Gerunds and multiple default inheritance*

RICHARD HUDSON

Abstract

In general, English gerunds such as (We were talking about) John having a sabbatical combine the internal characteristics of a clause with the external characteristics of a noun phrase. Previous analyses have tried to recognise the mixed character of gerunds by assigning them two separate nodes, one verbal and the other nominal, but all such analyses are problematic. This paper proposes an analysis similar to that of Malouf (1998) in which the verbal and nominal classifications are combined on a single node, and argues that the node can in fact inherit, by multiple default inheritance, from both the supercategories without any conflict. This is possible because 'noun' and 'verb' (as opposed to their respective sub-classes) place orthogonal restrictions on the phrase that they head: 'noun' restricts its external distribution, while 'verb' restricts its internal structure. (The analysis is expressed in terms of Word Grammar, so these facts are actually expressed in terms of dependencies: 'noun' and 'verb' appear only in restrictions on dependents and parents respectively.) The analysis is extended to explain some details of gerunds which confuse this very simple picture. On the one hand, a limited range of determiners is possible: a 'possessive' subject (e.g. about John's having a sabbatical) or no in prohibitions or existentials (e.g. No playing loud music! There's no mistaking that voice). And on the other hand, a very few constructions demand a gerund rather than a noun phrase (e.g. It's no use..., They prevented us from...).

1 Introduction and overview

One of the most troublesome areas of English grammar is illustrated in (1), which contains a gerund¹, the word *having*:

^{*}The ideas in this paper were first presented in a paper to the Linguistics Association of Great Britain conference in April 1999. It incorporates a number of suggestions made by participants, and has benefitted greatly from comments and bibliographical suggestions by David Denison. I am grateful to Chet Creider for drawing my attention to Malouf's excellent dissertation, and for sharing in the development of the ideas, which he has applied to similar data from Swahili in Creider (1999).

¹Terminology varies from author to author. What I am calling simply 'gerunds' are of ten called 'verbal gerunds', in contrast with 'nominal gerunds' which I shall call nominalizations. The term 'gerund' is used quite differently in romance linguistics, where it refers to verb forms which I would call 'present participles'. The term derives from Latin, where the form *gerundum* was in fact the gerund of the verb *gerere*, 'to do', so my usage is in line with that of traditional Latin grammars

(1) We were talking about John having a sabbatical.

The trouble with words like having in this example is that they are half-verb and halfnoun, which makes them a serious challenge for any theory of grammatical structure. The facts are well known and uncontentious, but there is a great deal of disagreement about precisely, or even approximately, how to accommodate gerunds. The history of modern linguistics is littered with attempts to do this: Rosenbaum (1967), Chomsky (1970), Jackendoff (1977), Baker (1985), Dienhart and Jakobsen (1985), Abney (1987), Johnson (1988), Milsark (1988), Pullum (1991), Lapointe (1993), Wescoat (1994), Yoon (1996), Bresnan (1997), Kaiser (1998), and especially Malouf (1998), which includes an excellent review. Meanwhile, and more or less independently of this debate about gerunds in present-day English, there has been a great deal of discussion of how they developed since Old English (which had no gerunds) - Tajima (1985), Donner (1986), Jack (1988), Houston (1989), Fanego (1996), Wurff (1993, 1997); Denison (1993:403-4) is a convenient summary. We shall see below that the historical development is important in evaluating any theory of modern gerunds because the same theory must also be able to accommodate the range of intermediate forms that are found in earlier stages of English.

We can easily summarise the main facts, as illustrated by *having* in the above example. It must be a verb, in fact an example of the ordinary verb HAVE, because it has a bare subject and a bare direct object and it can be modified by *not* or an adverb:

- (2) a. We were talking about John not having a sabbatical.
 - b. We were talking about John soon having a sabbatical.

These are characteristics which not only distinguish verbs from nouns but also distinguish them, at least in combination, from other word classes. On the other hand, it must also be a noun because the phrase that it heads is used as the object of a preposition (*about*), and could be used in any other position where plain noun-phrases are possible:

- (3) a. John having a sabbatical upset Bill.
 - b. Did John having a sabbatical upset Bill?
 - c. They discussed John having a sabbatical.
 - d. John not having a sabbatical and Mary's failure to get study-leave meant that we weren't short-staffed after all.

⁽Griffin 1991:82).

The word *having* must be a noun if these positions are indeed reserved for nounphrases and if noun-phrases must be headed by nouns - i.e. if we maintain the principle of endocentricity. As we shall see below, one way to reconcile these conflicting facts is to weaken this principle, but this is a high price to pay for a solution.

In addition to these main facts, however, there are three others which complicate the picture. The first is the well-known fact that the gerund's subject may be a 'possessive':

(4) We were talking about John's / his having a sabbatical.

What is not always recognised is that this pattern is not a straight-forward alternative to the bare subject. According to Quirk et al (1985: 1064, 1194) the possessive is more formal, so *his* and *your* in the following examples are described as 'awkward or stilted' in comparison with *him* and *you*:

- (5) a. I dislike him / his driving my car.
 - b. We look forward to you / your becoming our neighbour.

On the other hand, in American English possessives are (apparently) much more normal, and bare subjects may even be rejected (suggesting a somewhat more archaic grammar, as we shall see below). If this is true, it may explain why discussions of gerunds by American linguists have tended to take the possessive subject as the normal pattern (as witness the name 'POSS-ING' which was widely used for gerund clauses in the 1970s).

The second fact has been much less widely acknowledged, but it deserves to be taken seriously. Even in present-day English we find some patterns in which a gerund is used with an ordinary determiner, especially *no* or *any* (Quirk et al 1985:1066). This happens in two constructions. One construction consists of *no* and a gerund clause used as a main-clause prohibition:

- (6) a. No playing loud music!
 - b. No eating sweets in lectures!

The other construction is a clause whose subject is *there*, whose verb is a form of BE, and whose delayed subject is *no* or *any* followed by a gerund clause²:

 $^{^{2}}$ Examples (f) and (g) were provided by David Denison and the Collins Cobuild English Dictionary; the remaining examples are from Quirk et al (1985:1066).

- (7) a. There's no mistaking that voice.
 - b. There was no lighting fireworks that day.
 - c. There isn't any telling what they will do.
 - d. There must be no standing beyond the yellow line.
 - e. There was no turning the other cheek.
 - f. There's no pleasing some people.
 - g. There's no denying it.

It is true that these constructions are restricted in terms of what is possible outside the gerund clause; for example, in both patterns the negative is mandatory. However there is also no denying that they are fully productive as far as the gerund clause is concerned, so they cannot simply be listed as archaic relics of an earlier stage of the language (comparable with *come what may* or *if you please*). They have the classic characteristics of idiosyncratic but productive constructions - non-canonical syntax and semantics combined with productivity (Fillmore et al 1988, Goldberg 1995, Kay and Fillmore 1999). A complete account of present-day gerunds cannot ignore them.

A third detail which should be born in mind is the existence of constructions in which only a gerund phrase, and no other kind of noun phrase, may be used (Malouf 1998:34, quoting Quirk et al 1985:1231). On the one hand we have constructions where the gerund phrase is extraposed (examples from Quirk et al):

- (8) a. It's / There's no use telling him anything.
 - b. There's no point telling him anything.
 - c. It's scarcely worth(while) you / your going home.
 - d. It's pointless buying so much food.

In none of these examples is it possible to replace the gerund phrase by an ordinary noun phrase:

- (9) a. *It's no use a big fuss.
 - b. *There's no point anything else.
 - c. *It's scarcely worthwhile a lot of work.
 - d. *It's pointless purchase of food.

On the other we have at least one verb, PREVENT, which only allows a gerund phrase after its complement preposition.

(10) They prevented us from finishing it / *its completion.

In short, these are all cases where some construction selects specifically for gerund phrases, so it is important that these should be distinguishable from other noun phrases.

These facts about possessive subjects, *no/any* and gerund selection are important because they confuse the simple view of the relationship between the nominal and verbal characteristics of gerunds. If we think of a gerund in terms of the phrase that it heads, the following generalisation is almost true:

- (11) The phrase headed by a gerund is:
 - a. an ordinary clause as far as its internal structure is concerned, but
 - b. an ordinary noun-phrase (or DP) in terms of its external distribution.

