbs.) Again the dependency framework allows us to formulate this general constraint on grammars, but a pure constituency approach does not.

This completes the list of arguments in favour of including explicit representations of dependency relations in the structures that a grammar generates. It will be seen that they all have a similar form: phenomenon X applies to a wide range of constructions covering a number of different categories (such as clauses, prepositional phrases and noun-phrases), and is easy to formulate in terms of dependency relations, but if we cannot refer to the latter, there is no way of using the constituency framework to capture X as a unitary phenomenon.

3. What is dependency?

Having seen what dependency is useful for, we may now attempt a definition of the phenomenon which reconciles the needs of the various uses to which we want to put it. Once again we are entering very well-trodden terrain, as far as the literature on dependency theory is concerned, and the interested reader is referred to Korhonen (1977) for a good review of the main views that have been expressed. Suffice it to say that different scholars have focussed on different properties in defining dependency, some concentrating on syntax, others on semantics, others on morphology, and so on. As should be clear from the previous section, dependency may be seen as relevant to all these levels of language, and to different aspects (such as linear sequence and cooccurrence relations) on each level. (For the last point, see Haas, 1973.) The question is whether there is any way in which all these different kinds of relation can be brought together under a single definition, or whether we have to resign ourselves to concentrating on one level or aspect at the expense of the remaining ones. In other words, when we were talking about the role of dependency in stating word-order and the role of dependency in describing, say, semantic relations, were we in fact using 'dependency' in the same way, to refer to the same phenomenon in each case?

I believe it is possible to give an all-embracing definition of dependency which brings together all the relations described above. This definition must refer, once more, to the notions 'slot' and 'frame' alluded to in the
discussions of semantic relations and syntactic subcategorisation above (points D and E in the previous section). In the panlexicalist approach, the grammar is made up of a set of frames, plus rules for using these. Most frames correspond to lexical entries in an Aspects-type grammar, with one or more frames for each verb, for example, showing what kinds of complements it takes; but some frames give information about whole classes of words, such as finite verbs, nouns, or comparative adjectives. A frame contains a number of 'slots', specifying variables to be filled from the linguistic context (e.g. an 'object' slot), as well as a number of already filled variables such as pronunciation, syntactic class and certain aspects of meaning. The relevance of frames and slots to dependency is that we may say that, by definition, any filler of a slot is a modifier relative to the item to which the frame belongs. (Specifying to which item a frame belongs is easy when it is an ordinary lexical entry, since the head is the item whose pronunciation variable is filled; but more general frames require a somewhat more indirect definition of the head, which need not concern us here.)

This definition is offered here only as a tentative solution to the problem of defining dependency, but it does seem to give the right answer for all the constructions to which we have referred, and a number of others. The following is a list of constructions, with examples, showing which item qualifies as modifier (M) by virtue of filling a slot in the other's frame (H), assuming a reasonable application of the frame-and-slot approach.

<table>
<thead>
<tr>
<th>modifier</th>
<th>head</th>
<th>example</th>
</tr>
</thead>
<tbody>
<tr>
<td>object</td>
<td>verb</td>
<td>likes (H) biscuits (M)</td>
</tr>
<tr>
<td>subject</td>
<td>verb</td>
<td>John (M) drowned (H)</td>
</tr>
<tr>
<td>prep. obj.</td>
<td>preposition</td>
<td>in (H) London (M)</td>
</tr>
<tr>
<td>adjective</td>
<td>noun</td>
<td>big (M) hands (H)</td>
</tr>
<tr>
<td>genitive</td>
<td>noun</td>
<td>my (M) hands (H)</td>
</tr>
<tr>
<td>determiner</td>
<td>noun</td>
<td>the (M) end (H)</td>
</tr>
<tr>
<td>adverb</td>
<td>adjective</td>
<td>very (M) big (H)</td>
</tr>
<tr>
<td>rel. clause</td>
<td>noun</td>
<td>people (H) who think that (M)</td>
</tr>
<tr>
<td>comp. clause</td>
<td>comp. adj.</td>
<td>bigger (H) than Mary was (M)</td>
</tr>
</tbody>
</table>

The analysis is terms of modifiers and heads shown in (7) is 'right' in the sense that it conforms to the claims made in the previous section. In each case the position of the head is fixed without reference to that of the modifier (see point B in section 2); the semantic frame into which the meanings need to be integrated is provided by the head (point D), and similarly for the syntactic frame (point E). There are two constructions in which morphological concord applies, and in one of these
(determiner + noun, when realised by *this*/*these* or *that*/*those*) the direction is as predicted by point B; we shall have to return below to the other, the subject + verb construction. This leaves just point A, concerning cross-category generalisations about sequence, where we predicted that in some languages there would be a strong preference, across constructions, for one or the other of the orders of modifier and head. As far as the constructions listed in (7) are concerned, English does not seem to be such a language, as the two orders are about equal in representation, though it might turn out to be possible to generalise that if the head is a verb or preposition, its modifiers tend to follow it, and in other cases the modifier normally precedes (unless it is long, in which case it follows). However, point A did not claim that all languages would have a dominant order, but only that some languages would, so the data in (7) are irrelevant.

