**Grammar without functional categories**

Richard Hudson
University College London

Abstract
The paper considers the notion `functional category' and concludes that, at least as far as overt words are concerned, the notion is ill-founded. First, none of the definitions that have been offered (in terms of function words, closed classes or non-thematicity) are satisfactory, because they either define a continuum when we need a sharp binary distinction, or they conflict with the standard examples. And second, the two most commonly quoted examples of word classes that are functional categories cannot even be justified as word classes. Complementizers have no distinctive and shared characteristic, and Determiners are all pronouns which are distinguished only by taking a common noun as complement - a distinction which is better handled in terms of lexical valency than in terms of a word class.

## 1. Functional categories[1]

The notion `functional category' has played a major part in recent discussions of syntactic theory. For example, Chomsky introduces it as follows:

> Virtually all items of the lexicon belong to the *substantive* categories, which we will take to be noun, verb, adjective and particle, ... The other categories we will call *functional* (tense, complementizer, etc.), ... (Chomsky 1995:6).

He later suggests that only functional categories carry strong features (ibid:232), and that they

> `have a central place in the conception of language .. primarily because of their presumed role in feature checking, which is what drives Attract/Move.' (ibid:349)

Similarly, it has been suggested that functional categories cannot assign theta-roles (Abney 1987, Radford 1997:328), and that they can constitute the ‘extended projection' of their complement's lexical category (Grimshaw 1991, Borsley and Kornfilt, this volume). According to the `Functional Parameterization Hypothesis', functional categories are the special locus of the parameters that distinguish the grammars of different languages (Atkinson 1994:2942, Smith and Tsimpli 1995:24, Ouhalla 1991, Pollock 1989), and Radford (1990) has suggested that they are missing from child language.

Such claims have not been restricted to the Chomskyan school: in HPSG we find the suggestion that only functional categories may act as `markers' (Pollard and Sag 1994:45) and in LFG that functional categories always correspond to the same part of f-structure as their complements (Bresnan, this volume).

Any notion as important as Functional Category[2] should be subjected to the most rigorous scrutiny, but this seems not to have happened to this particular construct. Instead it has been accepted more or less without question, and has become part of mainstream theorizing simply through frequent mention by leading figures. I shall suggest in this paper that the notion is in fact deeply problematic. The attempts that have been made to define it are flawed, and all the individual categories that have been given as examples raise serious problems. The issues raised here should at least be considered by proponents of the notion. If the

criticisms are well founded, the consequences for syntactic theory are serious; but even if these worries turn out to be groundless, the debate will have made this key notion that much clearer and stronger.

To avoid confusion it is important to distinguish three kinds of `category', which we can call Word Category, Sub-word Category and Position Category. Word categories are simply word classes - Noun, Determiner and so on. Every theory accepts that there are words and that these fall into various classes, so Word Category is uncontroversial even if the validity of particular word categories is debatable. Sub-word categories are elements of syntactic structure which (in surface structure) are smaller than words - morphemes or zero. (Clitics are on the border between these types, but it makes no difference here whether we classify them as belonging to word or sub-word categories.) The obvious example of a sub-word category is INFL, to the extent that this corresponds merely to the verb's inflection or to zero. It is a matter of debate whether sub-word categories have any place at all in syntactic theory, and most theories at least restrict their use (e.g. by Pullum and Zwicky's principle of Morphology-free Syntax - Zwicky 1994:4477). This issue is orthogonal to the questions about functional categories that I wish to raise in the present paper, so I shall avoid it by focussing on word categories.

Position categories are a further extension of word and sub-word categories, where the name of the category is used to label a structural position. For example, the standard `Barriers' analysis of clause structure recognises C and I as positions in an abstract tree structure. The labels `C' and `I' are abbreviations of Comp (for Complementizer) and INFL (for Inflection), but the link to the original word and sub-word categories is broken because these positions may be either empty, or filled by a verb - which is not, of course, classified inherently as a complementizer or inflection, even if the relevant feature structures overlap. Such position categories are also controversial and raise problems of both fact (Hudson 1995) and theory which go beyond the scope of this paper.

The central question to be addressed, therefore, is the status of the construct Functional Word Category, **FWC**, rather than the more general question of functional categories. Given this focus it is important to acknowledge that sub-word and position categories are also central to the discussion of functional categories. The conclusion of this paper is that FWC is not justified, but even if this conclusion is correct it will still remain possible that some sub-word and position categories are functional. I shall argue, for example, that Complementizer is not a valid word category, but it could still be true that the position category C is valid. On the other hand, FWCs are part of the evidence which is normally adduced in support of the more abstract categories, so anything which casts doubt on the former must affect the credibility of the latter.

The paper will move from the particular to the general. Sections 2 and 5 will discuss the categories Complementizer and Determiner, which are among the most frequently quoted examples of FWC. The discussion of Pronoun in section 3 is needed as a preparation for the proposed analysis of determiners, as is the general theorizing about valency and classification in section 4. The conclusion of these sections will be that neither Complementizer nor Determiner is a valid word class, so (a fortiori) neither can be a functional word category. Sections 6 and 7 will consider two standard definitions of FWC: as a class of function words and as a closed class. It will be argued that Function Word is indeed an important and valid

construct with a good deal of empirical support, and similarly (but to a lesser extent) for Closed Class. However I shall also show that neither of these two concepts is suitable as a basis for FWC. The conclusion, in section 8, will be that FWC plays no part in grammar, though there may be a small role for Function Word. Encouragingly, Cann (this volume) reaches a similar conclusion by a different route.

## 2. Complementizer

The following argument rests in part on a general principle of categorisation which should be laid out before we proceed. The principle amounts to no more than Occam's Razor, so it should be sufficiently bland to be acceptable regardless of theoretical inclinations.

(1) **Principle 1**

A word-class should be recognised only if it allows generalisations which would not otherwise be possible.

The classic word-classes satisfy this principle well. Take Noun, for example. Without it, we could say that some words can head a verb's subject, and that some words can head its object, but in each case we should have to simply list all the words concerned. Given the category Noun, however, we can express the generalisation that the lists are the same - not to mention the lists needed for many other facts about distribution, morphology and semantics. Similarly for Auxiliary Verb, a word-class defined by a collection of characteristics which include negation, inversion, contraction and ellipsis. Without this word-class it would not be possible to show that these characteristics all applied to the same list of words. In contrast with these very well-established classes, some traditional word-classes have a rather uncertain status, with Adverb as the classic case of a `dustbin' that has very few characteristics of its own, though probably enough to justify it among the major word-classes.

In short, every word class must earn its place by doing some work in the grammar. How does the word-class Complementizer fare when tested against this principle? The history of this class is not encouraging, since its very existence escaped the notice of traditional grammarians; if it really does allow generalisations which would not otherwise be possible, how did traditional grammar manage without it? Even the name Complementizer suggests some uncertainty about the distinctive characteristics of its members: do they introduce complement clauses, or subordinate clauses in general?

In English, the words concerned are (according to Radford 1997:54) *that*, *if* and *for*. Every introductory book tells us that these form a class, with the possible addition of *whether*, but what precisely are the generalisations that this class allows? The answer seems to be that there are no such generalisations. This claim is controversial and requires justification, but before we consider the evidence I should reiterate that we are discussing the `word category', whose members are overt words, and not the `position category' which includes the structural position `C'. I have argued elsewhere (Hudson 1995) that this category is invalid as well, but that is a separate debate[3].

What, then, do all the three core complementizers have in common? As Radford points out (1997:54) they can all introduce a subordinate clause which is the complement of a higher verb or adjective. His examples are the following:

(2) a I think [that you may be right].
b I doubt [if you can help].
c I'm anxious [for you to receive the best treatment possible].

Radford's generalisation is that complementizers:
(A) indicate that the following clause is complement of some other word,
(B) show whether this clause is finite, and
(C) mark its semantic role in the higher clause (which Radford calls its illocutionary force).