Thus the gerund's nominal properties are all properties that it contracts as a dependent while its verbal ones are those that it has *qua* head. This description comes very close to being true, but it is falsified by examples like *his driving my car* and *no mistaking that voice*, both of which look as though they start with a determiner - part of the internal structure of noun phrases, not clauses. Similarly, the description has trouble with constructions like *prevent from*, which show that the external distribution of a gerund phrase is not totally identical with that of ordinary noun phrases.

On the other hand it would be wrong to take these exceptions too seriously. After all, it is almost true that gerund phrases are verbal inside but nominal outside, so we must not abandon this generalisation just because of the exceptions just noted. What is needed, therefore, is an analysis which solves two problems.

(12) **Problem A**: How to reconcile the nominal and verbal features found in straightforward examples, where verbal features control internal structure and nominal features control external distribution.

Problem A will turn out not to be a problem at all, thanks to the way that English and many other languages are organised. I shall argue for the simplest possible analysis, in which gerunds themselves are indeed both verbs and nouns; and I shall show that the characteristics of verbs and nouns never conflict because nominal features always control external distribution, but verbal features never do, whereas the reverse is true of internal structure. We shall also see that it is crucial to assign gerunds to specific sub-classes of both noun and verb in order to get the desired results; but given the right classification, nothing more needs to be said about straightforward gerunds.

(13) **Problem B**: How to reconcile

a. the fact that possessive subjects and *no/any* are determiners with the fact that they can introduce a gerund phrase, and

b. the fact that *prevent from* does not allow noun phrases with the fact that it does allow gerund phrases.

This is the problem of how to accommodate exceptional cases, and by definition exceptions must be stipulated so we must look for a solution which stipulates these; but the simpler the stipulations are, the better.

The inherent difficulties of gerunds are hard to separate from the limitations of the theories assumed, so the analysis proposed below will rest firmly on a number of general theoretical assumptions which I shall explain as they become relevant. In brief, the analysis will involve a combination of multiple default inheritance and dependency structure, so to the extent that it works it will count as evidence in favour of these theoretical assumptions and the more general theory which combines them, Word Grammar. However I should like to stress that the analysis is very similar to one which has been developed quite independently by Malouf (1998) in terms of a different theoretical framework (HPSG). What both theories share is multiple default inheritance, which plays a major part in both analyses, but I shall argue that the use of dependency (instead of phrase structure) leads to a simpler analysis.

2 Previous analyses: two nodes or one?

Most³ previous analyses have taken it for granted that no node in a sentence structure can be classified as both a noun and a verb - an assumption encouraged by the widely accepted analysis of word classes in terms of the features N and V. Since nouns and verbs carry opposite values for both these features it is logically impossible for 'verb' and 'noun' to combine; and the combination [+N, +V] is normally assumed to define the class of adjectives. As Malouf points out (1998:90), this is contrary to the Western grammatical tradition which has always recognised 'mixed' categories such as participles (so called because they 'participate' in the characteristics of both verbs and adjectives), and the analysis which I shall offer below is very much more traditional in this respect than any other recent one except Malouf's.

If one node cannot carry two conflicting classifications, the obvious solution is to assume two separate nodes, one for the nominal classification and the other for the verbal one. Moreover, the natural way to show that the nominal classification controls external distribution while the verbal classification controls internal structure is to make the verbal node subordinate to the nominal node: a verb phrase inside a noun phrase. This has the further attraction of providing a position for a possessive subject,

³I know only two analyses in which English gerunds are [+N, +V]: Malouf (1998) and van der Wurff (1993), which follows van Haaften et al (1985).

in the 'determiner' position within the higher noun phrase. Malouf surveys the various analyses that have been offered within the generative tradition, but observes (ibid: 87) that they all assign gerund phrases some variation on the structure shown in **Figure 1**, where VP is contained within NP.



Figure 1

Since we shall be considering an analysis below in which word-word dependencies replace phrase structure, it is worth noting here that the same logic has applied within the dependency tradition. For example, an early attempt to replace phrase structure entirely by dependency failed in part because of the problem of separating the nominal and verbal characteristics of gerunds (Hudson 1976:37-43, 204). More recently, a dependency analysis has been proposed in which the suffix *-ing* is treated as a separate word, a clitic whose host is the verb; this analysis allows the nominal classification to be assigned to the clitic (Hudson 1990:316-26). What these two analyses share with the phrase-structure analyses is that they offer a structural solution to what at least appears to be a problem of classification, not structure. Apart from the classification conflict there is no independent evidence that (say) *having* is two words. Similarly, there is very little evidence that a gerund phrase is a noun phrase as well as a verb phrase beyond the need for double classification, though in this case there is at least the evidence of the possessive subject.

For a detailed evaluation of earlier analyses I refer to the very thorough survey in Malouf (ibid:53-87). However it will be helpful to survey these analyses in terms of his helpful classification. He recognises three kinds of analysis:

- a. exocentric analyses where the NP has no head, including Chomsky (1970) and Jackendoff (1977);
- b. endocentric analyses in which the head is *-ing* (as in the dependency analysis mentioned above) or a zero suffix: Baker (1985), Abney (1987), Yoon (1996);
- c. endocentric analyses in which the head is the gerund phrase itself (excluding the possessive subject): Pullum (1991), Lapointe (1993), Wescoat (1994), Bresnan (1997), Kaiser (1998).

Each of the endocentric analyses exploits a different formal device for reconciling the verbal and nominal classifications:

- a. a transformation that lowers the nominal -ing onto the verb (Baker);
- b. an abstract category which is classified either as D or N and which selects either IP or VP, plus a rule which affixes this null suffix to a verb that already has the *-ing* suffix (Abney, Yoon);
- c. a weakened Head Feature Convention which allows the mother phrase and its head to have different values for N and V (Pullum);
- d. dual lexical categories <X|Y>, where X and Y determine external and internal properties respectively (Lapointe);
- e. one word projecting (as head) to two different phrasal nodes to an NP node and a VP node within the NP with the higher node unordered with respect to the lower one (Wescoat);
- f. as in Wescoat's analysis, but with a one-many mapping between c-structure and f-structure which allows a single c-structure N (the gerund) to map to an N and a V position in f-structure (Bresnan);
- g. lexical rules that convert a verb into a noun (Kaiser).

The survey is interesting as evidence not only for the ingenuity of linguists but also for the weakness of current theories. However Malouf also finds more or less serious empirical problems in all the proposed analyses.

What conclusion can we draw from this negative verdict? One possibility is that these analyses are all basically on the right lines but the crucial idea is still waiting to be found, maybe just round the next corner. Another possibility, however, is that this approach is a dead end. Maybe it is wrong to be looking for a structural solution; maybe we should be looking instead at the logic of classification, to see if a simpler solution can be found which recognises only as much structure as is visible on the surface. This is the approach which we shall now pursue.

3 Multiple default inheritance and orthogonality

The theory of classification on which the following analysis rests is MULTIPLE ORTHOGONALDEFAULTINHERITANCE. In a nutshell, this allows us to classify a dog not only as an instance of a mammal but also as an instance of a pet, which will guarantee that it will inherit all the characteristics of a typical mammal as well as those of a typical pet to the extent that these are (a) orthogonal - i.e. not in conflict with each other and (b) not overridden by more specific facts about dogs. This is a different logic from the one which underlies the feature structures which are used in most linguistic analysis, so the purpose of the following discussion is to clarify the nature of the assumed logic and to pick out the salient differences from feature-based logics.

Default inheritance is the inheritance of features 'by default', i.e. in the absence of any more specific alternative. Default inheritance has received a great deal of formal attention because of the important part that it plays in Artificial Intelligence and cognitive psychology. The classic discussion of its mathematical properties is Touretzky (1986), and it has been studied extensively in experimental cognitive psychology (Barsalou 1992:178-81, Reisberg 1997:267-70). This mechanism also explains the much-discussed phenomenon of 'prototype effects', in which a penguin counts as a worse instance of a bird than a sparrow does because the penguin overrides some of the bird defaults (Rosch 1976). The consensus in cognitive psychology is that default information is stored in connection with higher-level concepts so that it is available for inheritance if need be (e.g. at least in the case of unfamiliar instances), even though it can be overridden by information stored for more specific sub-concepts and defaults may be stored repeatedly for some subconcepts.