The main outstanding problem for the definition of dependency in terms of slots and frames is that certain facts require the verb and the predicative adjective to be treated as modifiers of the subject, even if other facts require the reverse relation, as we saw in the previous section. The former facts to which we referred were matters of morphology: in some languages the verb shows concord with the subject, as does the predicative adjective, and there are (presumably) no languages in which the converse is true. We might add, however, that if the adjective is considered to fill a semantic slot in the frame of the noun in, say, *big boy*, there should be some sense in which the same is true when the adjective is predicative, as in *The boy is big*, though I admit that this is arguable. Similarly, if *singing* fills a slot in the frame for *boy* in *the singing boy*, we might argue for the same analysis, mutatis mutandis, for *The boy is singing*. The differences between the constructions concerned involve the difference between Jespersen’s ‘junction’ and ‘nexus’, which he found so hard to put into words (see 1924:97). I shall not try at this stage to offer an account of the difference, but shall simply point out that it is not restricted to contrasts like those illustrated, where the ‘nexus’ member of the pair involves the subject-predicate relation. We can also bring in examples like (8), in which the item which, according to the proposed analysis, modifies the noun is not the first verb or adjective:

(8) a. *He* left the room *happy*.
   b. *He* left the *room empty*.
   c. *He* sat *watching* TV.
   d. *He* talked to her *sitting* at her feet.
   e. *She* had her *hair* loose.

In all these examples, the italicized noun is in some sense modified semantically by the italicized adjective or participle, and if it can be
accepted that this relationship is indeed one of modification, with the noun as head and the adjective or participle as modifier, then it should not be too hard to accept the same analysis for ordinary subject-predicate combinations. This analysis then allows us to predict that if any of the modifier items (adjectives, participles, finite verbs) are morphologically modifiable, then they will show concord with their respective head nouns, even if there is also a dependency relation going in the opposite direction between them. Thus the outstanding problems for our definition of dependency can be eliminated. However, I should emphasise that the proposed definition of dependency, and the solution to the problem of morphological dependencies that go 'the wrong way', are both very tentative and deserve much more careful consideration than I have been able to give them here.

4. Is constituency necessary?

If the arguments of section 2 were right, we need to make our grammars generate structures which represent dependency relations between words explicitly, and if the definition of dependency in section 3 is valid we shall presumably be able to identify dependency relations involving every word in a sentence. In other words, dependency relations do not involve just certain elements in the sentence, such as the complements of the verb, but every word that has a grammatical relation of any kind to another word in the sentence. The question we must now address is whether we need to show constituency relations in addition to these dependency relations, so one argument for constituency that we shall not be able to use is the argument that there are some words left over, outside the dependency structure of the sentence, which can only be integrated in terms of a constituency analysis. My view is that constituency is not needed as well as dependency (a view which I find I share with Korhonen (1977:83), but not with most other writers in the dependency tradition), and to support this view I shall now consider a number of arguments which have been used, or could be used, in favour of constituency. However, all I can do is to show that these particular arguments are unconvincing, and this is obviously very different from showing that no convincing arguments for constituency are possible. I hope that by presenting these arguments I shall stimulate other linguists to submit their favourite arguments to scrutiny.

A. Linear sequence

One of the main reasons given by dependency grammarians for having constituency as well as dependency is that the former, but not the latter,
allows one to state rules for linear sequence of elements. (For this view, see especially Baumgärtner, 1970.) This claim presumably arises from the historical development of the European dependency school and the American constituency school: for a variety of reasons, Tesnière had little to say about linear sequence (apart from the generalisations about typology referred to in section 2), whereas the constituency approach was developed by linguists such as Harris and Chomsky who wished to give rules covering the sequence of elements, and arrived at the phrase-structure rule for doing so. (It should be noted, incidentally, that more recent work by Harris is much nearer in spirit to the proposals in the present paper, see Harris, 1978.) However it should be clear that there is no necessary connection between constituency and the specification of linear sequence. For one thing, there is now a healthy movement within the transformational tradition which advocates unordered constituent structures at the deepest level of analysis (see for example Sanders, 1970; G. Hudson, 1972; Johnson and Postal, forthcoming), and for another there is no problem in writing linearisation rules for dependency structures, as I did in Hudson 1976a:108. An example of such a rule formulated in terms of dependency relations is (9).