Unfortunately these characteristics do not justify Complementizer, as we shall now see.

- Claim A (indicating complement-hood) is false because the clause introduced by a complementizer need not be the complement of another word. *That* and *for* allow a subject link:

(3) a [That you may be right] is beyond doubt.
b [For you to receive the best treatment possible] is important.

Moreover, *for* also allows an adjunct link:

(4) a I bought it [for you to wear].
b A good book [for you to read] is this one.

According to standard analyses of relative clauses, the same is even true of *that*, which is assumed to occur in combination with a zero relative pronoun (Radford 1997:305):

(5) He is someone [that we can identify with].

Furthermore, although it is true that all the complementizers may be used to introduce a complement clause, the same is also true of words which are not complementizers, most obviously the interrogative words.

(6) a I wonder [who came].
b I know [what happened].

It is true that standard analyses assume a zero complementizer in addition to the interrogative word, but the claim is that complementizers `indicate' the clause's function, which must be a claim about overt words.

- Claim B (indicating finiteness) is true, but again not unique to complementizers. The same is in fact true of every word which can introduce a clause: there is no word which allows a clause as its complement without placing some kind of restriction on its finiteness. For example, *why* requires a tensed clause whereas *how* allows either a tensed clause or an infinitival:

(7) a I wonder [how/why he did it].
b I wonder [how/*why to do it].

Similar remarks apply to all the traditional subordinating conjunctions such as *while*, *because* and *unless*, none of which are generally considered to be complementizers.

- Claim C (indicating semantic role) is only partially true, as Radford's own second example illustrates: after *doubt* either *that* or *if* is possible without change of meaning.

(8) I doubt [if/that you can help].

Moreover to the extent that it is true, this characteristic is again not peculiar to complementizers. Most obviously, the same is (again) true of interrogative pronouns.

Having considered and rejected Radford's generalisations we should

consider whether there are any other generalisations that might justify Complementizer. A plausible candidate concerns extraposition: all the complementizers allow extraposition.

(9) a It surprises me [that John is late].
b It is unclear [if it rained].
c It surprises me [for John to be late].

However, if Complementizer was valid this should be the end of the possibilities, but it is not. The same is also true for TO (which is not a complementizer) and to all the interrogative words, including *whether*:

(10) a It surprises me to see John here.
b It is unclear whether/when it rained.

Indeed, extraposition is even possible for some noun-headed phrases, such as those containing nouns like WAY (but not MANNER) and NUMBER:

(11) a It is astonishing the way/*manner she drinks.
b It is astonishing the number of beers she can drink.

These nouns can only be extraposed if they are modified by what is at least syntactically a relative clause:

(12) a *It is astonishing the clear way.
b *It is astonishing the incredibly large number.

Once again Complementizer does not prove particularly helpful. If there is a single thread running through all the phrases that can be extraposed, it may be semantic rather than syntactic.

In short, whatever all three core complementizers have in common does not distinguish them from interrogative words. This suggests that Radford's three claims can and should be handled without mentioning Complementizer. Let us consider how this can be done.

▪ Claim A. To the extent that complementizers do indicate a complement link between the following clause and some preceding word, this is because the latter selects it as the head of its complement. However, words that select complementizers always select specifically. This is illustrated in Table 1, which shows that *think* allows *that* or zero[4] but not *if* or *for*, and so on. Furthermore, almost every verb which allows *if* also allows *whether* and the full range of interrogative pronouns. (The only exception is *doubt*.) In short, no valency statement will ever say: `such-and-such word takes as its complement a clause introduced by a complementizer'[5].

Table 1. Selectional differences among complementizers.

| verb | complement clause | | |
|---|---|---|---|
| | that/zero | if (whether, who, ..) | for .. to .. |
| think | + | 0 | 0 |
| wonder | 0 | + | 0 |
| long | 0 | 0 | + |
| know | + | + | 0 |

■ Claim B. Precisely because different complementizers select different tenses, Complementizer as such will not help in constraining the choice of tense-inflection. This selection must be handled separately for different complementizers: tensed or subjunctive[6] after *that*, tensed after *if*, *to* after *for*.

(13) a I know that Pat is/*be/*to be leader.
b I recommend that Pat is/be/*to be leader.
c I wonder if Pat is/*be/*to be leader.
d I long for Pat to be/*is/*be leader.

■ Claim C. The same logic applies here too. Different complementizers indicate different semantic roles, so verbs will select specific complementizers rather than the general category Complementizer. As mentioned above, almost every verb that selects *if* also allows any interrogative word, which makes Complementizer even less relevant to semantic selection.

In short, Complementizer has no role to play in defining the use of the words *that*, *if* and *for*. It should be noticed that we arrived at this conclusion while considering only the `core' examples, so the status of Complementizer is not likely to be improved by including more peripheral examples like *whether*. On the contrary, in fact, since *whether* is even more like the interrogative words than *if* is. Unlike *if*, but like interrogative words, it allows a following infinitive and a subject link:

(14) a I wonder [whether/when/*if to go].
b [Whether/when/*if to go] is the big question.

However we analyse *whether*, it is unlikely that we shall gain by invoking Complementizer. The conclusion, therefore, must be that Principle 1 rules out Complementizer.

If these words are not complementizers, what are they? We might consider assigning them individually to existing word classes; for example, Haegeman classifies *for* as a `prepositional complementizer', or more simply as a preposition (1994:167), in recognition of the fact that it licenses an accusative subject NP. But even if this is the right analysis for *for*, it is certainly not right for *that* (nor for *if*, though this is less obvious), and in any case it raises other problems. If *for* is a preposition, its projection should presumably be a PP and yet it is said to head a CP. Its classification should explain why a for-clause can be used equally easily as complement, as subject or as adjunct, but no single established category has this distribution. The problems of classifying *that* and *if* are similar, but if anything even more acute.

The alternative to problematic classification is no classification at all - an analysis in which these words are each treated as unique (`syncategorematic'). This is my preferred analysis, as it reflects exactly the outcome of the earlier discussion in which we found that each word is, in fact, unique. Thus *that* is simply a word, and so are *if* and *for*; they are recognised as lexical items, but have no grammatical features and belong to no categories. When the grammar mentions them, it defines them simply as lexical items whose word class is irrelevant.

The only complementizer whose classification is at all straightforward is *whether*, whose similarities to interrogative pronouns have already been pointed out. At least some linguists (e.g. Larson 1985) argue that it is in fact a wh-pronoun, and I myself agree with this conclusion (Hudson 1990:374). Even this analysis is problematic, however, because *whether*, not being a true pronoun, has no

grammatical role within its complement clause. In this respect it is just like *if* and *that*, as well as all the subordinating conjunctions, so it is at best a highly unusual wh-pronoun.

In conclusion, we have found no justification for Complementizer because there seem to be no generalisations that apply to all the core members. This means that it is not enabling the grammar to express any generalisations, so according to Principle 1 Complementizer does not exist as a category, so (a fortiori) it is not a FWC.

**3. Pronoun**

We now make a slight detour from the main discussion in order to establish the controversial claim that pronouns are nouns, which will play an important part in the next section's discussion of determiners. As it happens Pronoun is itself claimed to be a FWC (Radford 1997:48-9) on the grounds that pronouns are determiners and that Determiner is a FWC. The status of Pronoun as a FWC is thus tied up with that of Determiner, which is the topic for the next section. If, as I shall argue, Determiner is not a FWC, Pronoun cannot be either. However, the argument there presupposes a specific set of analytical assumptions about the classification of pronouns which are non-standard: that Pronoun is a subclass of Noun, and that determiners are pronouns. Before we can consider the status of Pronoun as a FWC, therefore, we must attend to these analytical questions. Why should we take Pronoun as a subclass of Noun?