Within linguistics, default inheritance has played a surprisingly small part in formal studies. It has of course been taken for granted for centuries as a matter of common sense in traditional grammars, where generalisations have always been allowed to have exceptions - irregular morphology, exceptional uses of particular tenses or cases, and so on. Language is full of general or basic patterns to which there are exceptions and would seem to call out for treatment in terms of default inheritance, but until recently the dominant model in formal and theoretical linguistics has ignored this possibility. Instead, it has invoked special apparatus to derive exceptional patterns from basic ones; thus syntactic transformations handle deviations from the underlying syntactic structures, and morphological and phonological rules manipulate underlying phonological forms. Moreover, these rules may themselves have exceptions, and one recent suggestion (Pinker and Prince 1988, 1994) is that the exceptions are handled by one psychological mechanism (a neural network) and the general patterns by another (rules). These analyses reflect a very different theory of classification from default inheritance, where the exception and the default are 'facts' of exactly the same kind and differ only in terms of whereabouts in the inheritance hierarchy they are stored.

Nowadays, however, default inheritance is established as an important element in various linguistic theories: HPSG (Flickinger 1987, Pollard and Sag 1994:36, Malouf 1998), Network Morphology (Brown et al 1996), Cognitive Grammar (Langacker 1987, 1990), Construction Grammar (Goldberg 1995:5) and Word Grammar (Hudson 1990:30-52, Fraser and Hudson 1992). In all these theories it is possible to use the same mechanism to specify both the general pattern and its exceptions; for example, any of these theories that deal with syntax could formalise the following pair of rules:

- (14) a. A verb follows its subject.
 - b. An 'inverting verb' precedes its subject.

The first states the default pattern, but the second will override it if the verb is an 'inverting verb', because 'inverting verb' is a particular case of 'verb'. This approach removes the need to derive the inverted pattern by rule from the uninverted one. Of course this does not in itself prove that the derivational approach is wrong, but it does mean that any derivational analysis should be evaluated in relation to a default-inheritance analysis.

The proposed analysis of gerunds will use default inheritance in order to accommodate some of their exceptional characteristics, especially the messy facts about determiners. The possibility of doing this is an important benefit of adopting default inheritance. However another relevant difference between default inheritance and more familiar approaches is in the naming of categories. Since Chomsky (1965, 1970) it has become normal to define categories in terms of features, so the name of a category consists of a set of feature-values such as [+N, -V, +plural], which defines the category 'plural noun'; and previous discussions of gerunds have all assumed a feature-style classification. Default-inheritance systems on the other hand tend to use atomic names for categories - 'noun', 'plural' (or 'plural-noun', taken as an atom). This difference follows from the fact that default inheritance is defined over an inheritance hierarchy, a hierarchical arrangement of categories which stipulates their interrelationships. If the hierarchy already says that a plural is a (or more technically, 'ISA') noun, there is no need to build this information into the category name. We shall assume an inheritance hierarchy for English words like the simplified one shown in Figure 2. Although this uses the same tree structure as the phrase structure in Figure 1, its interpretation is quite different: the lines show 'isa' links (i.e. classification) rather than part-whole relationships. The small triangles are used to signal the isa relationship iconically, since the (large) base of the triangle rests on the (large) supercategory while the (small) apex points to the instance (Hudson 1998:34).



Figure 2

We can call the two kinds of category name 'featural' and 'atomic'. The difference between them has some important consequences for general theory as well as for the analysis of gerunds.

Sub-classification. The use of featural names is well suited to cross-classification, but not for sub-classification. This approach may be well suited to phonology, where it started, but it misrepresents most syntactic classification. Apart from inflectional categories such as gender, number and case it is very hard to find genuine cross-classification in syntax, but sub-classification is common-place. Consequently a feature system must always be supplemented by rules⁴ which specify subordination relations among features - for example, a rule that 'plural' only combines with the features for 'noun'. In short, when applied to sub-classification a featural system contains just the same stipulations about isa relationships as does an atomic one, but it also has more complex category names, so it has no formal advantage in terms of elegance or economy.

Markedness. One of the advantages of default inheritance is that it captures formally the traditional notion of 'markedness', in that the default category and its properties are those that are 'unmarked'. In a featural system, on the other hand, markedness is much harder to capture, especially under the normal assumption that every feature is a combination of an attribute and a value⁵. This means that the marked and unmarked members of a pair have just the same status in a feature analysis, whereas a default-inheritance analysis gives them quite different statuses. For example, in the analysis of English nouns it is possible to dispense entirely with the category 'singular', since the singular is simply the unmarked default - unmarked in terms of both morphology and semantics. This means that **Figure 2** is in fact a complete analysis of the number classification of English nouns; so *dog* is a noun, whereas *dogs* is a plural-noun. For verbs the default is probably finite present tense plural (Hudson forthcoming a), so gerunds are marked exceptions which override some of the default properties of verbs. This status is clear in their morphology but it is most dramatic in their classification, as we shall see.

⁴Such rules are explicit in GPSG, as Feature Cooccurrence Restrictions (Gazdar et al 1985:17), but in some theories they are simply taken for granted.

⁵ Admittedly it would be possible to define '-' (or 'u' or '0') as the universal unmarked value, but labels for values are arbitrary so this convention would be a stipulation of an unmarked value for each feature. Another approach is to treat features themselves as atoms which are either present or absent (as in Dependency Phonology, for example - Anderson and Durand 1986); but this is probably just a notational variation on the use of atomic names for which I argue here.

Multiple inheritance. Finally, default-inheritance makes it possible for gerunds to be nouns as well as verbs, in contrast with the feature analyses which rule this out as a logical impossibility on the grounds that [+N, -V] is incompatible with [-N, +V]. If the inheritance system allows multiple inheritance in principle, and if nouns are labelled simply 'noun' and verbs 'verb', then there is no logical objection to a sub-category which belongs to both. Of course it could be objected that the problem lies not in feature analysis as such, but in one particular application of it (the cross-classification by N and V); but cross-classification is the primary motivation for features, and once multiple inheritance is permitted this motivation disappears.

These differences between the two naming systems justify the approach that I shall take below, in which I shall assume that grammatical categories have atomic names (e.g. 'gerund') and that their relationships are stipulated in an inheritance hierarchy which allows default characteristics to be inherited unless overridden. Furthermore, as we have just seen this hierarchy will allow one category to inherit from more than one super-category; in short, I shall assume the logic of multiple default inheritance.

However, we have to recognise the logical problem that is inherent in multiple inheritance: how to deal with inheritance conflicts. The problem does not arise in most examples of multiple inheritance simply because the characteristics inherited from the different super-categories are 'orthogonal' to each other - they classify the world along completely different lines. For example, a dog can inherit comfortably from both 'mammal' and 'pet' because they each specify quite different characteristics one has to do with biology, while the other relates to social roles. However characteristics need not be orthogonal in this sense, and the possibility of conflict is the main topic of Touretzky (1986), who illustrates the problem with the now famous 'Nixon diamond' presented in Figure 3. The American president Richard Nixon was both a Republican and a Quaker, so he should have inherited the characteristics (i.e. the principles) of both; but on at least one issue Republican and Quaker principles conflict: Republicans countenance warfare, whereas Quakers are pacifists. Nixon clearly could not inherit both of these competing principles, so at least one had to lose. In a simple default inheritance hierarchy, a more specific property has priority over a more general one, which guarantees a single winner in every conflict; but in a multiple inheritance hierarchy concepts may be at the same level of generality so neither will emerge as winner. This is the case with the concepts 'Quaker' and 'Republican' because neither is a special case of the other, so the conflict cannot be resolved.



Figure 3

What does this example mean for the theory of default inheritance? One reaction is that the theory should rule out such cases as logically impossible. This approach is most clearly espoused in Evans and Gazdar (1996), the standard definition of the programming language DATR. According to this theory multiple inheritance should never give rise to this kind of conflict. This is ensured in DATR by excluding genuine multiple inheritance from equal super-categories altogether; only one ordinary 'isa' relationship is allowed, so any other relevant super-categories must be stipulated attribute by attribute. This approach has the great advantage for a computer system of a clean logic and a guaranteed single outcome.

