



Title	Phrase Structure, Lexical Sharing, Partial Ordering, and the English Gerund
Author(s)	Wescoat, Michael T
Citation	OUPEL(Osaka University Papers in English Linguistics). 1995, 2, p. 197-210
Version Type	VoR
URL	<a href="https://doi.org/10.18910/72958">https://doi.org/10.18910/72958</a>
rights	
Note	

*The University of Osaka Institutional Knowledge Archive : OUKA*

<https://ir.library.osaka-u.ac.jp/>

The University of Osaka

# PHRASE STRUCTURE, LEXICAL SHARING, PARTIAL ORDERING, AND THE ENGLISH GERUND\*

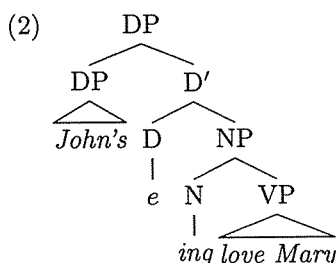
## 1 INTRODUCTION

The English gerund is simultaneously noun-like and verb-like, as (1) suggests:

- (1) a. I was surprised at *John's cleverly having sent his boss an extra copy*.  
 b. \*I was surprised at *the clever having sent his boss an extra copy*.

Occurring as object of a preposition is a property of nouns; allowing adverbs, auxiliaries, and multiple prepositionless objects, while rejecting articles and adjectives is reminiscent of verbs. The gerund's dual nature has disrupted many elegantly principled theories, especially with respect to ENDOCENTRICITY or 'headedness' in  $\bar{X}$ -syntax. Some studies resort to positing EXOCENTRIC or 'headless' NPs dominating VPs. However, at least two linguists, Abney (1987) and Pullum (1991), tackle the construction without compromising endocentricity. I present my own endocentric analysis and observe that it overcomes some theoretical and empirical difficulties to which the others are prone.

Baker (1985) likens the morpheme *ing* to INFL in sentences and proposes it as the head of gerunds. Abney (1987) significantly refines this analysis, positing the structure in (2), whose head is D, a functional category like INFL:



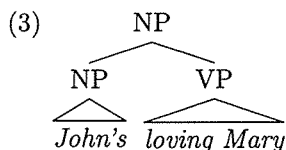
However, (2) has some controversial features: I propose to avoid inaudibilia like  $[D\ e]$  and to maintain a strong lexicalist prohibition against syntactically altered

---

\*This is a slightly revised version of a paper previously presented at the 20th Annual Meeting of the Berkeley Linguistic Society on February 20, 1994 and included in the proceedings of the conference (pp. 587–598). I would like to express my sincerest gratitude to Ernest W. Adams, without whose kindness and sustained guidance I could not have undertaken this study. My thanks also go to Tom Wasow and Peter Sells for their detailed comments on an earlier draft. Of course, responsibility for remaining inadequacies rests entirely with me.

morphology, which is violated when the verb and *ing* are joined transformationally.

Pullum adopts a structure that is traditionally considered exocentric:



His contribution is to note that the logic of GPSG allows syntactic rules to stipulate VP as the head of NP, thereby making (3) endocentric. However, this violates a traditional assumption of  $\bar{X}$ -theory: the 'projects to' relation on syntactic categories should be a disjoint union of discrete linear orders. For instance, NP should be prevented from admitting both  $N^{(i)}$  and VP heads.

Pullum's analysis also over-zealously excludes articles and adjectives from gerunds; it makes no allowance for a variant grammar of English that admits archaic forms like (4)–(7), attested between the 15th and early 20th centuries:

- (4) the untrewē forgyng and contryvyng certayne testamentys and last wyll  
[*Paston Letters*, 15th cent.]
- (5) the uttering sweetly and properly the conceit of the minde [Sidney]
- (6) my wicked leaving my father's house [Defoe]
- (7) the being weighed down by the stale and dismal oppression of remembrance  
[Dickens]

By making VP the head of NP, Pullum denies the determiners and adjectives in these forms the nominal head they require, rendering (4)–(7) impossible to generate.<sup>1</sup> However, such structures coexisted with all modern gerund forms, so it is only plausible that the current and former grammars of gerunds should be largely compatible, in a way that Pullum's approach cannot model.