(9) The modifiers of a head should not be separated from it by any other items except other modifiers of the same head.

By this very general convention we immediately guarantee that sisters will be next to each other unless some rule ordered before (9) requires or allows some modifier to be moved away from its head (as, for example, in the case of extraposition of a relative clause). Other more specific linearisation rules may easily be formulated in terms of dependency relations between specified categories of item, such as determiners and nouns.

B. Boundary positions

One particular type of sequence rule raises special problems for a grammar without constituency. This is the type of rule which, in a constituency-based grammar, would refer to the position at the start or (less usually?) at the end of a constituent. Such a rule would seem to be needed, for example, for positioning subordinating conjunctions, relative pronouns and wh-interrogative words in English, which all occur at the start of the clause. If there is no constituency, then there are no clauses, and consequently no starts of clauses, so the position of these elements needs to be formulated in some other way. One way of doing this is simply
to use a rule which requires the items concerned to come before the verb and any other of the latter’s modifiers. There are other possibilities that suggest themselves and which may turn out to be better, but it should be clear that constituent structure is not essential for defining boundary positions.

C. Features on higher nodes

In Hudson, 1976a (Appendix 2), I argued that constituent structure was necessary in addition to dependencies between sisters because otherwise there would be no ‘higher nodes’ to which features could be attached, as I had earlier argued was necessary (op cit.: 36). I now believe I was wrong in arguing that features are needed on higher nodes, since my principal piece of evidence turns out to have been wrong. I based my argument largely on the analysis of constructions containing gerunds in English, which I argued needed to be treated both as noun-phrases and as clauses. In this I was following Chomsky’s position in his 1970 paper, but unlike him I drew the conclusion that it was unsatisfactory to analyse such constructions as noun-phrases containing clauses as their only constituents. Rather, I argued, they should be treated as noun-phrases which were at the same time clauses, and consequently it was necessary to cross-classify such constructions by means of features, and to have a higher-level node representing the construction as a whole to which such features could be added. However, there is an easier solution to the problem of showing the similarity between a construction like *his solving the problem so quickly* and both clauses and nounphrases: we simply cross-classify the word *solving* as both a verb and a noun, then it will automatically share the properties of both. If this solution is correct (and clearly it needs more support and explanation than I have given it here), then there is no need to postulate an item in the sentence structure containing both the gerund and all its complements, to which the cross-classifying elements could be assigned. Another example of the need for features on higher nodes would be the traditional features of syntactic ‘mood’ (declarative, interrogative, imperative, exclamative). One of the standard problems for dependency theory has been where to attach such features (Bauer, 1976), since it seems unnatural to attach them to the verb, as though it was the latter’s form which distinguished, say, between a declarative and interrogative clause. (Similar problems have of course arisen in transformational grammar, since there was no mechanism for attaching mood features to clauses rather than to their constituents — Hudson, 1976a:44.) However I am no longer convinced that the traditional mood features are syntactic features...
at all, but believe rather that they are semantic, and the relevant contrasts may be labelled in some way in the semantic structure, having been ‘read off’ the syntactic structure by interpretive rules (in the form of pan-lexical boxes). I have in any case argued elsewhere that the semantic structure of a sentence should show such contrasts, and that they can be clearly distinguished from the pragmatic notions of illocutionary force (Hudson, 1975). In conclusion, I no longer believe that cross-classifying features are needed on nodes representing complex constructions like clauses and phrases.

D. Categorisation of higher nodes

Even if features are not needed on higher nodes, it could be argued that Chomskyan ‘categories’ such as ‘noun-phrase’ and ‘clause’ are needed, in order to define the domains of certain syntactic rules. For example, the rule of gapping applies to clauses, but not to other categories (pace Jackendoff, 1971; for discussion see Hudson, 1976b):

(10) a. John invited Mary, and Bill, Sue.
   b. *my picture of Mary, and your of Sue.

It might be thought that this counted as a reason for marking clauses as such in the grammar, but this is not so, since the rule for gapping simply refers to verbs, and says that two items, each extracted from one of two coordinated pairs (e.g. John and Bill, and Mary and Sue, in (10a)) may be extraposed together, provided only that they both depend on the same verb. This is so in (10a), but not in (10b), since there is no verb for your and of Sue to depend on. Thus there is no need to refer to clauses as such at all. Similarly, unbounded leftward movement rules (so-called) apply only to clauses, but once again the relevant rules may be formulated with reference to verbs, rather than clauses. (At the present this must remain as hand-waving, but I believe it can be shown that it is so.)