Another way to react to the Nixon diamond is to say that it is not a logical problem at all, but a moral one. If we are trying to model human cognition, then we must accept that Nixon's dilemma was actually a very typical and normal one and therefore one that our model must accommodate. As it happens, the example is a poor one because in real life Nixon solved the problem by prioritizing his principles, and this too is a typical human reaction in the face of conflicting beliefs or principles: we stipulate one of the competing characteristics as the winner. This outcome can be envisaged as filling in one of the values 'Yes' or 'No' for Nixon's 'war?' attribute. Since 'Nixon' is a more specific category than either 'Republican' or 'Quaker', this entry automatically overrides both the inheritable ones. Since Nixon did in fact wage war we must assume that this is the strategy that he adopted. The alternative is to leave the conflict unresolved, which leads to paralysis - a total inability to decide between them. Common experience suggests that this is a common situation, but rather surprisingly perhaps there is very clear evidence for the 'blocking' effect of an unresolved conflict in grammar. This is the only plausible explanation for at least one 'morphological gap': the English gap where we should expect *I amn't. The explanation runs as follows: the desired form is a the first-person form am, but it also is a the negative form, *aren't*; neither of these is a the other so the conflict cannot be resolved and we have no usable form (Hudson forthcoming b).

In view of examples like this, then, it seems that we must allow the logical system to fail because of irresoluble inheritance conflict, rather than protect it against failure

by requiring orthogonality in the database. This is why the logic that we shall assume in the rest of this paper is multiple orthogonal default inheritance - it works only when the characteristics concerned are orthogonal, and crashes otherwise. This conclusion is important if we are considering an analysis of gerunds in which they inherit simultaneously from 'verb' and from 'noun' because it seems likely a priori that the characteristics that can be inherited from these two super-categories will conflict, so the analysis ought to fail. We shall explore this possibility and explain why the problem does not in fact arise in sections 5 and 6, but first we must consider what kinds of structural analyses are available for gerunds.

4 Dependency or phrase structure?

Almost all previous discussions of gerunds have assumed that syntactic structures must be phrase structures, based on the part-whole relationship between words or phrases and larger phrases, but this is not the only imaginable possibility, and there are indeed a number of advantages in basing syntactic analysis on dependencies between individual words. For example, this allows a much more natural treatment of lexical selection relationships because the words concerned are related directly, whereas phrase structure analysis means that they must be related only via at least one mother node. The contrast can be seen clearly in **Figure 4**, which shows how the relationship between a verb and its dependent preposition (e.g. between DEPEND and ON) is shown in these two kinds of analysis; in contrast with the single dependency link there are no fewer than three links in the phrase-structure analysis. For discussion and justification of the dependency approach, see Mel'čuk (1997), Weber (1997), Hudson (1990, 1998).



Figure 4

The choice between these theories is important for our treatment of gerunds, but not vital because they both allow the noun and verb classifications to be assigned either

to two nodes or to one node. The two-node analysis is natural in phrase structure, and the one-node one in dependency, but the reverse pairings are also possible. Malouf (1998) proposes a single-node analysis in a phrase-structure theory (HPSG), while (as mentioned earlier) Hudson (1990:316-26) is a dependency analysis which divides the classes between two nodes by treating the suffix *-ing* as a separate word (a 'gerund' noun). The four possibilities are illustrated in **Figure 5**, where (a) the two-node, phrase-structure analysis is based on Bresnan 1997, (b) on Malouf (1998:94), and (c) on Hudson (1990:316), while (d) is the analysis towards which we are moving.



Figure 5

Although both approaches allow both kinds of analysis, the choice between dependency and phrase structure is nevertheless important for the analysis of gerunds because of its implications for endocentricity. In phrase-structure theory, endocentricity is generally recognised as an important principle, but it is possible to infringe it (as witness the earliest analyses of gerunds, which treated them as NPs whose heads were not nouns). In dependency theory, on the contrary, endocentricity is not negotiable. One of the basic assumptions underlying this theory must be that every phrase has a head, because an exocentric construction simply cannot be analysed in terms of dependencies. By definition, dependencies link a phrase's head to its dependents, so if there is no head, there can be no dependents either. This is one area in which dependency theory is more restrictive than 'plain-vanilla' phrase-structure theory, though of course the X-bar version of phrase-structure theory is restrictive in the same way.

I shall adopt the dependency approach in the following analysis of gerunds because of its advantages in other areas of grammar, but I should first explain how it can be applied to noun phrases since this question is so important for gerunds. At first sight a dependency approach may appear to be unsuitable for explaining the similarities

between disparate examples such as *you*, *John*, *boys* and *these boys*, which appear to have very little in common in terms of their internal structures although their external distributions are precisely the same. These phrases appear to have different kinds of heads (a pronoun, a proper noun, and so on), so how can they have the same distribution? However the principle of endocentricity means that the same problem arises for a phrase-structure analysis, so in each phrase there must be a head whose classification projects to the whole phrase, and if the phrases all share the same classification the same must be true of their heads, a fact which can be expressed even more directly in a dependency analysis than in phrase structure.

The real theoretical issue is not the choice between dependency and phrase structure, but whether or not we allow empty heads. Accepting empty heads permits a 'DP' analysis in which each of these phrases has a determiner which is overt only in these books and you (assuming that pronouns and determiners belong to the same category - a widespread assumption with which I agree; Postal 1966, Greenbaum 1996:163) but which is covert in John and books. However, this analysis faces empirical difficulties (Hudson 2000b) and in any case does not seem to be necessary since an easy alternative is available: an analysis in which the pronoun/determiner class is a type of 'noun' (Huddleston 1988:85, Hudson 1990:268, Pollard and Sag 1994:249). For simplicity we can call pronoun/determiners simply pronouns (after all, traditional pronouns far outnumber determiners). Given these assumptions and this terminology, the analyses are easy: you, John and books are respectively a pronoun, a proper (noun) and a common (noun), all of which are nouns, and these books is a pronoun (these) with a common noun as its complement. The analyses are shown in Figure 6. This, then, is the analysis of noun phrases which will be assumed in the rest of the paper.



Figure 6

The claim of this paper, therefore, is that gerunds can be analysed as single words which have a double classification, and which do not project to a higher phrase node. To the extent that the analysis works it will serve as support for the theoretical and descriptive assumptions on which it rests.

5 Gerunds as nouns

The proposed analysis takes gerunds as examples of both nouns and verbs, so the present section will consider the consequences of analysing them as nouns, leaving the verb half of the analysis till the next section. The crucial point for the present section is the subclassification of nouns discussed in the previous section. This will be the basis for explaining why gerund phrases are nominal externally but not internally. This part of the analysis is virtually the same as the one in Malouf (1998:154).

If gerunds are nouns, how do they fit into the three-way contrast among proper, common and pronoun? The obvious answer is that though they are nouns, they do not belong to any of these three sub-classes of noun, so we must add 'gerund' as a fourth sub-class. This gives the hierarchy shown in **Figure 7**, where the dotted line stands for a relationship that will be made more precise in the next section.



Figure 7

This classification immediately explains why a gerund has the external distribution of a noun: because it is a noun. However it also allows gerunds to be used (as dependents) in contexts where other kinds of noun are not allowed. In section 1 we noticed two such contexts. One was where the gerund phrase is extraposed in examples like the following (repeated from (8)), where a gerund is used in a context where other kinds of noun phrase are not allowed.

- (15) a. It's / There's no use *telling him anything* / *a big fuss.
 - b. There's no point *telling him anything* / *anything else.
 - c. It's scarcely worth(while) *you / your going home / **a lot of work.
 - d. It's pointless *buying so much food* / *purchase of food.

The other was after at least one verb, PREVENT, which only allows a gerund phrase after its complement preposition.

(16) They prevented us from finishing it / *its completion.

The possibility of distinguishing gerunds from other kinds of noun allows us to prevent over-generation in these areas by permitting only gerunds in these contexts. (The details of the rules concerned are irrelevant, the main point being that they can apply to 'gerund' rather than more generally to 'noun'.) The analysis seems to give us just the right combination of specificity and generality in defining the contexts in which gerunds may act as dependents.

However, the noun classification also introduces a new problem: if gerunds are nouns, why do gerund phrases not have the internal structure of noun phrases? As we know, the fact is that gerund phrases have the internal structure of clauses, as witness all the evidence for their being verbs: their use with direct objects and predicative complements, with non-possessive subjects, with adverbs rather than by adjectives and with *not*, plus the fact that a gerund may itself be an auxiliary verb. The gerund phrase (italicized) in the following sentence illustrates all these well-known facts:

(17) I object to him not yet being ready.