I will propose here an analysis that accounts for the simultaneously noun-like and verb-like character of the gerund, while still satisfying the various theoretical desiderata outlined above. Also my analysis is capable of accounting for either modern or archaic gerunds with only a single minor change. The approach involves a somewhat radical re-evaluation of the basic assumptions about phrase-structure trees. The new ideas concern the ordering of phrase-structure nodes and the connection between nodes and lexical tokens.<sup>2</sup>

<sup>1</sup>In contrast, Abney seems to make allowances for archaic gerunds. As for adjectival modification of gerunds, he says (1987:198) "This suggests a structure in which the VP is inside of N-bar," implying that D-structure could, if necessary, be made to accommodate the modifiers. He also indicates (ibid.:231) that lexical determiners may replace the empty category [ $\bar{D}$  e].

<sup>2</sup>Formally I distinguish 'lexical tokens' from 'lexical forms.' For instance, in the sentence *the boy likes the girl*, the two instances of *the* constitute two different lexical tokens, both of which correspond to the same lexical form. More will be said on this matter below.

## 2 A NEW AXIOMATIZATION OF SYNTACTIC PHRASE-STRUCTURE TREES

My approach stems from a reaxiomatization of phrase-structure trees. It is rare to pursue syntactic analyses by revising the basic assumptions of tree structure, though McCawley (1982) provides a notable exception. However, I believe this leads to new insights into various constructions, including gerunds.

The new notion of phrase structure trees differs from the standard one in two ways. First, lexical tokens may be shared by multiple leaf nodes,<sup>3</sup> i.e., the function  $\lambda$ , associating leaf nodes to lexical tokens, is not one-to-one. Secondly, leaf nodes are only partially, rather than linearly, ordered by the left-to-right precedence relation  $\prec$ . This allows for the possibility of EXTRA-SEQUENTIAL nodes, which do not participate in left-to-right ordering.

I implement these notions by applying the classical axiomatic method; I first set out a range of primitive notions for a theory of syntactic phrase-structure trees, and then I state a series of axioms to govern the theory.

## 2.1 Primitive Notions

A syntactic phrase-structure tree, denoted as in (8), comprises the primitive notions described in (9):

- (8)  $\mathcal{T} = (N, L, \Lambda, \Sigma, l, \lambda, \sigma, D, \prec, P, \leq, A)$   
 (9)  $N$ , a set of phrase structure nodes;  
 $L$ , a set of syntactic category labels;  
 $\Lambda$ , a set of lexical tokens;  
 $\Sigma$ , a set of lexical forms;  
 $l$ , the label function from  $N$  to  $L$ ;  
 $\lambda$ , the lexical-token function from leaf nodes to  $\Lambda$ ;  
 $\sigma$ , the lexical-form function from  $\Lambda$  to  $\Sigma$ ;  
 $D$ , the ‘dominates’ relation, a strict partial ordering<sup>4</sup> of  $N$ ;  
 $P$ , the ‘projects to’ relation, a strict partial ordering of  $L$ ;  
 $\prec$ , the phrasal ‘precedes’ relation, a strict partial ordering of  $N$ ;  
 $\leq$ , the lexical ‘precedes’ relation, a linear ordering<sup>5</sup> of  $\Lambda$ ;  
 $A$ , the ‘argument/adjunct of’ relation on members of  $N$ .

For clarity, I choose as the idiom in which to state the axioms of the theory a many-sorted, first-order logic with distinct variables for nodes, category labels, lexical tokens, and lexical forms.<sup>6</sup> The notational conventions for the sorted

<sup>3</sup>Leaf nodes, formally defined later, are nodes with no children. Note that I call ‘leaves’ what might elsewhere be termed ‘pre-terminals,’ since I do not regard lexical tokens as nodes.

<sup>4</sup>A relation  $\Pi$  is a strict partial ordering iff  $\Pi$  is asymmetric and transitive.  $\Pi$  is asymmetric iff  $\forall x \forall y (x \Pi y \rightarrow \neg y \Pi x)$ .  $\Pi$  is transitive iff  $\forall x \forall y \forall z ((x \Pi y \wedge y \Pi z) \rightarrow x \Pi z)$ .

<sup>5</sup>A relation  $\Pi$  is a linear ordering iff  $\Pi$  is asymmetric, transitive, and connected.  $\Pi$  is connected iff  $\forall x \forall y (x \neq y \rightarrow (x \Pi y \vee y \Pi x))$ .