Radford's discussion considers only one kind of pronoun, personal pronouns, but it is uncontroversial that there are other subclasses including reflexive, reciprocal, interrogative, relative, demonstrative, negative, distributive and compound. These subclasses are presented in Table 2, a reminder that our Pronoun is the traditional word class, not the much smaller category that Chomsky calls Pronoun (1995:41). His category excludes reflexive and reciprocal pronouns and is claimed always to refer, so it presumably excludes the indefinites as well.

Table 2. Subclasses of pronoun

| class | definiteness | examples |
|---|---|---|
| personal | definite | I/me, you, he/him, one(?) |
| reflexive | | myself, yourself, himself |
| reciprocal | | each other, one another |
| relative | | who, which, whose, where, when |
| demonstrative | | this/these, that/those |
| possessive | | mine, yours, his; -'s[7] |
| distributive | | each |
| universal | indefinite | all, both |
| existential | | some, any, either |
| negative | | none, neither |
| interrogative | | who, what, which, how, why |
| compound | | someone, anybody, nothing, everywhere |

What all these words share is the ability to be used in the range of phrasal environments that are available for a full noun phrase; for example, *they/them* has almost exactly the same overall distribution as *the students*:

(15) a They/the students have finished.
b Have they/the students finished?
c I like them/the students.
d We talked about them/the students.
e I saw them and the students.

There are obvious differences of detail which apply to specific subclasses (e.g. personal pronouns cannot be used before possessive *-'s*, reflexives cannot be used as subject) but the overall similarity between pronouns and noun phrases is striking. On the other hand, pronouns are different from common nouns in several ways, but in particular in their inability to combine with a preceding determiner (**the I*, **a who*, etc). The range of possible modifiers is also strictly limited compared with common nouns, though some allow adjuncts (*someone nice*, *who that you know*), and depending on one's analysis, some allow complements (*who came*, *this book*)[8].

Traditional grammar recognises Pronoun as a supercategory, one of the basic parts of speech alongside Noun, and links the two classes by the enigmatic statement that pronouns `stand for' nouns. In modern phrase-structure analyses the similarity is shown at the phrase level by giving the same label to the phrases that they head. This label is either DP or NP, depending on analysis, but this choice is crucial to the following argument so we shall keep it open by temporarily adopting

the neutral term `Nominal Phrase'. Thus *they* is not only a pronoun but also a nominal phrase, and *the students* is a nominal phrase.

Suppose we accept this analysis, and also the general X-bar principle that a phrase's classification must be that of its head word. Given these two assumptions, what follows for the classification of *they* and *students*? (The next section will discuss the classification of *the*.) We have to choose between two answers:
A1 (the standard analysis) *They* belongs to the same class as *the*, which (for the time being) we can call `determiner'; we shall revise the name in the next section.
A2 (my preferred analysis) *They* belongs to the same class as *students*: Noun.
The choice between the two analyses revolves around the analysis of the one-word phrase *students*, where there is no determiner:

(16) a I like students.
b I found students waiting for me.

If *students* really is the only word in this phrase (as I shall argue), its classification must project to the phrase which must therefore be a noun phrase, so *they* must also head a noun phrase and must itself be a noun. If, on the other hand, the phrase *students* is headed by a covert determiner, it must be a determiner phrase and *they* must be a determiner. The standard analysis stands or falls with the covert determiner. We shall now consider some arguments for it, and an argument against it.

One argument for the covert determiner is that it is required by the DP analysis (Abney 1987), which is widely accepted. If *the* is the head of *the students*, the phrase *the students* must be a projection of *the*: DP. Therefore the one-word phrase *students* must be a DP, with a covert determiner. However this argument rests on the assumption that *the* is not a noun. If it were, both phrases would be NPs and there would be no need for a covert determiner. The next section will be devoted to this claim, so I shall simply note at this point that (so far as I know) the category Determiner has always been taken for granted in discussions of the DP hypothesis so there is no `standard' evidence for it. I cannot prove that such evidence does not exist, but I shall prove that a coherent analysis can be built without this assumption.

Radford gives some more direct evidence in support of the covert determiner (1997:152). He points out that (if it exists) it selects the same kind of complement as *enough*[9], namely a plural or mass common noun:

(17) a I found things/stuff/*thing.
b I found enough things/stuff/*thing.

It also has an identifiable determiner-like meaning, which is either `existential', as in the above examples, or `generic', as in the following:

(18) I collect things/stuff/*thing.

In short, the covert determiner is a normal determiner in its semantic and syntactic characteristics, so its only peculiarity is its lack of phonology.

However this argument is open to various empirical objections:

- It is because of the semantics of *enough* that its complement must be either plural or mass; so we might predict that the word meaning `enough' in any other language will have the same restriction. In contrast, the restrictions on the hypothetical covert determiner vary between languages - and in particular there are many languages where it would not be restricted as in English. So even if the covert determiner selects the same kind of complement as some determiners, it

does not select in the same way.

- The word *sufficient* imposes the same restriction on the semantics of the noun as does *enough*, and also, like *enough*, it excludes (other) determiners.

(19) a enough/sufficient things/stuff/*thing
b some/the *enough/*sufficient stuff

But to judge by the adverb *sufficiently*, *sufficient* is an adjective, which weakens the argument for a covert determiner wherever the noun must be plural or mass. It could equally be argued either that *enough* is an adjective, or that there is a covert adjective whose meaning and distribution are like those of *sufficient*.

- The fact that the covert pronoun allows generic and existential reference only shows that it places no restrictions on that aspect of reference. In contrast, overt determiners typically do restrict it; for example, *some* excludes generic reference.
- What the covert pronoun does exclude is `definite' reference - reference to an object which is already known to the addressee; but definiteness is one of the main differences between common nouns and proper nouns, which are inherently definite. This suggests that the indefiniteness of the one-word phrase *students* is inherent to the common noun, rather than due to a covert determiner.

We now consider the alternative of the one-word phrase *students* in which there is no covert determiner, my analysis A2. The syntactic restrictions can be reversed: instead of saying that the covert determiner selects plural or mass commoun nouns, we have Rule 1[10].

(20) **Rule 1**
Singular countable common nouns must be the complement of a determiner[11].

As for the indefinite meaning of *students*, we can follow the suggestion made above: just as proper nouns are inherently definite, so common nouns are inherently indefinite. In both cases the default meaning may be overridden - in the terms of Pustejovsky (1991), `coerced' - by the meaning imposed by a determiner; so a common noun may be coerced into definiteness (*the students*), and a proper noun into indefiniteness (*a certain John Smith*).

In this analysis, *a* has a special status similar to that of the dummy auxiliary *do*. There are purely syntactic patterns such as subject inversion which are only available for auxiliary verbs; so if the meaning to be expressed does not require any other auxiliary verb, *do* can be used because it has no meaning of its own, and therefore does not affect the meaning. Similarly for *a*: Rule 1 sometimes requires a determiner, so if no other determiner is required by the meaning to be expressed, a `dummy' determiner is needed which will not affect the meaning. This is *a*, whose only contribution to meaning is to restrict the sense to a single individual (thereby excluding plural and mass interpretations).

This analysis of *a* leads to a positive argument against the covert-determiner analysis. The argument involves predicative uses such as the following:

(21) a They seem (*some/*no/*the) good linguists.
b He is a/*any/*no/*the good linguist.

One of the restrictions imposed by verbs like *seem* is that a complement nominal must not contain a determiner other than *a*[12]. Ignoring the exception of *a*, this restriction is quite natural if we assume, first, that determiners relate semantically to a nominal's referent rather than to its sense, and, second, that a predicative nominal has no referent (as witness the oddity of the question *Which good linguists*

*do they seem?*). On that assumption, it is natural to assume that *good linguists* has no determiner at all, rather than that it has a covert one: so in Radford's terms, it is an NP, not a DP. But in that case it is hard to explain the determiner *a* in the second example - why is *a* not only possible, but obligatory? The DP analysis forces a disjunction: the predicative complement of verbs like *seem* is either an NP or a DP headed by *a*. (Alternatively, the disjunction may be between *a* and the covert determiner as the head of the DP.)