This gerund phrase clearly has nothing at all in common with ordinary noun phrases such as *the idea of chocolate* or *his irrational anxiety*.

However, this problem disappears as soon as we notice that NOTHING has 'the internal structure of a noun phrase'. The only thing that all noun-headed phrases have in common is that their head is a noun - a generalisation which does apply to those headed by gerunds, including the italicized phrase in (17), whose head is the (gerund) noun *being*. Beyond this, the phrase's structure depends on whether its head is a pronoun (i.e. pronoun/determiner), a common noun or a proper noun. Pronouns generally take complements but not modifiers, common nouns take modifiers but typically not complements, and proper nouns take a very restricted range of modifiers:

(18)	a.	pronoun:	this book, every student
	b.	common:	big <i>book</i> about linguistics that cost ten pounds
	c.	proper:	(poor) John, North London

These differences are the reason for distinguishing pronouns, common nouns and proper nouns, and the result is that the rules which control the dependents found inside noun phrases never refer to the head simply as a noun.

In short, the grammar of nouns (as such) says nothing at all about their dependents, so there are no dependent-facts to be inherited by gerunds. This is why I started this section by saying that the sub-classification of nouns is the key to the analysis, and why more generally I claimed that the double-classification of gerunds "will turn out not to be a problem at all, thanks to the way that English and many other languages are organised." If nouns had all been of one type, all taking the same range of

dependents, these facts would have been stored at the level of 'noun' and would therefore have been inherited by gerunds. Given the logic of multiple default inheritance the result would have been a clash with the structures inherited from 'verb', a clash which could only have been solved by stipulating a winner. As it is, however, the classification of gerunds as nouns is almost entirely 'free' as far as the phrase's internal structure is concerned, because (for most gerunds) there is no need for special rules or apparatus to resolve conflicts between nominal and verbal features. The two exceptions are the very limited possibilities for determiners (possessive subjects and *no/any*), which will be discussed in sections 8 and 9.

The outcome of this section, therefore, is that the classification of gerunds as nouns has important consequences for how they are themselves used as dependents, but none at all for their own dependents - in other words, gerund phrases have the external distribution of noun phrases, but not their internal structure. In the next section we shall see how the converse is true of their classification as verbs.

6 Gerunds as verbs

As nouns, gerunds contrast with common nouns, proper nouns and pronouns, all of which are word-classes - i.e. classes of LEXEMES. The same is not true of their relationship to verbs, where gerunds differ from other verbs in their INFLECTIONS. Any verb which can be non-finite (i.e. any verb other than a modal and a handful of full verbs such as BEWARE) can be a gerund, but gerunds are distinguished by their inflectional suffix *-ing*. In Word Grammar, 'inflection' and 'lexeme' are subcategories of 'word', so an inflected lexeme inherits from both an inflection and a lexeme (Creider and Hudson 1999). We can now complete the diagram given earlier in which the link from 'gerund' to 'verb' was left unspecified, as in **Figure 8**.





At the same time, of course, a gerund is an instance of whatever lexeme provides its stem - *having* is an instance of HAVE, *walking* is an instance of WALK, and so on which means that gerunds are basically verbs being used as nouns, rather than nouns being used as verbs. It is the verb lexeme that determines its meaning and its valency (subcategorisation) as well as its stem; and the fact that the verb lexeme is a verb has implications for the kinds of modifier that are possible - in particular, a verb may be modified by an adverb but not by an adjective, which is why the same is true of gerunds. All the noun classification contributes is the possibility of being used as a dependent where a noun is required. This, then, is the explanation for why gerund phrases have the internal structure of clauses: because they are clauses (i.e. phrases headed by a verb).

This part of the analysis is different from Malouf (1998), where gerunds are not verbs at all but a sub-class of 'relational', a category which includes adjectives as well as verbs. It is not clear why he adopts this analysis, but it is made possible because gerunds are either derived by means of a 'lexical rule' which takes a verb and turns it into a gerund (ibid:90) or are an inflectional class which overrides the 'verb' classification (ibid:163). Either way, the result is that gerunds are not verbs, although they are based on verbs. The analysis proposed here avoids this change of class.

But if gerunds really are verbs, why don't their phrases have the external distribution of a verb phrase? This is similar to the question in the previous section about why gerund phrases do not have the internal structure of noun phrases, and the answer is also similar: because NOTHING has the external distribution of a verb phrase. The fact is that there are no rules (or principles) which permit some position to be occupied by 'a verb phrase'; every rule that allows a verb phrase also requires the head verb to have some particular inflection - tensed, participle, infinitive or whatever. As long as gerunds are distinct from other inflections, they will never be subject to rules which pick out these inflections. However, as we saw in the previous section, they are subject to rules which apply to nouns, because these do not specify inflections.

The conclusion to which the last two sections have led us is that the grammar of gerunds is very simple indeed. They are inflected by the addition of the same *-ing* suffix as present participles, but they are not present participles: they constitute a unique word class, 'gerund'. This word class isa both 'noun' (where it contrasts with 'proper', 'common' and 'pronoun') and 'non-finite' (which is a sub-class of 'verb'). Having said this, all the main facts about gerunds follow automatically, without any stipulations or special provisions at all: as heads, they are ordinary non-finite verbs, but as dependents they are ordinary nouns.

7 The debris of history: possessives and no/any

The simplicity of gerunds in present-day English lies at the end of many centuries of gradual evolution whose beginnings in Old English were entirely different. In Old English there were no gerunds, but there were nominalisations ('verbal nouns') comparable to modern nouns like NOMINALISATION, ARRIVAL and READING, as in (19):

(19) Fast reading of linguistics articles is difficult.

In Old English the regular verbal noun ended in either *-ing* or *-ung*. The following example is from Denison (1993:387).

(20) ac gyrstandæg ic wæs on huntunge but yesterday I was at hunting 'But yesterday I was hunting'

We shall consider the rise of gerunds in section 8, but the aim of the present section is to correct the impression of perfection and simplicity which the previous two sections have left. Gerunds developed out of a purely nominal pattern, and this history is still visible in the peculiarities of modern gerunds which were described in section 1.

The most obviously nominal relic is the possibility of possessive subjects, as in *John's knowing the answer*. As I mentioned in section 1, this strikes British speakers as rather forced and formal, though it seems to be more acceptable to Americans. In Britain the bare 'accusative' subject is more normal, as in (a) below, and the only possibility in (b):

- (21) a. John knowing the answer surprised us.
 - b. Our visit was spoilt by there being no-one at home

This bare subject is the form to be expected given the rules so far, if we assume that non-finite verbs allow the subject to be either overt or a covert PRO (Hudson 1999a). The dependency structure for this example is as shown in **Figure 9**. The choice of non-subject pronoun forms (*him knowing the answer*, not **he knowing the answer*) is as expected, since subject forms are used only with tensed verbs.



Figure 9

Where the gerund's subject is 'possessive' (aka 'genitive', though this is clearly not an inflected case), it is less clear what the structure is. On the one hand, it could be argued that the structure is the same as when a possessive is used as a determiner in a noun phrase - i.e. with the possessive as head (Hudson 1990:321). This has the advantage of revealing the similarity between these gerunds and ordinary noun phrases, and gives structures like that in **Figure 10**, where 's is the head of the whole noun phrase. (The classification of 's is presumably the same as that for ordinary possessive pronouns: 'possessive pronoun'; Hudson 1990:276). If this is the correct structure, then the grammar must accommodate it by allowing possessives (unlike other determiners) to have a gerund as their complement. Bearing in mind that determiners are assumed here to be pronouns which take a complement noun, the rules required for this are as follows (Hudson 1990:268-76):

- (22) a. A pronoun's complement is a common noun.
 - b. A gerund-complement is a complement.
 - c. A pronoun's gerund-complement isa gerund.
 - d. A possessive pronoun may have a gerund-complement.
 - e. A possessive pronoun's predependent is the subject of its gerund-complement.