<sup>6</sup>Sorted logics are notational conveniences; there are equivalent, ordinary logics with unary predicates to mark distinctions encoded in sorted variables.

variables are summarized in (10):

- |                               |                                      |
|-------------------------------|--------------------------------------|
| (10) nodes in $N$ ,           | letters late in the Roman alphabet;  |
| category labels in $L$ ,      | letters early in the Roman alphabet; |
| lexical tokens in $\Lambda$ , | letters early in the Greek alphabet; |
| lexical forms in $\Sigma$ ,   | letters late in the Greek alphabet.  |

Next I provide various axioms of the theory along with some auxiliary notions.

## 2.2 Nodes and Domination

The notions of nodes and domination are well known, and the axioms below are essentially ordinary. The set  $N$  contains nodes comparable to vertices in a graph-theoretic directed tree. However, where a directed tree employs arcs, I instead use the primitive ‘dominates’ relation  $D$ . I also require the ‘immediately dominates’ relation  $ImD$  defined in (11):<sup>7</sup>

$$(11) \quad \forall x \forall y (ImD(x, y) \stackrel{\text{def}}{=} (D(x, y) \wedge \neg \exists z (D(x, z) \wedge D(z, y))))$$

Note that domination is often described with metaphors of ancestry: an ANCESTOR dominates a DESCENDANT; a PARENT immediately dominates a CHILD. I model the axioms in (12)–(14) on one of the standard graph-theoretic definitions of directed trees (Thulasiraman & Swamy 1992:106ff.):

- (12) ROOT AXIOM  
 $\exists x \forall y (D(x, y) \vee x = y)$   
 There must be a ROOT, i.e., a node which dominates all other nodes.
- (13) ANCESTORLESS NODE AXIOM  
 $\exists x \neg \exists y D(y, x)$   
 There must be a node (inevitably the root) that has no ancestors.
- (14) SINGLE PARENT AXIOM  
 $\forall x \forall y \forall z ((ImD(x, z) \wedge ImD(y, z)) \rightarrow x = y)$   
 A node may have at most one parent.

## 2.3 Partial Left-to-Right Ordering

I next present a new view of left-to-right ordering. Usually siblings are linearly ordered, and that ordering is projected down the tree with an axiom not unlike (16). Thus, the leaves—nodes without descendants, as defined in (15)—are linearly ordered. Instead, I assume merely that  $\prec$  is a partial ordering, and I

---

<sup>7</sup>Making  $ImD$  primitive would more closely reflect the graph-theoretic notion of arcs, but  $D$  would then be  $ImD$ ’s transitive closure, which would not be first-order definable. Similar comments hold for  $P$  and  $ImP$ , to be defined later.

project it up and down the tree in a consistent manner by means of the axioms in (16) and (17):

- (15)  $\forall x(Leaf(x) \stackrel{\text{def}}{=} \neg \exists y D(x, y))$   
 (16) DOWNWARDS PRECEDENCE CONSISTENCY AXIOM  
 $\forall x \forall y \forall z ((D(x, y) \wedge x \Pi z) \rightarrow y \Pi z)$  where  $\Pi$  is  $\prec$  or  $\succ$ .  
 A node precedes or follows every node that its parent precedes or follows, respectively.  
 (17) UPWARDS PRECEDENCE CONSISTENCY AXIOM  
 $\forall x \forall y ((\forall z (D(x, z) \rightarrow z \Pi y) \wedge \neg Leaf(x)) \rightarrow x \Pi y)$  where  $\Pi$  is  $\prec$  or  $\succ$ .  
 A non-leaf node precedes another node if all children of the first node precede the second node, and similarly for the ‘follows’ relation.

The result of these definitions is to make  $\prec$  less restrictive than the usual left-to-right precedence relation. Note, for instance, that leaf nodes need not be linearly ordered. In other words, some pairs of nodes that would have to be ordered on a traditional approach needn’t be under this axiomatization. More ordering constraints will come from the linear ordering of lexical tokens.