Now consider the no-determiner analysis. Suppose we assume, first, that *seem* requires its complement to have no referent, and secondly, that most determiners have a referent. These two assumptions immediately explain the basic fact, which is the impossibility of determiners. The other fact is the appearance of *a*, which is also explained by two assumptions that we have already made: that Rule 1 requires a determiner before singular countable common nouns, and that *a* does not have to have a referent - it is a semantically empty, dummy, word like auxiliary *do*. The result is that *a* is both obligatory and possible with *linguist*, but neither needed nor possible with *linguists*.

Rather tentatively, therefore, we may be able to conclude that nominal phrases need not contain a determiner, so they must be projections of Noun. Therefore pronouns too must be nouns. We can still distinguish Pronoun from Common Noun and Proper Noun as subclasses of Noun. As shown in Table 2, Pronoun has its own subclasses, and no doubt the same is true for Common Noun. The hierarchical structure is shown (using Word Grammar notation) in ?, and more details of the assumed analysis can be found in Hudson (1990:268-335).

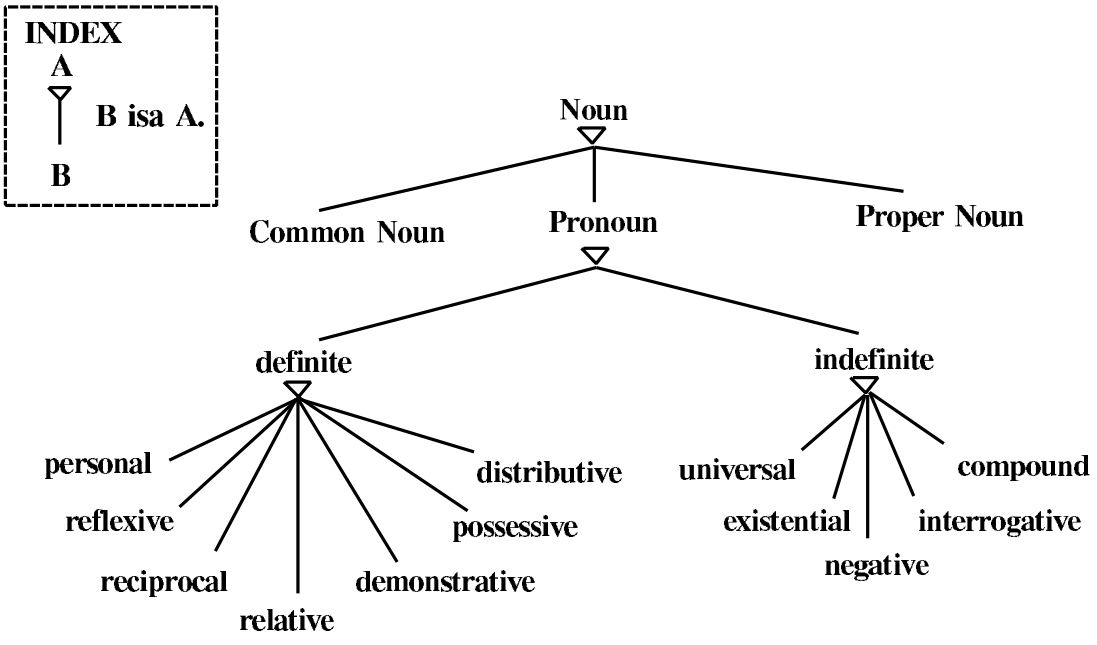


**Figure 1**

The conclusion of this section is that Pronoun is a subclass of Noun. This view is quite traditional and fairly widely held (Huddleston 1984:96), but it is controversial. On the one hand the traditional part-of-speech system treats Pronoun

as a separate superclass, and this tradition persists in most modern descriptive grammars (e.g. Quirk et al 1985:67). On the other hand, the modern DP analysis treats it as a subclass of Determiner, which itself is a distinct supercategory (Radford 1997:154). If the present section is right, at least this part of the DP analysis must be wrong.

**4. Valency and its irrelevance to classification**

In preparation for the discussion of determiners we must establish another general principle, which is merely a particular application of Principle 1 (Occam's Razor). It concerns the treatment of valency (alias subcategorization), the restrictions that words place on their complements[13]. Various devices are available for stating these restrictions: subcategorization frames, Case-marking, SUBCAT lists, linking rules and so on. However formulated, these restrictions can be stated on an item-by-item basis, as facts about individual lexical items. There is no need to recognise a word class for every valency pattern unless the pattern correlates with some other shared characteristic. In fact, more strongly, it would be wrong to recognise a word class because the class would do no work. As we can see in the following abstract example, it would actually make the grammar more complex without permitting any additional generalisation.

Given some valency pattern V which is shared by words A, B and C the simplest grammar relates V directly to A, B and C, giving at most[14] three rules:

(23) a A has V.
b B has V.
c C has V.

Now consider the alternative grammar in which A, B and C are grouped into a word class W, whose sole defining characteristic is that its members have valency pattern V. In this grammar there must be four rules, because the membership of A, B and C in W must be stipulated[15]:

(24) a A isa W.
b B isa W.
c C isa W.
d W has V.

So long as V is the sole characteristic shared by these three words, the grammar with W is clearly inferior to the one without it.

In short, valency facts have the same status as any other facts. The above conclusion follows directly from Principle 1, but we shall state it as a separate principle:

(25) **Principle 2**
A word class should not be recognised if its sole basis is in valency/subcategorisation.

For example, if a verb's lexical entry shows that it needs a particle (e.g. *give up*), there is no point in duplicating this information by also classifying it as a `particle-taking verb' unless such verbs also share some other characteristic. So long as the complementation pattern is their only shared feature, the class Particle-taking Verb is redundant.

In most cases this principle is quite innocuous. It does, however, conflict with the traditional Bloomfieldian idea that differences of `distribution' justify differences

of word class. This idea is still widely accepted (or at least taught):

> `The syntactic evidence for assigning words to categories essentially relates to the fact that different categories of words have different *distributions* (i.e. occupy a different range of positions within phrases or sentences).' (Radford 1997:40; author's emphasis)

Principle 2 means that some distributional differences are not relevant to categorisation, because they are best handled by means of lexical valency.

Consider, for instance, the traditional classification of verbs as transitive or intransitive. These terms are simply informal descriptions of valency patterns which can be stated better in valency terms so far as they correlate with nothing else. Indeed, valency descriptions may be preferable to word classes even when there are other correlating characteristics. For example, as the classic Relational Grammar analyses showed (Blake 1990) it is better to describe the facts of the French `*faire-faire*' construction in terms of valency than in terms of transitive and intransitive verbs.

(26) a Paul fait rire Marie.
Paul makes laugh Mary.
`Paul makes Mary laugh.'

b Paul fait lire la lettre à Marie.
Paul makes read the letter to Mary
`Paul makes Mary read the letter.'

Described in terms of verb classes, as in a) below, the facts appear arbitrary; but explanation b) allows them to follow from the assumption that a verb cannot have two direct objects.

(27) a The direct object of *faire* is demoted to an indirect object if its infinitive complement is transitive.

b The direct object of *faire* is demoted to an indirect object if it also has a direct object raised from its infinitive complement.

It should be noted that Principle 2 is not a wholesale `rejection of distributionalism' (as one reader complained), but simply a recognition that the syntactic distribution of a word has two separate components. One component involves its relations to `higher' structures, via the phrase that it heads. Thus, when seen as head, a preposition is used in `prepositional' environments, a noun in nominal environments, and so on. These are the distributional facts for which word classes are essential. However the other component involves its valency, its relations to `lower' structures. Here word classes are less helpful because the facts concerned vary lexically and/or semantically: different members of the same word class allow complementation patterns which vary in complex ways that have little to do with word classes, but a great deal to do with semantics.