Rule (a) is the default for all pronouns that take a complement noun (i.e. for all determiners). This rule is separate from the various rules which specify which pronouns do allow a complement - e.g. *we* does (*we linguists*) but *they* does not (**they linguists*). Rule (b) introduces a sub-type of complement; unlike other theories, Word Grammar allows dependents (and other relationships) to be classified in an inheritance hierarchy (Hudson 1990:189-211). Rule (c) therefore overrides the default (a), allowing gerunds as exceptional complements. Rule (d) allows possessive pronouns (exceptionally) to have a gerund as complement, and Rule (e) says that where this option is adopted, the gerund's subject is the predependent of the possessive - i.e. *John* in the case of *John's knowing the answer* or the syntactic word *him* in *his knowing the answer* (Rosta 1997). This subject link creates a complex dependency structure, as shown in the diagram, but such structures are permitted in Word Grammar (Hudson 1994, Rosta 1994) provided they include a tangle-free 'surface structure' which can be diagrammed on top of the words.



Figure 10

On the other hand, Malouf points out (ibid:51, following Abney 1987:245) that ellipsis of the gerund is not possible, unlike a common noun:

(23) a. *John's passing the exam was surprising, and Bill's [] even more so.b. John's success in the exam was surprising, and Bill's [] even more so.

One way to explain this would be to reject the analysis outlined above, and to assume instead that the possessive is merely the gerund's subject, with just the same structural status as the bare subject in **Figure 9**. This would certainly predict that the possessive cannot occur without the gerund, but it would also throw out the baby with the bathwater by losing the comparison with ordinary noun phrases. Moreover we shall see below that the same is true of the other gerund-taking determiner, *no*, where this alternative is not available; and it would leave unexplained why gerunds can be coordinated with ordinary common nouns after a possessive:

(24) John's passing the first exam and failing of the second surprised everybody.

On balance, then, the structure in **Figure 10** seems preferable. It is in any case normal for complements to differ in optionality; for example, the complement of *every* is obligatory whereas that of *each* is optional. This means that there are no serious descriptive or theoretical objections to stipulating that a gerund complement is obligatory. Given the dependent-type 'gerund-complement' the stipulation is easy:

- (25) a. A pronoun's complement is optional.
 - b. *Every*'s complement is obligatory.
 - c. A possessive pronoun's gerund-complement is obligatory.

Once again Rule (a) is the default, and deals only with the complement's optionality, leaving its word class and selection to be handled by other rules. This default is

overridden by the other two rules, one for *every* and the other for gerund complements.

It is easy to see how possessive subjects formed a necessary stage in the development of modern gerunds from ordinary nominalisations, which use possessives as 'subjects'. It is in this sense that I describe possessive subjects as 'the debris of history'. Another item of debris is the determiner *no/any* which we discussed in section 1, using examples that included the following:

- (26) a. No playing loud music!
 - b. There's no mistaking that voice.
 - c. There isn't any telling what they will do.

These can be analysed along the same lines as the gerunds with possessive subjects. The determiner can be treated in the usual way, as the head of its phrase, but its gerund complement is unusual in being obligatory (i.e. not subject to ellipsis). This can be seen from the following examples, where the gerund is contrasted with a common noun:

(27)	a.	A: No noise, please!
		B: What, none at all?
	b.	A: No being noisy, please!
		B: *What, none at all?
(28)	a.	A: There's no possibility of mistaki
		B: No, none at all!
	h	A. There's no mistaking that voice!

B: *No, none at all!

In this construction there does not appear to be any alternative to a stipulation about optionality.

mistaking that voice!

The rules for no/any will include the following, but I shall not try to offer a semantic analysis of the 'prohibition' use of no or a proper analysis of the pattern *There's no* / not any + gerund construction. I assume that each combination of meaning and valency involves a distinct lexeme, but there is no loss of generality because lexical items themselves may be arranged in an inheritance hierarchy. For example, we can treat 'prohibitive *no*' as a lexeme which isa 'ordinary *no*', thereby guaranteeing that any facts shared by the two uses can be stated just once as a fact about 'ordinary *no*'. (It is even possible that *no* isa *any*, but I shall not explore this possibility here.)

- (29) a. *No* is a pronoun.
 - b. *No*_{prohibition} isa no.
 - c. *No_{prohibition}* has a gerund-complement.
 - d. A pronoun's gerund-complement is obligatory.
 - e. No_{prohibition} has no parent.

Rules (a) and (b) define the inheritance hierarchy just described; rule (c) assigns a gerund complement to the prohibitive no, while rule (d) stipulates that this is obligatory. However rule (d) is made to apply to all pronouns, not just prohibitive no, so it replaces the more specific rule (c) in (25) and has no extra cost. Rule (e) says that prohibitive no does not depend on any other word - i.e. it must be the root of a main clause.

Once again it is obvious why these uses of *no/any* with a gerund exist in current English, given the origin of gerunds in ordinary common nouns which are also possible, with similar meanings, after *no/any*:

- (30) a. No noise, please!
 - b. There's no doubt about his intentions.
 - c. There isn't any way of telling his intentions.

But however understandable their origins may be, the fact remains that these patterns, like the possessive subjects, are exceptional and special uses of gerunds which cannot be explained as simply as was possible with ordinary gerunds.

8 The route from Old English to now, and other languages.

It is essential to evaluate any analysis of current English in relation to a much broader context. First, a diachronic question; Does it explain the origins of current English in earlier forms of English? And second a typological question: does it explain the variation in similar patterns found in other languages? These two criteria are defined and discussed at length in Malouf (1998:75, 119-75), and I shall build here on his suggestions.

The diachronic question arises because the development has been very gradual, so that slightly different grammars have had to coexist over long periods. This means that it should be possible to trace a route back from current English to a much older stage (perhaps late Middle English - Denison 1993:404) via a series of grammars with only minimal differences between adjacent stages. Unfortunately the early history of gerunds is very complex, unclear and hotly disputed - not least because the suffixes used for nominalisations (*-ing* and *-ung* in Old English) merged in Middle English

with those of the participle (formerly *-ende*), to give the Modern English situation where the difference between *-ing* and *-in'* is grammatically irrelevant (both are ambiguous between participle and gerund) but socially important (Denison 1993:387, Malouf 1998:116, Labov 1989). An analysis of current English must therefore generalise, with only minor changes, to the intermediate grammars that are known to have existed in the past. The following discussion rests heavily on the data in van der Wurff (1993,1997), and as in his more recent account (1997) I shall show that the changes involved a gradual evolution of fine details rather than a major reorganisation of the grammar; however van der Wurff assumes a structural analysis which is quite different from the one proposed here⁶.

The relatively 'pure' system of current English stands at the end of a long period of gradual evolution (which van der Wurff dates as starting in the 11th century), during which gerunds shed their nominal 'internal' characteristics - i.e. the characteristics expected within a noun phrase. As we have seen, even today they still have two such characteristics - possessive subjects and occurrence after *no/any* - but until as recently as the end of the 19th century they could also occur with *the* and with adjectives. In the following examples from van der Wurff (1993), I have italicized the relevant words:

- (31) a. Between rheumatism and *constant* handling the rod and gun ... (1853)
 - b. *The* managing an argument handsomely being so nice a Point, ... (1711)
 - c. *The* writing the verbs at length on this slate, will be a very useful exercise (1829)
 - d. *the due* placing them adapts the rhyme to it. (1684)

Malouf (1998:75) quotes similar examples:

- (32) a. *the untrewe* forgyng and contryvyng certayne testamentys and last wyll [15th century]
 - b. my *wicked* leaving my father's house [17th century]
 - c. *the* being weighted down by the stale and dismal oppression of the rememberance [19th century]

Denison (1999) quotes other examples which are worth repeating because of their relatively recent dates:

⁶Van der Wurff assumes an abstract phrase-structure analysis similar to the one in Yoon (1996), in which a zero nominalizing node combines with a present participle. This decision produces structures which are admitted to be "rather complicated, with a bottom-up succession of nominalization, verbalization and nominalization" (ibid:187).

- (33) a. *The* copying them has been and still is my occupation; ... and I am trying to get the printing done also while I am finishing the copying. (1873)
 - b. At least I can't fix on any tangible object or aim in life which seems so desirable as *the* having got it finally over and *the* remaining in perpetuo without desire or aim or consciousness whatsoever. (1890)

Conversely, during this period nominalizations have a verbal characteristic, modification by adverbs, which Malouf claims to be impossible (1998:121). Again the examples are from van der Wurff (1993):

- (34) a. The *quickly* doing of it, is the grace. (1610)
 - b. he finds that bearing of it *patiently* is the best way. (1664)
 - c. the shutting of the gates *regularly* at ten o'clock ... (1818)

Indeed, van der Wurff (1997) even gives an example where an adverb is used with a derived nominalization:

(35) but on an examination more *strictly* by the justices of the peace, and at the Lord Mayor's request, it was found there were twenty more. (1722)

The question, then, is what these examples tell us about the grammar.