#### 2.4 Syntactic Category Labels

Syntactic category labels from the set  $L$  are associated with each node by the function  $l$ . The members of  $L$  are organized according to a version of Kornai and Pullum’s (1990) axiomatization of  $\bar{X}$ -theory. The set  $L$  is ordered by the ‘projects to’ relation  $P$ . The treatment of  $P$  here renders by traditional formal means what is usually encoded notationally with diacritics in most works on  $\bar{X}$ -theory. For instance, given the usual assumptions about the syntactic category labels in (18a), one could diagram  $P$  as in (18b) with respect to these items:

- (18) a.  $N, N', N'', V, V', V''$   
 b.  $P(N, N'), P(N, N''), P(N', N''), P(V, V'), P(V, V''), P(V', V'')$

The category labels in (18) separate into discrete linear orders; one involves  $N, N', N''$ , and the other  $V, V', V''$ . Let us call these discrete linear orders on subsets of  $L$  PROJECTIONS. Since I desire precisely this sort of disjoint union of projections, it is necessary to ensure that each category label has at most one immediate predecessor and successor. I therefore define the ‘immediately projects to’ relation  $ImP$  in (19) and then make it one-to-one with the axioms in (20) and (21):<sup>8</sup>

$$(19) \forall a \forall b (ImP(a, b) \stackrel{\text{def}}{=} (P(a, b) \wedge \neg \exists c (P(a, c) \wedge P(c, b))))$$

<sup>8</sup>Kornai and Pullum take as basic a partial function essentially equivalent to  $ImP$ . They then require that the function be ‘invertible’ and ‘acyclic.’ Mutatis mutandis, the axioms stated here ensure invertibility in Kornai and Pullum’s sense. Furthermore, since  $P$  is a strict partial ordering, it is asymmetric, and that satisfies Kornai and Pullum’s notion of acyclicity.

- (20) UNIQUE  $P$ -IMMEDIATE PREDECESSOR AXIOM  
 $\forall a \forall b \forall c ((ImP(a, c) \wedge ImP(b, c)) \rightarrow a = b)$   
 A syntactic category label has at most one  $P$ -immediate predecessor.
- (21) UNIQUE  $P$ -IMMEDIATE SUCCESSOR AXIOM  
 $\forall a \forall b \forall c ((ImP(a, c) \wedge ImP(b, c)) \rightarrow a = b)$   
 A syntactic category label has at most one  $P$ -immediate successor.

Studies on  $\bar{X}$ -theory often assume that all projections involve a uniform number of categories, but the present axiomatization imposes no such constraint. Thus, the following  $L$  and  $P$  conform to the axioms, even though the categories DET and DEG have no successors:

- (22) a.  $L = \{I, IP, N, NP, V, VP, A, AP, P, PP, ADV, ADVP, DET, DEG\}$   
 b.  $P(I, IP), P(N, NP), P(V, VP), P(A, AP), P(P, PP), P(ADV, ADVP)$

At this time, I have no evidence to suggest that any category needs more than one successor. I will assume the  $L$  and  $P$  in (22) in the present study.

Also part of  $\bar{X}$ -theory are constraints on the distribution of category labels throughout the tree. To state the relevant axioms, I require the unary predicates in (23) and (24), which pick out the  $P$ -minimal and  $P$ -maximal elements of each of the discrete linear orderings, or projections, that make up  $P$ . These are called LEXICAL and MAXIMAL categories, respectively:

$$(23) \forall a (Max(a) \stackrel{\text{def}}{=} \neg \exists b P(a, b))$$

$$(24) \forall a (Lex(a) \stackrel{\text{def}}{=} \neg \exists b P(b, a))$$

The axioms are all familiar ones from the literature on  $\bar{X}$ -theory:

- (25) LEXICALITY AXIOM  
 $\forall x (Leaf(x) \leftrightarrow Lex(l(x)))$   
 All and only leaf nodes are labeled with lexical categories.
- (26) SUCCESSION AXIOM  
 $\forall x (\neg Leaf(x) \rightarrow \exists y (ImD(x, y) \wedge ImP(l(y), l(x))))$   
 A non-leaf node's label is the  $P$ -immediate successor of the label of one of the node's children, known as the HEAD.
- (27) MAXIMALITY AXIOM  
 $\forall x \forall y \forall z ((ImD(x, y) \wedge ImD(x, z) \wedge \neg Max(l(y)) \wedge \neg Max(l(z))) \rightarrow y = z)$   
 A node has at most one child labeled with a non-maximal category.

Pullum (1991) observes that the logic of GPSG makes these axioms a default that may be overridden; I shall take them as absolute.

## 2.5 Lexical Material

In this axiomatization I distinguish between lexical tokens in  $\Lambda$  and lexical forms in  $\Sigma$ . This stems from a need to distinguish instances of words in a syntactic structure from words 'in the lexicon.' For instance, in the sentence *the boy likes the girl*, there are two instances of *the* separated by two other words. If one is to have any sensible notion of lexical ordering, the two instances of *the* must be recognized as separate individuals. I call such individuals lexical tokens. However, for expressing grammatical generalizations, one would prefer to regard the two instances of *the* as representing a single individual: this I call a lexical form.