## 5. Determiner

Turning then to Determiner, this is another class which Radford presents as a functional category (1997:45-8). In this case I shall use Principle 2 to argue that there is in fact no word class Determiner because Determiner would be a subclass that was defined solely by valency. As in other analyses I shall classify determiners with pronouns, but I shall also argue that the superclass that contains both determiners and pronouns is actually Pronoun, not Determiner. The analysis will build on the idea of section 3 that pronouns are nouns.

The first step in the argument is to establish that many determiners can also be used as pronouns. This overlap of membership has often been noticed, and has led to various analyses in which pronouns are treated as determiners (Radford 1997:154). The earliest of these analyses was Postal (1966), which argued that pronouns are really `articles' which are generated in what would nowadays be called the Determiner position, and which may be followed by a noun as in *we linguists*. The DP analysis of nominals continues the same tradition in which the classes of determiners and pronouns overlap, so the general idea is now very familiar and widely accepted. As Radford points out (1997:49) some of the words vary morphologically according to whether they are used `as pronouns' or `as determiners'; for example, *none* is the pronoun form that corresponds to the determiner form *no*, and similarly for *mine*/*my*, *yours*/*your*, *hers*/*her*, *ours*/*our* and *theirs*/*their*. However the recent tradition assumes that these variations can be left to the morphology and ignored in the syntax. This seems correct (Hudson 1990:269).

The second step involves the syntactic relationship between the determiner and the common noun. The `DP' tradition that Radford reviews takes the determiner as the head of the phrase. This is also the analysis that I have advocated for some time (Hudson 1984:90), so I shall simply accept it as given. One of the benefits of the analysis is that it explains why the common noun is optional after some determiners but obligatory after the others (namely, *the*, *a* and *every*): each determiner has a valency which not only selects a common noun but decides whether it is obligatory or optional, in just the same way that a verb's valency determines the optionality of its object. In other words, the lexical variation between determiners that we shall review below is just what we should expect if the common noun is the determiner's complement.

The final step is to show that this analysis is only partially right. Determiners are in fact pronouns, rather than the other way round. This may sound like a splitting of hairs but it will make a great deal of difference to our conclusion. This alternative is not considered in the DP tradition so there is no case to argue against. In favour of it we can cite the following facts:

- When a determiner/pronoun occurs without a common noun, it is traditionally called a pronoun, not a determiner. It seems perverse to call *she* and *me* determiners in *She loves me*, as required by the DP analysis. In contrast the traditional analysis would (incorrectly) treat *this* in *this book* as an adjective, so it is no more perverse to call it a pronoun than to call it a determiner.
- Almost every determiner can also be used as a traditional pronoun, but most traditional pronouns cannot also be used as determiners. As mentioned earlier, the only determiners that cannot be used without a common noun are the articles *the* and *a*, and *every*. If we ignore the morphological variation discussed above, all the others can be used either with or without a common noun:

(28) a I like this (book).
b Each (book) weighs a pound.
c I found his (book).
d I found the *(book).
e Every *(book) weighs a pound.

Given that Determiner is almost entirely contained in Pronoun, it seems perverse to call the superset Determiner.

It seems, therefore, that `determiner' may be just an informal name for a particular kind of pronoun, namely a pronoun whose complement is a common noun (or, in more orthodox terms, a phrase headed by a common noun). I believe I may have been the first to suggest this (Hudson 1984:90), but others have arrived independently at the same conclusion (Grimshaw 1991, Netter 1994).

As promised, this change of terminology has far-reaching consequences. The pronouns that allow nominal complements are scattered unsystematically across the subclasses which we distinguished in Table 2, with representation in nine of the subclasses: personal (*we*, *you*), relative (*whose*), demonstrative (*this/these*, *that/those*), possessive (*my*, *your*, etc), distributive (*each*, *every*), universal (*all*, *both*), existential (*some*, *any*, *either*), negative (*no*, *neither*) and interrogative (*which*, *what*). (The classification also needs to accommodate the articles *the* and *a*, but this is irrelevant here.) So far as I know there is no other characteristic that picks out this particular subset of pronouns. It is true, for example, that determiners are also distinguished by the fact of being limited to one per nominal phrase - e.g. unlike Italian, we cannot combine both *the* and *my* to give **the my house*. But this simply reflects a more general fact about pronouns: that they are not allowed as the complement of a pronoun. So long as the complement of a pronoun is limited to a phrase headed by a common noun, one-per-NP restriction will follow automatically[16].

What this observation suggests is that valency is the sole distinguishing characteristic of determiners; in short, that determiners are just the subset of pronouns which happen to be `transitive'. If this is so, Determiner is ruled out by Principle 2 as redundant. Instead of classifying *this* as a determiner, therefore, we just say that it allows a complement, and similarly for all the other determiners. Admittedly this misses an important shared characteristic of determiners, which is that their complements are common nouns; but this generalisation can be captured in terms of Pronoun. Indeed, we can even generalise across subclasses of pronoun about their valency, but without invoking Determiner. The following mini-grammar suggests how determiners should be treated[17]:

(29) a *which* is an interrogative pronoun.
b *which* allows a complement.
c *this* is a demonstrative pronoun.
d A demonstrative pronoun allows a complement.
e The complement of a pronoun is a common noun.

Rule (b) treats the valency of *which* as an arbitrary fact about this pronoun, whereas (d) generalises across both the demonstrative pronouns; and (e) defines the possible complements for all pronouns (without, however, implying that all pronouns allow a complement).

If determiners really are transitive pronouns, two things follow. First, so-called DPs must in fact be NPs because their head, the determiner, is a pronoun and pronouns are nouns. Therefore the head projects to NP, not DP. Secondly, and most importantly for present purposes, at least in English there is no category Determiner so Determiner cannot be any kind of category, functional or otherwise. If correct this conclusion should be worrying for those who advocate FWCs because we have now eliminated the two most quoted examples of FWC, Complementizer and Determiner.

## 6. FWCs as classes of function words

We turn now to a more general consideration of FWC in terms of its defining characteristics: what is it about a category that makes it qualify as a FWC? A satisfactory answer will draw a boundary around FWCs which satisfies three criteria.

Firstly, the boundary should be `in the right place', corresponding to the general intuitions that Auxiliary Verb and Pronoun are candidates for FWC but Full Verb and Common Noun are not. How can we decide where the right place is, other than by invoking these intuitions? Fortunately we already have a criterion: Principle 1. The category FWC will be justified to the extent that it allows generalisations which are not otherwise possible. This means that we must look for distinct characteristics that correlate with each other, comparable to those which distinguish (say) Verb from Noun. Without such correlations, any choice of criteria is arbitrary; but with them it becomes a matter of fact. If it turns out that criteria A and B correlate, and both characterise categories X, Y and Z, then it is just a matter of fact, not opinion, that X, Y and Z belong to a supercategory. If A and B are among the standard characteristics of FWCs, then we can call this supercategory `FWC' - provided it satisfies the remaining criteria.

Secondly, the boundary should have `the right degree of clarity'. By this I mean that it should probably be an all-or-nothing contrast, and not a matter of degree. This criterion is important because of the kinds of generalisations that are expressed in terms of FWC - categorical generalisations such as the one quoted earlier about the links between FWC and feature checking. It would be impossible to evaluate such generalisations if categories had different degrees of `functionality'.

And thirdly, FWC should have the `right kind of membership', because as a supercategory its members must be categories, not words. If it includes some members of a category, it must include them all.

We start then with a widely accepted definition of FWC which links it to the traditional notion Function Word (**FW**):

> `... functional categories. The lexical/functional dichotomy is rooted in the distinction drawn by descriptive grammarians ... between two different types of words - namely (i) contentives or content words (which have idiosyncratic descriptive content or sense properties), and (ii) function words (or functors), i.e. words which serve primarily to carry information about the grammatical properties of expressions within the sentence, for instance information about number, gender, person, case, etc.' (Radford 1997:45)

As Radford says, this distinction is often drawn by descriptive grammarians, and is beyond dispute in the sense that there is an obvious difference between the meaning of a word like *book* or *run* and that of a function word such as *the* or *will*.