One important fact is that 'mixed' gerunds of the kinds illustrated here were not at all common. In a collection of 400 clear gerunds or nominalizations from the 18th and 19th centuries that van der Wurff studied (1997), only 8% showed mixed characteristics by the most generous definition of this category. All the rest were either consistently verbal (82%) or consistently nominal (11%). These figures suggest that the mixed patterns may have been archaic and perhaps even impossible for most writers.

Another observation is that only two areas of grammar are involved: the use of *the*, and the choice between adverbs and adjectives. The first is easily accommodated as yet another determiner which allows a gerund complement, in addition to possessives and *no/any*; in other words, the range of determiners which allow such complements has gradually reduced over time. This is hardly surprising given the origins of gerunds.

The change in the use of adverbs and adjectives is harder to explain, but seems to suggest that there was a period when the choice was less rigidly determined than in current English. Example (34) above shows that adverbs could at least sometimes modify ordinary nouns in 18th century English, and according to van der Wurff (1997) adverbs such as *telkens*, 'continually', can modify nominalizations in modern Dutch:

(36)	a.	door het telkens breken van je beloften	
		by the continually breaking of your promises	
		'Because of the continual breaking of your promises'	
	b.	het telkens geven van geld aan hem	
		the continually giving of money to him	

'The continual giving of money to him'

It is worth pointing out that there is at least some flexibility even in current standard English; some adverbs may modify some nouns, and the choice between adverb and adjective is optional in some verb-modifier collocations (Swan 1980/1995:16-9).

- (37) a. The weather *recently* in London has been appalling.
 - b. I held it *tight/tightly*.
 - c. You guessed *wrong/wrongly*.

However the fact remains that the examples quoted earlier, in which adverbs modified nouns and adjectives modified gerunds, would all be rejected in present-day English.

What has changed is clearly that both adverbs and adjectives are more tightly restricted now than they were in earlier periods. On the one hand, adverbs are (in general) not allowed to modify nouns, and on the other adjectives are (in general) only allowed to modify common nouns (and compound pronouns like *someone*). Without more facts it is hard to know exactly what the restrictions in earlier periods were, but one possibility is that adjectives could modify all nouns, including gerunds, while the restriction on adverbs was semantic rather than syntactic (e.g. *quickly* can modify any word which refers to an event that has a speed). Whatever the facts and the correct analysis, it seems clear that the relevant changes in the grammar can be accounted for by changes to the rules for adjectives and adverbs, and without any change to the analysis of gerunds.

The discussion so far has shown that the proposed analysis for gerunds in current English will accommodate earlier varieties of English by means of minor adjustments, so it satisfies the diachronic criterion. The typological criterion requires it to be compatible with the known characteristics of gerund-like patterns in other languages. This is clearly much harder to apply because there is so much more potentially relevant data, so I shall restrict myself to the generalisation called the 'Deverbalisation Hierarchy', which is summarised in Malouf (1998:121-5). This is an implicational hierarchy on which finite verbs form one pole and nominalisations (i.e. common nouns derived from verbs) form the other. The following examples based on Malouf (ibid:121) illustrate the possibilities, but I have added examples (d) and (e) to represent the intermediate grammars for English which we recognised above, with a star to show that it is no longer grammatical though it is clearly possible. However it

is important to recognise that Malouf denies the possibility of examples like (e), as I mentioned earlier; the alleged impossibility of such examples plays an important part in his analysis.

- (38) a. The DA was shocked that Pat illegally destroyed the evidence.
 - b. The DA was shocked by Pat illegally destroying the evidence.
 - c. The DA was shocked by Pat's illegally destroying the evidence.
 - d. *The DA was shocked by Pat's illegal destroying the evidence.
 - e. *The DA was shocked by Pat's illegally destroying of the evidence.
 - f. The DA was shocked by Pat's illegal destroying of the evidence.
 - g. The DA was shocked by Pat's illegal destruction of the evidence.

According to the Deverbalisation Hierarchy, these examples represent all the known possible ways of combining verbal dependents (subject, object, modifying adverb) and nominal dependents (modifying adjective, object with *of*). At one extreme (a), a finite verb allows all verbal dependents and rejects all nominal ones; at the other extreme (f, g) a nominalisation (regardless of its morphological transparency - i.e. *destruction* or *destroying*) allows nominal dependents but rejects all verbal ones; but in between we have the possibility of mixtures. A gerund may be as purely verbal as a finite verb (b), but it may also allow a nominal (possessive) subject (c) and/or a nominal (adjective) modifier; but (d, e) it may combine nominal and verbal dependents: a modifying adjective and a bare direct object, and a modifying adverb and an object with *of*.

The question, then, is precisely what combinations are not compatible with the Deverbalisation Hierarchy. As we have seen, one of Malouf's supposedly impossible combinations did in fact occur in English, so the remaining exclusions may be wrong as well. Two patterns that do not seem to be attested in either Malouf's or van der Wurff's data are the following:

- (39) a. *... Pat destroying of the evidence.
 - b. *... Pat illegal destroying the evidence.

Unfortunately van der Wurff explicitly excludes examples with overt subjects such as these, so we cannot be sure whether they actually occurred; but we might well expect examples like (b), on the basis of examples where an adjective modifies a gerund (31). The only reason why they might not have occurred is that the use of a bare subject seems to be a relatively new possibility, but these two changes are structurally independent of one another so it is easy to imagine a stage of English (or another language), in which gerunds allowed both a bare subject and also a modifying

adjective. It seems, then, that the Deverbalisation Hierarchy should in fact allow pattern (b).

As for pattern (a), this does in fact conflict with the analysis that I have proposed. On the one hand we have to assume that *destroying* is a noun, and not a gerund, because otherwise we cannot explain the *of*. On our analysis, a gerund inherits the valency of the verb, so in the case of DESTROY this should be an ordinary bare direct object, not *of*. But the bare subject is only sanctioned if *destroying* is a verb (i.e. a gerund); therefore example (a) requires two conflicting analyses of *destroying* and must be impossible for exactly the same reason as **Pat destruction of the evidence*. In this rather limited sense, then, there is a Deverbalisation Hierarchy, which excludes precisely one pattern; and this one exclusion is predicted by the proposed analysis. However it remains to be seen whether the exclusion, and the prediction, are factually correct.

9 Conclusions

The parochial conclusion is that English gerunds are indeed just what the traditional grammarians said: both verbs and nouns. Once this has been said, nothing more is needed in order to generate ordinary gerunds, though special provisions are needed for possessive subjects and *no/any*. In particular there is no need to take precautions in order to prevent verbal and nominal characteristics from conflicting because English is organised in such a way that these characteristics are always orthogonal: nominal features are exclusively concerned with relations external to the gerund phrase, and verbal features with its internal patterns. It is hard to imagine how the analysis could be more parsimonious, elegant or explanatory.

However, it is also worth pointing out some more general conclusions to which the discussion seems to have led. The analysis is possible only because of a number of general assumptions, so if it is right, those assumptions may be right as well. For convenience I will list the assumptions which seem to me to be the most important:

- Word classes are generally identified by atomic names ('verb', 'noun', 'adjective'), and not by a combination of the features V and N; this must be so precisely because the combination [+V, +N] would have to be used for gerunds, so it would not be available (as usually assumed) for adjectives.
- Word classes are fundamental categories, not just convenient abbreviations for bundles of categories (as suggested by Malouf 1998:120). At least in the analysis of gerunds there is no evidence in favour of teasing apart the 'dimensions of categoriality' that contribute to their definition.

- Phrases are redundant all that is needed in syntax is an account of the dependencies between individual pairs of words. It is important to have maintained this position throughout a discussion of gerunds because at one time they looked like particularly clear evidence for the need to identify phrases (NP) as well as individual words (the gerund verb).
- The logic of default inheritance can allow one sub-category to inherit freely from more than one super-category; if the database is self-consistent the logic will work smoothly, as it does in the case of English gerunds.

References

Abney, S. (1987). The English noun phrase in its sentential aspect. MIT PhD dissertation.