The function  $\lambda$  associates each leaf node with a lexical token in  $\Lambda$ . Each lexical token is in turn associated with a lexical form in  $\Sigma$  by the function  $\sigma$ . In traditional phrase-structure trees,  $\lambda$  would be a one-to-one mapping from leaf nodes into  $\Lambda$ . However, I impose no such restriction and thereby allow LEXICAL SHARING:  $\lambda$  may associate multiple leaf nodes with the same, shared lexical token.

I provide the linear ordering  $<$  on lexical tokens to reflect their obvious temporal sequence. I propose the axioms in (28) and (29) to relate  $<$  to the precedence relation on nodes,  $\prec$ :

(28) PHRASAL-LEXICAL PRECEDENCE CONSISTENCY AXIOM

$$\forall x \forall y ((\text{Leaf}(x) \wedge \text{Leaf}(y) \wedge x \prec y) \rightarrow \lambda(x) < \lambda(y))$$

One leaf node precedes another only if their respective lexical tokens mirror the ordering.

(29) LEXICAL-PHRASAL PRECEDENCE CONSISTENCY AXIOM

$$\forall \alpha \forall \beta (\alpha < \beta \rightarrow \exists x \exists y (\lambda(x) = \alpha \wedge \lambda(y) = \beta \wedge x \prec y))$$

One lexical token precedes another only if there is a pair of nodes related to the lexical tokens, and the nodes reflect the same ordering.

The gist of (28) is that the orderings on leaf nodes and lexical tokens cannot contradict each other by having leaf nodes ordered one way and their corresponding lexical tokens ordered the other. In contrast, (29) says that the linear ordering on lexical tokens cannot go completely unreflected in the ordering relations among leaf nodes. However, (29) stops short of making the ordering on leaf nodes linear; the use of existential quantification in (29) leaves the possibility of having pairs of unordered leaf nodes.

## 2.6 Argument- and Adjuncthood

The only remaining primitive notion is the 'argument/adjunct of' relation  $A$ :  $A(x, y)$  means  $x$  is either an argument or an adjunct of  $y$ . I use the  $A$  relation for both arguments and adjuncts, because these seem to form a natural class with regard to such phenomena as participating in *wh*-constructions. I assume that arguments and adjuncts can be adequately distinguished according to the



category of the node with which they stand in the  $A$  relation: lexical categories for arguments, and non-lexical ones for adjuncts. Finer distinctions among argument types, e.g., grammatical relations such as subject (*Sub*), object (*Obj*), second object (*Ob2*), and oblique (*Obl*) may be handled with language-particular definitions. For instance the English object might be treated as follows:

$$(30) \forall x \forall y (Sibling(x, y) \stackrel{\text{def}}{=} \exists z (ImD(z, x) \wedge ImD(z, y)))$$

SIBLINGS are nodes with the same parent.

$$(31) \forall x \forall y (ImSib(x, y) \stackrel{\text{def}}{=} (Sibling(x, y) \wedge \forall z ((x \prec z \rightarrow z \not\prec y) \wedge (y \prec z \rightarrow z \not\prec x))))$$

IMMEDIATE SIBLINGS are siblings contiguous with respect to  $\prec$ .

$$(32) \forall x \forall y (Obj(x, y) \stackrel{\text{def}}{=} (A(x, y) \wedge (ImSib(x, y) \vee \dots)))$$

Objects are arguments that are immediate siblings or...

The definition in (32) is left incomplete, because a full definition would have to take into account such matters as long distance dependencies. I will not attempt to flesh out the definitions of *Sub*, *Obj*, *Ob2*, and *Obl* here; though I mention them in the discussion below, their exact definitions are not crucial.

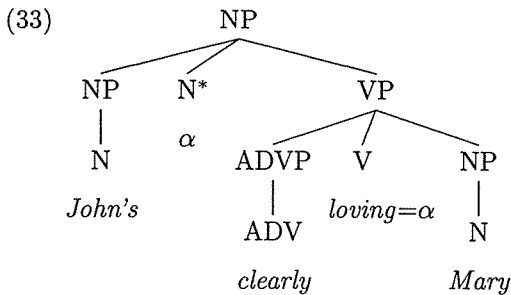
## 2.7 The Grammar

The foregoing axioms provide a theory of syntactic phrase-structure trees, which may be viewed as a sub-theory, in the technical, logical sense, of a theory of universal grammar. This in turn is to be regarded as a sub-theory of theories of grammars of particular languages. Given these relationships, I find it most natural to formulate universal grammar and write grammars for individual languages with the same axiomatic method and the same logical idiom as were employed in the foregoing discussion. Hence, my syntactic rules are axioms, and first-order logic is my 'notation.' However, I will not attempt to provide a complete axiomatization of even a small fragment of English. Instead, I will content myself with providing a few relevant axioms where they are needed to support the empirical discussion below.