Cann (this volume) surveys a wide variety of criteria by which Function Word (his `FE') has been defined, and which all tend to coincide. The criteria are semantic, syntactic and formal, and in each case FWs are in some way `reduced' in comparison with typical words. Semantically they lack what is variously called the `denotational sense' or `descriptive content' of words like *tomato*, syntactically their distribution is more rigidly restricted, and formally they tend to be short. We could even add to Cann's list of formal distinctions that are relevant to English:

- In terms of phonology, only FWs may have /ə/ as their only vowel.

- In terms of spelling, only FWs may have fewer than three letters.[18]
- In terms of orthography, only FWs are consistently left without capital letters in the titles of books and articles.[19]

In short, there can be no doubt about the reality and importance of FW as the carrier of a large number of correlating characteristics.

Nevertheless, FW does not justify FWC because it fails the second and third criteria. As far as clarity of boundaries is concerned, there are too many unclear borderline cases which either have one of the characteristics only to some degree, or which have some characteristics of FWC but not all of them. These uncertainties and conflicts are well documented by Cann. For example there is no clear cut-off between having descriptive content and not having it, so there are borderline cases such as personal pronouns. Unlike clear FWs these each have a distinct referent and they also involve `descriptive' categories such as sex, number and person. Perhaps because of this they tend to be capitalized in book titles (contrary to the pattern for FWs mentioned above). Similarly, it is hard to see where to draw the line, on purely semantic grounds, between the FW *and* and what is presumably not a FW, *therefore*: where for example would *so* belong? If even one of the characteristics of FW is indeed a matter of degree, determined by the `amount of meaning' carried, it cannot be mapped onto the binary contrast between functional and substantive categories. This is not an isolated case: formal brevity is even more obviously a matter of degree, and it is hard to imagine that there is a clear threshold for syntactic limitation.

Similarly for the third criterion, concerning the membership of FWC. If it is indeed a set of word classes, then for any given word class either all of its members should belong to FWC, or none of them should. There should be no split classes (what Cann calls `crossover' expressions). And yet, as Cann points out, split classes do seem to exist. The clearest example of a split word class is Preposition. Some prepositions are very clear FWs - for example, *of*, *at*, *in*, *to* and *by* all qualify in terms of both meaning and length. Indeed, all these prepositions have regular uses in which they could be said to have no independent meaning at all:

(30) a the city of Glasgow
b at home
c believe in fairies
d take to someone
e kidnapped by gangsters

In each example the preposition is selected by the construction, and does not contrast semantically with any other. On the other hand, there are also prepositions like *during* and *after* which have just as much meaning as some adverbs - indeed, there are adverbs which are synonymous except for being anaphoric (*meanwhile*, *afterwards*). If the adverbs are content words, the same should presumably be true of their prepositional synonyms. But if this is right, some prepositions are content words and some are FWs. This should not be possible if it is whole word classes that are classified as FWCs.

A similar split can be found even among nouns and verbs, though in these classes there are very few members that have the `wrong' classification. The anaphoric *one* is an ordinary common noun (with the regular plural *ones*):

(31) a He lost the first game and won the second one.

b He lost the first game and won the other ones.

*One* (in this sense) has no inherent meaning except `countable', since it borrows its sense from its antecedent by identity-of-sense anaphora. But it behaves in almost every other respect just like an ordinary common noun such as *book* - it accepts attributive adjectives, it inflects for number, and so on. Similarly for the British English anaphoric *do*, which is an ordinary non-auxiliary verb:

(32) a He didn't call today, but he may do tomorrow.

b A. Does he like her? B. Yes, he must do - just look how he talks to her.

This too is completely empty of meaning - it can borrow any kind of sense from its antecedent, stative or active - and yet we use it syntactically exactly as we use an ordinary verb like *run*. Their lack of meaning suggests that both these words are function words - an analysis which is further supported by their shortness (*do* has only two letters, and *one* has a variant with /ə/ which is often shown orthographically as *'un*: *a big 'un*). And yet they are clear members of classes whose other members are content words.

A similar problem arises with a widely accepted definition of FW in terms of `thematicity' (Radford (1990:53), quoting in part Abney (1987:64-5)). According to Radford, FWs are `nonthematic', by which he means that even if they assign a theta role to their complement, they do not assign one to their specifier: for example, consider the auxiliary *may* in the following:

(33) It may rain.

This has a thematic relationship to its complement `it..rain', but not to its subject. This may well be a general characteristic of FWs, but it does not apply in a uniform way to all members of the two main verb classes, Full Verb and Auxiliary Verb. On the one hand, as Radford recognises, there are full verbs which are non-thematic - e.g. *seem* - and on the other there are auxiliary verbs that are thematic. Perhaps the clearest example of a thematic auxiliary is the possessive auxiliary *have*, which for most British speakers can be used as in the following examples:

(34) a Have you a moment?

b I haven't time to do that.

The inversion and negation show that *have* is an auxiliary, but it means `possess' and assigns a thematic role to both its subject and its complement. Some modal verbs also qualify as thematic auxiliaries. The clearest example is *dare*, but we might even make the same claim regarding *can* if we can trust the evidence of examples like the following:

(35) a Pat can swim a mile.

b Pat's ability to swim a mile is impressive.

The words *Pat's ability to swim a mile* are a close paraphrase of example (a), so they should receive a similar semantic (or thematic) structure; but if the ability is attributed directly to Pat, as in (b), there must be a thematic link not only in (b) but also in (a). If this is true, *can* in (a) must be thematic, because it assigns a role to its subject as well as to its complement. In short, although most auxiliary verbs qualify as FWs, there are some that do not. This is not what we expect if Auxiliary Verb, as a whole, is a FWC.

In conclusion, we cannot define FWC in terms of FW because the latter does not have suitable properties. As Cann says, FW is a `cluster concept' which brings together a range of characteristics that correlate more or less strongly with

one another, but which does not map cleanly onto word classes. The boundary of FW runs through the middle of some word classes, and the criteria that define FW are themselves split when applied to word classes. There is no doubt that a grammar should accommodate FW in some way, but not by postulating FWC.

**7. FWCs as closed classes**

Another definition which has been offered for Functional Category refers to the distinction between open and closed classes, which again is part of a fairly long tradition in descriptive linguistics (Quirk et al 1985:71, Huddleston 1984:120). For example, Haegeman (1994:115-6) invokes the contrast between closed and open classes when she first introduces the notion Functional Projection (her equivalent of Functional Category). It is also one of the criteria in Abney (1987:64).

This distinction is different from the function/content distinction because it applies to classes rather than to their members. A class is open if it can accept new members, and closed if it cannot, regardless of what those members are like. This looks promising as a basis for the definiton of FWC - at least it should satisfy our third criterion of having whole classes rather than individual words as members. However, this definition fares badly in relation to the other two criteria.

Once again one problem is that the closed/open distinction is a matter of degree, whereas categories must be either functional or substantive; in short, this criterion fails on the second test, clarity of the boundary. The distinction really belongs to historical linguistics because the addition of new vocabulary changes the language and involves two diachronic processes: creative word formation and borrowing. Borrowing is the most relevant of these because it is the one most usually mentioned in discussions of the closed/open distinction.

The question, then, is whether there is, in fact, a clear distinction between word classes which do accept loans (or calques) from other languages, and those which do not. Among historical linguists the answer is uncontroversial (e.g. Bynon 1977:255, Hudson 1996:58-9). There is no such distinction, only a gradient from the most `open' class, Noun, to the most closed ones (such as Coordinating Conjunction). Even the most closed classes do accept some new members. For example, in English the list of personal pronouns has seen some changes through time, with the recent addition of *one*, `people' and the much older addition of *they*, *them* and *their*; and even Coordinating Conjunction has a penumbra of semi-members (*yet*, *so*, *nor* - see Quirk et al 1985:920) which may presage a future change of membership.