- Anderson, J. and Durand, J. (1986). Dependency phonology. In J. Durand (ed.) *Dependency and Non-linear Phonology*. London: Croom Helm. 1-54.
- Baker, M. (1985). Syntactic affixation and English gerunds. *Proceedings of the West Coast Conference on Formal Linguistics* **4**: 1-11.
- Barsalou, L. (1992). *Cognitive Psychology: An Overview for Cognitive Scientists*. Hillsdale, NJ: Erlbaum.
- Bresnan, J. (1997). Mixed categories as head sharing constructions. In M. Butt and T. Holloway King (eds.), *Proceedings of the LFG97 Conference*. Stanford: CSLI.
- Brown, D., Corbett, G., Fraser, N., Hippisley, A. and Timberlake, A. (1996). Russian noun stress and network morphology. *Linguistics* **34**: 53-107.

Chomsky, N. (1965). Aspects of the Theory of Syntax. Cambridge, MA: MIT Press.

- Chomsky, N. (1970). Remarks on nominalizations. In R. Jacobs and P. Rosenbaum (eds.), *Readings in Transformational Grammar*. Waltham, MA: Ginn. 184-221.
- Creider, C. (1999). Mixed categories in Word Grammar: Swahili infinitival nouns. *Linguistica Atlantica* 21: 53-68.

Creider, C. and Hudson, R. (1999) Inflectional morphology in Word Grammar. *Lingua* **107**: 163-87. Denison, D. (1993). *English Historical Syntax*. London: Longman.

- Dienhart, J. and Jakobsen, L. (1985), On clauses, syntagms and the English gerund. *Journal of English Linguistics* 18: 127-52.
- Donner, M. (1986), The gerund in Middle English. English Studies 67: 394-400.
- Evans, R. and Gazdar, G. (1996) DATR: A language for lexical knowledge representation. *Computational Linguistics*
- Fanego, T. (1996), The development of gerunds as objects of subject-control verbs in English (1400-1700). *Diachronica* **13**: 29-62.
- Fillmore, C., Kay, P. and O'Connor, M. (1988). Regularity and idiomaticity in grammatical constructions: the case of *let alone*. *Language* **64**: 501-38.

Flickinger, D. (1987). Lexical rules in the hierarchical lexicon. Stanford PhD dissertation.

Fraser, N. and Hudson, R. (1992). Inheritance in Word Grammar. *Computational Linguistics* **18**: 133-58.

- Gazdar, G., Klein, E., Pullum, G. and Sag, I.(1985). *Generalized Phrase Structure Grammar*. Oxford: Blackwell.
- Goldberg, A. (1995). *Constructions. A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.
- Greenbaum, S. (1996). The Oxford English Grammar. Oxford: Oxford Univesity Press.
- Houston, A. (1989). The English gerund: syntactic change and discourse function. In R. Fasold and D. Schiffrin (eds.), *Language Change and Variation*. Amsterdam: John Benjamins. 173-95.
- Huddleston, R. (1988). English Grammar. An Outline. Cambridge: Cambridge University Press.
- Hudson, R. (1976). Arguments for a Non-transformational Grammar. Chicago: University of Chicago Press.
- Hudson, R. (1990). English Word Grammar. Oxford: Blackwell.
- Hudson, R. (1994) Discontinuous phrases in dependency grammar. UCL Working Papers in Linguistics 6: 89-124.
- Hudson, R. (1998). English Grammar. London: Routledge.
- Hudson, R. (1999a). Encyclopedia of English Grammar and Word Grammar. http://www.phon.ucl.ac.uk/home/dick/enc.htm
- Hudson, R. (1999b). Subject-verb agreement in English. *English Language and Linguistics* **3**: 173-207.
- Hudson, R. (2000a). *I amn't. Language 76:297-323.
- Hudson, R. (2000b). Grammar without functional categories. In R. Borsley (ed.), *The Nature and Function of Syntactic Categories*. New York: Academic Press, 7-35.
- Jack, G. (1988). The origins of the English gerund. Nowele 12: 15-75.
- Jackendoff, R. (1977). X-bar Syntax: A Study of Phrase Structure. Cambridge, MA: MIT Press.
- Johnson, K. (1988). Clausal gerunds, the ECP and government. Linguistic Inquiry 19: 583-609.
- Kaiser, L. (1998). *The morphosyntax of clausal nominalization constructions*. Yale University PhD dissertation.
- Kay, P. and Fillmore, C. (1999). Grammatical constructions and linguistic generalisations: the *What's X doing Y* construction. *Language* **75**: 1-33.
- Labov, W. (1989). The child as linguistic historian. Language Variation and Change 1: 85-97.
- Langacker, R. (1987). *Foundations of Cognitive Grammar I: Theoretical Prerequisites*. Stanford: Stanford University Press.
- Langacker, R. (1990). *Concept, Image and Symbol. The Cognitive Basis of Grammar.* Berlin: Mouton de Gruyter.
- Lapointe, S. (1993). Dual lexical categories and the syntax of mixed category phrases. In A. Kathol and M. Bernstein (eds.) *Proceedings of the Eastern States Conference of Linguistics*. 199-210.
- Malouf, R. (1998). *Mixed Categories in the Hierarchical Lexicon*. Stanford University PhD dissertation.
- Mel'čuk, I. (1997). Vers une Linguistique Sens-Texte. Collège de France: Chaire Internationale.

Milsark, G. (1988). Singl-ing. Linguistic Inquiry 19: 611-34.

- Pinker, S. and Prince, A. (1988). On language and connectionism: Analysis of a Parallel Distributed Processing model of language acquisition. *Cognition* **28**: 73-193.
- Pinker, S. and Prince, A. (1994). Regular and irregular morphology and the psychological status of rules of grammar. In S. Lima, R. Corrigan and G. Iverson (eds.), *The Reality of Linguistic Rules*. Amsterdam: Benjamins. 321-51.
- Pollard, C. and Sag, I. (1994). *Head-driven Phrase Structure Grammar*. Chicago: University of Chicago Pressl.

- Postal, P. (1966). On so-called 'pronouns' in English. *Monographs on Languages and Linguistics* **19**: 177-206.
- Pullum, G. (1991). English nominal gerund phrases as noun phrases with verb-phrase heads. *Linguistics* **29**: 763-99.
- Quirk, R., Greenbaum, S., Leech, G. and Svartvik, J. (1985). A Comprehensive Grammar of the English Language. London: Longman.
- Reisberg, D. (1997). Cognition. Exploring the Science of the Mind. New York: Norton.
- Rosch, E. (1976). Classification of real-world objects: origins and representations in cognition. In S. Ehrlich and E. Tulving (eds.) *La Mémoire Sémantique*. Paris: Bulletin de Psychologie; reprinted in P. Johnson-Laird and P. C. Wason (eds.) (1977) *Thinking: Readings in Cognitive Science*. Cambridge: Cambridge University Press. 212-222.
- Rosenbaum, P. (1967). *The Grammar of English Predicate Complement Constructions*. Cambridge, MA: MIT Press.
- Rosta, A. (1994). Dependency and grammatical relations. UCL Working Papers in Linguistics 6: 219-58.
- Rosta, A. (1997). English Syntax and Word Grammar Theory. London University PhD dissertation.
- Swan, M. (1980/1995). Practical English Usage. Second Edition. Oxford: Oxford University Press.
- Tajima, M. (1985). The Syntactic Development of the Gerund in Middle English. Tokyo: Nanun-do.

Touretzky, D. (1986). The Mathematics of Inheritance Systems. Los Altos, Cal.: Morgan Kaufman.

- van der Wurff, W. (1993). Gerunds and their objects in the Modern English period. In J. van Marle (ed.) *Historical Linguistics 1991*. Amsterdam: Benjamins. 363-75.
- van der Wurff, W. (1997). Gerunds in the Modern English period: structure and change. *The History* of English **3**: 163-96.
- Weber, H. (1997). Dependenzgrammatik. Ein interaktives Arbeitsbuch. Tübingen: Gunter Narr
- Wescoat, M. (1994). Phrase structure, lexical sharing, partial ordering and the English gerund. *Proceedings of the Berkeley Linguistics Society* **20**: 587-98.
- Yoon, J. (1996). Nominal gerund phrases in English as phrasal zero derivations. *Linguistics* **34**: 329-56.