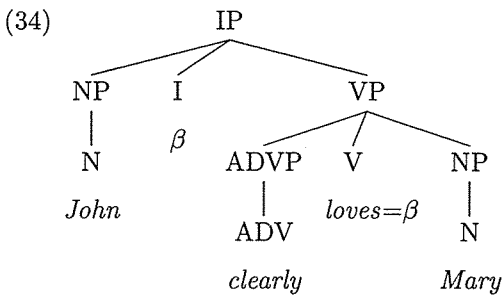
## 3 THE REPRESENTATION OF GERUNDS

The new notion of syntactic phrase-structure trees axiomatized above allows one to provide gerunds with the structural representation in (33):<sup>9</sup>

<sup>9</sup>Of course, if one assumed only the axioms stated here, one could assign a host of undesirable structures to the same string. Further axioms of universal and English grammar are required to limit the possibilities to just (33).



When the  $\lambda$  associates a leaf node with a lexical token, I indicate this fact by writing the latter in close proximity under the former; the arc notation will be reserved for the  $D$  relation among nodes. The use of Greek variables and '=' indicates lexical sharing; for instance, the function  $\lambda$  associates the lexical token *loving* in (33) with two leaf nodes,  $N^{*10}$  and  $V$ . The structure in (33) mirrors that in (34), which I would posit for a sentence:



Assume that the precedence relations on nodes and lexical tokens are reflected in the left-to-right arrangement of nodes in the (33) and (34), with one provision:  $N^*$  and  $I$  are unordered or extra-sequential. Careful consideration will reveal that these structures conform to the various axioms presented so far.

The most obvious novel feature of the trees in (33) and (34) is the use of lexical sharing to maintain endocentricity; the finite verb or participle is the head not merely of a verb phrase but also of a sentence or NP. This effect is comparable to that produced in modern transformational grammar with amalgams formed by head movement. However, this is achieved here without employing inaudibilia or movement rules. The lexical tokens *loving* and *loves* are not composites built in the syntax; rather, they are integral units which require that they be associated with two leaf nodes of the categories shown in (33) and (34). Thus, the strong lexicalist hypothesis may be preserved. Furthermore, these results are achieved without retreating from the notion of the 'projects to' relation as a disjoint union of discrete linear orders, or otherwise weakening the principles of  $\bar{X}$ -theory in any way. Thus the structures in (33) and (34) satisfy all the theoretical desiderata outlined in the introduction. It remains, though, to evaluate (33) on various empirical points, including the matter of archaic gerunds.

<sup>10</sup>The asterisk has no meaning; it merely facilitates unambiguous reference.

### 3.1 *Issues of Syntactic Category*

The tree in (33) clarifies various issues concerning phenomena that are sensitive to syntactic category. Gerunds are clearly NPs, as evidenced by the fact that they may occur as objects of prepositions, whereas non-NPs with similar semantics cannot, as shown in (35):

- (35) a. I was shocked at John's rejecting the application.  
 b. I was shocked at John's rejection of the application.  
 c. \*I was shocked at (for John) to reject the application.  
 d. \*I was shocked at that John rejected the application.

The structure proposed here for gerunds models these facts by positing NP as the syntactic category of the construction as a whole.

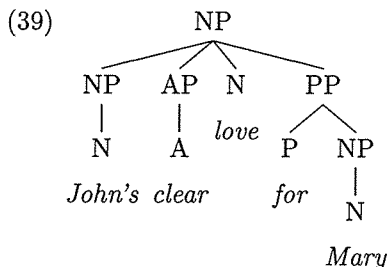
The verb-like characteristics of gerunds tend to be restricted to their internal structure. One such characteristic is the occurrence of OBJECTOIDS, i.e., objects and second objects. Examples (36) and (37) show that objectoids are allowed in gerunds but not in deverbal nominals, the latter being more thoroughly noun-like:

- (36) a. I was surprised at John's destroying the evidence.  
 b. \*I was surprised at John's destruction the evidence.  
 (37) a. I was surprised at John's giving her a valentine.  
 b. \*I was surprised at John's gift her a valentine.