Another way to approach the closed/open distinction would be to consider the actual size of the membership, giving a distinction between `large' classes and `small' ones. However, this is obviously likely to be a matter of degree as well, and it is precisely in the classes of intermediate size that uncertainty arises. Preposition is a clear example, with about seventy clear single-word members (Quirk et al 1985: 665), several of which are loans (*via, per, qua, circa, versus, vis-a-vis, save*). Chomsky (1995:6) appears to classify Preposition as a functional category[20], but Radford does not (1997:45). Quirk et al (1985:67) classify Preposition as a closed class, in spite of the evidence in their own list, and Haegeman (1994:115) recognises that it is a `relatively closed class', but nevertheless classifies it as a substantive category. As we saw in the last section, Preposition is also a troublesome borderline case for the definition of FWC in terms of FW.

The closed-class definition of FWC also fails on the first test by not putting the boundary in the right place. The problem is that it is easy to find examples of closed classes which are not FWC by any other criteria - and in particular, not FWs. Cann lists a number of examples such as points of the compass and days of the week. These have some idiosyncratic syntactic characteristics, but in most respects they are ordinary common or proper nouns (*to the north*, *on Wednesday*). The membership of these classes is rigidly closed, so should we conclude that they are FWCs in spite of being nouns?

This discussion has suggested that FWC is not the natural extension of Closed Class that it may seem to be. Classes are more or less closed, but categories are not more or less functional, and a closed class may be a subset of an open one.

**8. Grammar without FWCs**

If the previous arguments are right, the notion FWC has never been defined coherently, so we cannot be sure which categories are functional and which are not. Moreover, we have found that two of the clearest examples, Complementizer and Determiner, are not even word classes, let alone functional word classes. It therefore seems fair to conclude (with appropriate reservations about sub-word and position categories) that there may not in fact be any functional categories at all.

However, it would be wrong to end on such a negative note, because the discussion also has a positive outcome: the validity of the notion Function Word as a cluster concept defined by a combination of characteristics. Even if FW does not justify FWC, it deserves some place in a grammar, but what place should it have? The basis for FW is that words which have very little meaning tend also to have very little form and very little syntactic freedom. One possibility is that this is a fact which is available only to academic observers of language, comparable with the facts of history and world-wide variation; but this seems unlikely as the raw data are freely available to every speaker, and the correlations are both obvious and natural - indeed, iconic. It seems much more likely that FW is part of every speaker's competence; but cluster concepts are a challenge for currently available theories of grammar,[21] especially when some of the concepts are quantitative (amount of meaning, amount of form, amount of freedom).

**References**


Abney, Stephen. (1987). "The English Noun Phrase in its Sentential Aspects." MIT dissertation.

Albrow, Kenneth. (1972). "The English Writing System: Notes towards a description." Longman, London.

Atkinson, Martin. (1994). "Parameters." *In* "Encyclopedia of Language and Linguistics." (R. Asher, ed.), pp. 2941-2. Pergamon Press, Oxford.

Blake, Barry. (1990). "Relational Grammar." Routledge, London.

Borsley, Robert (ed.). (Forthcoming). "The Nature and Function of Syntactic Categories." Academic Press, New York.

Borsley, Robert and Kornfilt, Jaklin. (This volume). "Mixed extended projections." *In* Borsley (forthcoming).

Bresnan, Joan. (This volume). "Mixed categories and head-sharing constructions." *In* Borsley (forthcoming).

Bynon, Theadora. (1977) "Historical Linguistics." Cambridge University Press: Cambridge.

Cann, Ronald. (This volume). "Functional versus Lexical: a cognitive dichotomy." *In* Borsley (forthcoming).

Carney, Edward. (1997). "English Spelling." Routledge, London.

Crystal, David. (1995). "The Cambridge Encyclopedia of the English Language." Cambridge University Press, Cambridge.

Chomsky, Noam. (1995). "The Minimalist Program." MIT Press, Cambridge, MA.

Creider, Chet and Hudson, Richard. (Forthcoming). "Inflectional morphology in Word Grammar." *Lingua.*

Grimshaw, Jane. (1991). "Extended projection." Unpublished manuscript.

Haegeman, Liliane. (1994). "Introduction to Government and Binding Theory. 2nd edition." Blackwell, Oxford.

Huddleston, Rodney. (1984). "An Introduction to the Grammar of English." Cambridge University Press, Cambridge.

Hudson, Richard. (1984). "Word Grammar." Blackwell, Oxford.

Hudson, Richard. (1990). "English Word Grammar." Blackwell, Oxford.

Hudson, Richard. (1995). "Competence without Comp?" *In* "The Verb in Contemporary English" (B. Aarts and C. Meyer, eds.), pp. 40-53. Cambridge University Press, Cambridge.

Hudson, Richard. (1996). "Sociolinguistics." Cambridge University Press, Cambridge.

Hudson, Richard. (1997). "The rise of auxiliary DO: verb-non-raising or category-strengthening?" *Transactions of the Philological Society* 95:1, 41-72.

Hudson, Richard. (1998). "An Encyclopedia of Word Grammar." Accessible via http://www.phon.ucl.ac.uk/home/dick/wg.htm.

Larson, Richard. (1985) "On the syntax of disjunction scope." *Natural Language and Linguistic Theory* 3, 217-64.

Netter, Klaus. (1994). "Towards a theory of functional heads: German nominal phrases." *In* "German Grammar in HPSG. CSLI Lecture Notes 46." (John Nerbonne, Klaus Netter and Carl Pollard eds.). CSLI, Stanford.

Ouhalla, Jamal. (1991). "Functional Categories and Parametric Variation." Routledge, London.

Pollard, Carl and Sag, Ivan. (1994). "Head-driven Phrase Structure Grammar." University of Chicago Press, Chicago.
Pollock, Jean-Yves. (1989). "Verb-movement, universal grammar and the structure of IP." *Linguistic Inquiry* 20, 365-424.
Postal, Paul. (1966). "On so-called `Pronouns' in English." *Georgetown Monographs on Languages and Linguistics*, 177-206.
Pustejovsky, James. (1991). "The generative lexicon." *Computational Linguistics* 17, 409-41.
Quirk, Randolph; Greenbaum, Sidney; Leech, Geoffrey and Svartvik, Jan. (1985). "A Comprehensive Grammar of the English Language." Longman, London.
Radford, Andrew. (1990). "Syntactic Theory and the Acquisition of English Syntax." Blackwell, Oxford.
Radford, Andrew. (1997). "Syntactic Theory and the Structure of English. A Minimalist Approach." Cambridge University Press, Cambridge.
Rosta, Andrew. (1997). "English Syntax and Word Grammar Theory." London PhD.
Smith, Neil and Tsimpli, Ianthi-Maria. (1995). "The Mind of a Savant. Language learning and modularity." Blackwell, Oxford.
Trudgill, Peter. (1990). "The Dialects of England." Blackwell, Oxford.
Zwicky, Arnold. (1994). "Syntax and phonology." *In* "Encyclopedia of Language and Linguistics" (R. Asher, ed.), pp. 4476-81. Pergamon Press, Oxford.

**Endnotes**

1. This paper has changed considerably since my presentation at the Bangor conference on syntactic categories in 1996. It has benefitted greatly from discussion at that conference and at a seminar at UCL, as well as from the individual comments of And Rosta, Annabel Cormack and, in particular, Bob Borsley. It also takes account of what two anonymous referees said about an earlier version. I am grateful to all these colleagues who have helped me along the way, and especially to Ronnie Cann for generously showing me several versions of his paper.