E. Headless constructions

In straightforward cases, clauses are represented in dependency terms as verbs with their dependents, and noun-phrases as nouns with their dependents; but what about the non-straightforward cases, where ‘noun-phrases’ have no head noun (e.g. the poor, the biggest), and ‘clauses’ have no verb (e.g. the longer the stem, …; …, (with) his hat over his eyes)? If we want to cover these constructions by the same rules as we formulate for
constituency and dependency clauses with verbs and noun-phrases with noun-heads, we clearly face problems. However, in each case I think the problem is soluble. Taking first noun-phrases with no apparent head, we must accept that words like poor in the poor and biggest in the biggest are not zero-derived nouns, but are genuine adjectives; but this need not worry us, since we may simply replace the term 'noun' by the well-established term 'substantive' as the name for words like boy, and broaden the denotation of 'noun' to include both substantives and adjectives (thus following Jespersen's usage; see 1924:72). Thus if a rule needs to cover the constructions traditionally referred to as noun-phrases, with or without heads, it just refers to 'nouns', and this will automatically allow the constructions to have either substantives or adjectives as heads. If we then want to specify, for example, that the modifier of a noun may be either an adverb or an adjective, then we may invoke the distinction between substantives and adjectives. The consequence of this analysis is that constructions like the poor and the biggest are treated as syntactically complete, just like the boy, and any generalisation about the distribution of nouns (and their dependents) will automatically cover them too. Turning now to the other apparent problem, verbless clauses, we already have the outlines of a solution in the proposed analysis of the relation between subjects and predicate adjectives. If we say that the adjective in The boy is big depends directly on the subject (as well, perhaps, as on the copula), then there is no need for a verb to hold the subject and the predicate adjective together; and the same will be true if the predicate is not an adjective but a prepositional phrase like over his eyes, or a noun-phrase. This leaves the problem of how to bring together verbless clauses and clauses with verbs for rules that need to refer to them all, indiscriminately. To solve this problem it may be that we need to consider a radical revision to standard dependency analyses, perhaps somewhat along the lines suggested by Werner (1975), with the subject as the pivot on which the rest of the clause depends instead of the verb, but this revision takes us well beyond the bounds of the present paper. Suffice it to say, then, that both 'headless noun-phrases' and 'verbless clauses' can be handled in a satisfactory way without invoking constituent structure and the categories 'noun-phrase' and 'clause'.

F. Propositions and terms

Finally it could be argued that constituent structure is necessary in syntax because it is necessary in semantics, and the two levels need to be kept in step as far as their treatment of sentence structure is concerned. According
to this view, constituent structure is appropriate for semantics because semantic structure is made up of propositions and terms, each of which has a complex internal structure. This is a particularly weak argument as there is no reason to suppose that the 'proposition' is any more necessary in semantics than the 'clause' is in syntax, and similarly for the 'term' and the 'noun-phrase', and indeed dependency analysis seems to be especially appropriate for semantics, as dependency theorists have often pointed out (e.g. Tesnière, 1959:42; Heringer, 1970:78). (One of the attractions of dependency theory as a model for semantic representations is that it seems to offer the possibility of semantic representations in terms of cognitive networks, which are more flexible than the strictly hierarchical organisation implied by constituent structure.) Once again we must leave this particular issue without proper discussion, but it is at least clear that semantic structure does not offer specially obvious evidence for constituency in syntax.

5. Conclusions

The conclusion to which this paper has led is that syntactic analysis needs dependency but does not need constituency. In other words, the only items to which syntactic rules need refer are (classes of) words (including clitics), and similarly the only items to be represented by nodes in a structural diagram are words. (I assume that the internal structure of words in terms of morphemes or morphological processes should be dealt with by means of separate rules of morphology rather than by the syntax; here too I am following the European rather than the American tradition.)

I have also argued (section 3) that 'dependency' is just another name for the relation between a 'frame' and its 'slots' — in other words, for the relation of 'strict subcategorisation', which in transformational grammar is relegated to the lexicon. If the claims of this paper are right, then, they constitute strong further support for the 'pan-lexical' model in which the grammar is virtually identified with the lexicon. It has another implication for this model, however: one source of complexity is removed in as much as there no longer need to be 'pan-lexical entries' referring to internally complex items such as clauses and noun-phrases. Like the traditional lexicon, the pan-lexicon refers only to words.

As far as the theory of 'Daugher-dependency grammar' is concerned, the implications seem to be fairly serious. Contrary to what I argued in Hudson, 1976a, there is no need to combine the dependency approach with the constituency approach, and what we need is something much nearer to traditional dependency theory than what I offered. Nor do we
need rules relating clauses and phrases to their constituents, so we may dispense with the whole apparatus of ‘daughter-dependency rules’. The useful life of the name ‘Daughter-dependency grammar’ seems to have definitely come to an end, but the notion of dependency, which plays a large part in the theory, lives on.

References