I propose the axiom in (38) to constrain the occurrence of objectoids:

- (38) OBJECTOID SUBCATEGORIZATION AXIOM  
 $\forall x \forall y ((Obj(x, y) \vee Ob2(x, y)) \rightarrow l(y) = V)$   
 Objectoids may be subcategorized only by verbs.

Now, I assume that the structure of deverbal nouns involves no VP structure. Compare (39) with the structure posited for gerunds:



The lack of V and VP nodes in deverbal nominals and their inclusion in gerunds combine with the axiom in (38) to predict the correct distribution of objectoids in (36) and (37).

Similar explanations could be applied to other verb-like characteristics of gerunds, such as the occurrence of adverbs and auxiliary verbs. Suitable axioms, which I will not spell out, would constrain adverbs to occur only inside of VP or IP and require auxiliary verbs to be associated with a V node. This would suffice to predict the distribution of the items in question.

### 3.2 *Issues of Phrase Structure Constituency*

The structures I propose for gerunds, sentences, and deverbal nouns differ with regard to their internal constituency. The two former constructions have a VP node which groups together the predicate and any arguments that follow, whereas deverbal nouns lack such a constituent. This is a felicitous state of affairs, since one of the standard constituency tests suggests that gerunds and sentences differ in the predicted way from deverbal nouns and also from participial forms with prepositionally marked logical objects, which I take to have the same structure as deverbal nouns. Only where the analysis posits a VP can predicates and subsequent arguments participate as a unit in *both...and* coordination:

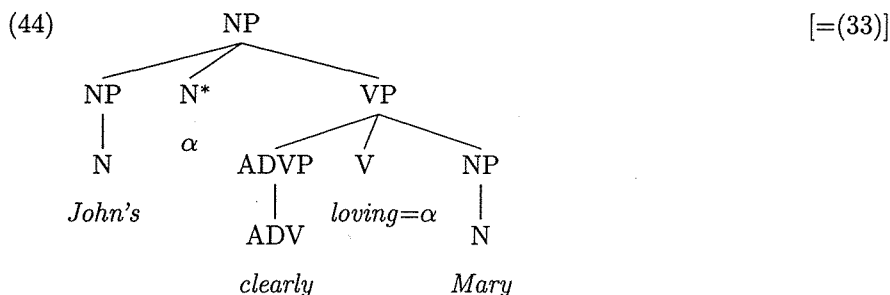
- (40) John both stole a car and killed a bystander.
- (41) It resulted in John's both stealing a car and killing a bystander.
- (42) \*It resulted in the both stealing of a car and killing of a bystander.
- (43) \*It resulted in the both theft of a car and murder of a bystander.

It might be observed that without *both* (40)–(43) are all grammatical. However, as McCawley (1988) notes, coordination without *both* is a significantly weaker test of constituency, admitting a variety of elliptical structures. Thus, the more accurate constituency test supports my predictions.

### 3.3 *Archaic Gerunds and Precedence*

Next I return to the fact that archaic gerunds had verb-like internal syntax and yet allowed determiners and adjectival modifiers. I will propose a single constraint based on left-to-right ordering that applies in Modern English to eliminate these determiners and adjectives. Archaic gerunds would be manifestations of an alternative grammar that simply lacks this constraint.

If one considers the structure for gerunds in (33), repeated below in (44), one will be led to the conclusion that the node V is necessarily unordered with respect to N\*:



This is guaranteed by the theorem in (45), which follows directly from the phrasal-lexical precedence consistency axiom in (28) and the asymmetry of  $\prec$ :

(45) LEXICAL SHARING PRECEDENCE THEOREM

$$\forall x \forall y (\lambda(x) = \lambda(y) \rightarrow x \not\prec y)$$

Nodes sharing a common lexical token do not stand in the  $\prec$  relation.

Otherwise, the axioms do not compel or prevent the ordering of any other pair of leaf nodes. Rather, certain ordering relations will be imposed by separate licensing principles. For instance, the grammar will require English-particular axioms like (46) and (47) to the effect that subjects must precede their verbs, while non-subjects must follow them:

(46) SUBJECT ORDERING AXIOM

$$\forall x \forall y (Sub(x, y) \rightarrow x \prec y)$$

Subjects must precede their verbs.