2. Grammatical terms may be used either as common nouns (e.g. *a noun*; *two nouns*) or as proper nouns (e.g. *(The class) Noun is larger than (the class) Adjective*). I shall capitalise them when used as proper nouns.

3. To give a flavour of the debate, consider the argument that the position C allows a simple explanation for word-order facts in Germanic languages. In V2 order, the verb is said to be in C, so if C is filled by an overt complementizer, the verb cannot move to C - hence clause-final verbs in subordinate clauses. Unfortunately for this explanation, it is not only complementizers that trigger clause-final verb position - the same is true of all traditional `subordinating conjunctions', relative pronouns, interrogative pronouns and so on, none of which are assumed to be in C position. If they are (say) in Spec of C, why can't the verb move to C, as in a main clause?

4. The treatment of `zero' is irrelevant to our present concerns because any solution will pair zero with just one complementizer, *that*. My own preferred solution was suggested by Rosta (1997), and involves a special relationship `proxy'. Verbs that allow either *that* or zero select a proxy of a tensed verb, which is either the tensed verb itself, or the instance of *that* on which it depends. In this way we avoid

positing a `zero complementizer' while also avoiding the repeated disjunction `a tensed clause introduced by *that* or by nothing'.

As Rosta points out, another advantage of this analysis is that it allows us to refer directly to the finiteness of the lower verb. One part of finiteness is the contrast between indicative and `subjunctive', as in (1).

(1) I recommend that Pat be the referee.

A verb such as *recommend* may select the proxy of a subjunctive verb as its complement, which means *that* followed by a subjunctive verb. If it turns out that subjunctive verbs almost always occur with *that*, this fact can be built into the definition of `proxy of subjunctive verb', just as the optionality of *that* is built into that of `proxy of indicative verb'.

5. A referee comments that the same is true of Preposition: no verb selects generally for PP, but many verbs select either individual prepositions (e.g. *depend* selects *on*) or some meaning which may, inter alia, be expressed by a PP (e.g. *put* selects a place expression such as *on the table* or *here*). This is true, but all it shows is that Preposition is not relevant to subcategorization. It is not a reductio ad absurdum of Principle 1, because Preposition can be justified in other ways (e.g. in terms of preposition stranding and pied-piping).

6. See footnote 4 for my preferred treatment of subjunctive selection.

7. The analysis of possessive *'s* is controversial. I shall simply assume that it is a possessive pronoun; for evidence see Hudson (1990: 277).

8. According to the analysis for which I shall argue in section 5, determiners are pronouns which take common nouns as their complements. Wh-pronouns also take complements, though their complements are the finite verb in the clause that they introduce (Hudson

1990:362). Given these two analyses, it follows that one pronoun may even have two complements, a common noun and a finite verb, as in *which students came?*.

9. The word *enough* is a poor example of a determiner, as it is more like a quantity expression such as *much* and *many* - indeed, Rosta has suggested (1997) that the surface word *enough* corresponds to a pair of syntactic words, *much/many enough*.

10. Rule 1 ignores examples like the one Radford cites, our (8a):

(1) Pat is head of the department.

This shows that some singular countable common nouns can sometimes be used without a determiner, but this possibility depends both on the noun itself and on the containing phrase. It is possible for names of professions (as in French), but not generally; and it is possible after the verb *be* or *become*, but not generally:

(2) a Pat is head/*bore/*admirer of the department.

b Pat is/became/*introduced/*looked for head of the department.

If anything, the pattern confirms Rule 1, because the exceptions also involve a complement noun selecting the word on which it depends:

Rule 1 (exception)

A profession noun may be the complement of the verb *be* or *become*.

Similar remarks apply to other well-known examples like the following:

(3) a We were at school/college/*cinema.

b He was respected both as scholar and as administrator.

11. This analysis reverses the usual relationship between complements and heads. In general, heads select complements, but in this analysis it is the complement (the common noun) that selects the head (the determiner). This relationship is not without precedent, however. In Romance languages, the perfect auxiliary is

selected by its complement (e.g. unaccusative verbs such as `go', select `be' while other verbs select `have'). In English, the adjective *same* selects the determiner *the* (*the/*a/*my same person*), and *own* selects a possessive determiner (*my/*the/*an own house*).

12. The possibilities are different if the complement contains a superlative adjective:

(1) a That seems the best solution.

b That seems my best option.

Thanks to Annabel Cormack for this detail. And Rosta also points out the possibility of *no* before certain adjectives:

(2) a She seems no mean linguist.

b That seems no small achievement.

13. It could be argued that valency should also cover subjects/specifiers, but this is a separate issue.

14. It makes no difference to this argument whether valency patterns are stipulated or derived by general linking rules from argument structure. In either case, the links between the words and the word class have to be stipulated.

15. It makes no difference how the class-membership is expressed. I use Word Grammar terminology here, but the same objection would apply to a feature notation because the relevant feature [+W] would be stipulated lexically.

16. If the equivalent of **the my house* is permitted in some other language, this must be because the valency of the article allows a disjunction: either a common noun or a possessive pronoun. Similar minor variations are found in English - for example, universal pronouns (*all*, *both*) allow a pronoun or a common noun (*all (the) men*, *both (my) friends*). This is to be expected if the complement is selected

lexically by the determiner, as claimed here.

17. This little grammar speaks for itself, but illustrates some important principles of Word Grammar, such as the independence of rules about `possibility' and `identity', and generalisation by default inheritance. For more details see Hudson 1990, 1998.

18. The only non-function words whose only vowel is /ə/ are the interesting pair *Ms* and *Saint* (pointed out by Piers Messun); and those which have fewer than three letters are *go*, *do* and *ox*. The observation about spelling is in Albrow 1972 and Carney 1997.

19. For example, we all write *Aspects of the Theory of Syntax*, with the FWs *of* and *the* treated differently from the others. Even quite long FWs such as *since* and *with* may be treated in this way, but some short FWs tend to be capitalized, as in the following examples:

(1) a But Some of Us Are Brave.

b What Do We Mean by Relationships?

c Up to You, Porky.

d Cosmetics: What the Ads Don't Tell You.

Usage is remarkably consistent across authors, but not completely consistent; for example I found *its* treated in both ways:

(2) a Huddersfield and its Manufacturers: Official Handbook.

b The English Noun Phrase in Its Sentential Aspect.

20. More accurately, Chomsky lists just four substantive categories, including `particle', but gives an incomplete list of functional categories which does not include Preposition. His view of Preposition therefore depends on whether or not he intends it to be subsumed under Particle. To add to the uncertainty, in another

place (1995:34) he includes Pre- or Post-position among the categories which are defined by the features [N, V], which presumably means that it is a substantive category, but Particle is not. He does not mention Adverb in either place.

21. I believe that Word Grammar offers as good a basis as any current theory for the treatment of cluster concepts. The `isa' relationship allows the members of FW to be either whole word classes (e.g. Auxiliary Verb) or individual words (e.g. *that*). Default inheritance allows individual FWs to lack any of the default characteristics, as required by any cluster concept. Functional definitions allow FW to have the same range of attributes as any specific word - a default meaning schema, a default valency pattern, even a default phonological and spelling schema. Thus the definition of FW might look something like this (where `whole' is the full inflected form - Creider and Hudson forthcoming):

(1) a FW has a complement.

b FW's sense and referent are the same as those of its complement.

c FW's whole is one syllable.

d FW's vowel may be /ə/.

e FW's written whole may be one letter.

This set of characteristics defines the typical FW, such as *a*, *do* or *of*. Less typical FWs override one or more characteristics, so the more characteristics are overridden, the less typical they are. This captures one aspect of the quantitative variability of FWs. The other aspect involves the amount of variation on each of the individual characteristics; for example, a word that contains two syllables is clearly less exceptional than one with three, and a word whose sense supplies only one feature (e.g. `countable') is less exceptional than one that supplies two (e.g. `male', `singular').