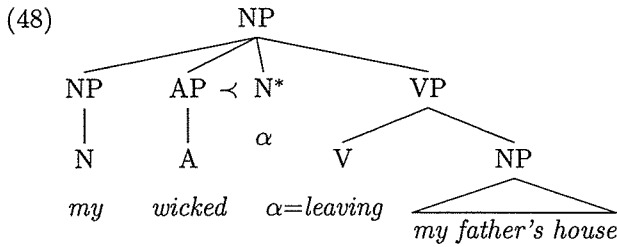
(47) NON-SUBJECT ORDERING AXIOM

$$\forall x \forall y ((Obj(x, y) \vee Obj2(x, y) \vee Obl(x, y)) \rightarrow y \prec x)$$

Non-subjects must follow their verbs.

As a consequence of these axioms, V cannot go unordered with respect to its arguments, *John* and *Mary*. Adverbs, furthermore, conform to a variety of complex ordering constraints with respect to the verbs they modify. Ultimately these regularities should be somehow reflected in axioms, but I shall not attempt to state such constraints here. In any case, V would plausibly also have to be ordered with the projections of *clearly* to satisfy these rules. Thus, V has to be ordered with its neighbors, but nothing I have said so far compels N\* to be ordered with any other element of the gerund. Hence, N\* may be completely unordered and thus extra-sequential.

If one considers now the archaic gerund in (6), whose phrase structure would be that given in (48), one finds that N\* must surely be ordered with AP, since simple adjectives are constrained to precede head nouns:



Articles are similarly constrained and would have to be ordered with their head noun. I will forgo exhibiting either of the relevant axioms. Consequently the articles and adjectives that were allowed in archaic gerunds could be excluded from the grammar of Late Modern English by adding a stipulation that the N associated with the *ing*-form of gerunds must be extra-sequential. This might take the form of the underlined expression in the partial axiom in (49), suggesting how the syntactic properties of *loving* could be stipulated:

- (49)  $\forall \alpha (\sigma(\alpha) = \textit{loving} \rightarrow$   
 $\exists x \exists y (\lambda(x) = \lambda(y) = \alpha \wedge l(x) = V \wedge l(y) = N \wedge \forall z \neg (z \prec y \vee y \prec z)))$   
 A lexical token associated with the lexical form *loving* is related to two nodes, a V and an N, and the N is unordered with respect to all other nodes.

Consequently the syntactic representation with lexical sharing and partial ordering can readily represent either state of the language simply by adding or removing this extra-sequentiality stipulation.

#### 4 CONCLUSION

The reaxiomatization of syntactic phrase-structure trees proposed here allows for an endocentric analysis of gerunds that satisfies all of the desiderata set out in the introduction. In conclusion, it appears that any way one analyzes gerunds will involve some degree of abstraction, be it in the form of positing inaudibilia and movement (Abney 1987), adopting default logic (Pullum 1991), or revising the axioms governing phrase-structure trees. I believe the path I have chosen is novel for concentrating on the association of nodes and lexical tokens and for putting renewed emphasis on the importance of precedence relations in syntactic representations.

#### REFERENCES

- Abney, Steven Paul: 1987, *The English Noun Phrase in Its Sentential Aspect*, Ph.D. thesis, MIT, Cambridge, MA.  
 Baker, Mark C.: 1985, 'Syntactic Affixation and English Gerunds', *WCCFL* 4, 1-11.

- Kornai, András, and Geoffrey K. Pullum: 1990, 'The X-bar Theory of Phrase Structure', *Language* **66**, 24–50.
- McCawley, James D.: 1982, 'Parentheticals and Discontinuous Constituent Structure', *Linguistic Inquiry* **13**, 91–106.
- : 1988, *The Syntactic Phenomena of English*, University of Chicago Press, Chicago.
- Pullum, Geoffrey K.: 1991, 'English Nominal Gerund Phrases as Noun Phrases with Verb-Phrase Heads', *Linguistics* **29**, 763–799.
- Thulasiraman, K., and M. N. S. Swamy: 1992, *Graphs: Theory and Algorithms*, John Wiley & Sons, New York.

*Michael T. Wescoat*  
*Faculty of Letters*  
*Osaka University*  
*1–5 Machikaneyama*  
*Toyonaka, Osaka 560*  
*Japan*  
*wescoat@lisa.lang.osaka-u.ac.jp*