Abstract

Making some fundamental innovations in Word Grammar theory, the thesis explores the nature of syntax (focusing on English) in the areas of (I) relationships between word order and surface constituency, and grammatical relations, (II) the syntax—morphology interface.

(I) Dependency and surface phrase structure ('skeletal trees') get defined; mediating between them is s-dependency, essentially a relation of linear precedence but also crucially involved in (among much else) determining extractability and landing sites for extraction, and defining subordination (demonstrated on prepositional passives). A range of constructions where phrases are in some sense ‘many-headed’ are examined. A relation ‘Proxy’ is motivated for: relative and interrogative pronouns; that clauses; extraposition; cognate objects. WG’s traditional constituency-based analysis of coordination is refined but ultimately rejected and instead a highly detailed dependency-based analysis is provided for all/most varieties: standard, complex, gapping, appended, asyndetic. Partially coordination-like constructions studied include: pied piping; other conjunctions than AND/BUT/OR (e.g. than, as, like, instead of, so, though, nor); ‘subjunctions’ (e.g. negators (not, other than); subjunct adverbials (e.g. even, only, sort of); all but, more/less than; some prepositions (e.g. over thirty went; Romance partitives; these type of dogs; ‘binominal’/‘helluva’ constructions (that peach of a film))).

(II) Issues of what words, syntax and morphology are are discussed in the light of an argument that a sequence of more than one word can be enounced simultaneously (e.g. French DE+LE=du), a fusion-like phenomenon dubbed ‘coenunciation’. Morphological—phonological entities are syntactic words’ symptoms, not their components. Solving persistent problems or yielding drastic simplifications, coenunciation analyses are applied to: Romance and German articulated prepositions; articles; contracted auxiliaries; clitics; right node raising; interrogatives clauses; underlying demoted subjects of passives and mediopassives/middles; gerunds; possessive/genitive ’s; -ly adverbs; depictives (eat meat {raw naked/*naked raw}); deictics (e.g. today); possessive determiners/pronouns; numerics (e.g. twice, second, half); comparatives; superlatives; morphosyntactic forms of verbs; and gallimaufries of lexeme-specific examples.
Thanks

above all and in very large quantities to Maria Novella and Dick.
Also to Nik and other interlocutors.

Dedicated to Edwin H.,
conceived years later but delivered months sooner,
easier to make but far more perfect.

Dedicated also to the memory of Sid Greenbaum.
## Contents

1 Introduction to a version of Word Grammar ......................................................... 11
   1.1 On being a version of Word Grammar .............................................................. 11
   1.2 Grammatical categories ...................................................................................... 15
      1.2.1 Wholism ...................................................................................................... 15
      1.2.2 Inheritance hierarchy .................................................................................. 17
      1.2.3 Default overriding ....................................................................................... 17
      1.2.4 Derivational ‘processes’ and lexical rules .................................................. 18
   1.3 Relationism ........................................................................................................... 22
   1.4 Other key tenets of WG syntax .......................................................................... 22
   1.5 WG and universalism .......................................................................................... 24
   1.6 Notation: rules and diagrams ............................................................................. 26
2 Skeletal dependency .................................................................................................. 28
   2.1 Associacy and dependency .................................................................................. 28
   2.2 Skeletal trees ....................................................................................................... 31
      2.2.1 Projectivity and skeletal constituency .......................................................... 31
      2.2.2 Branch-dependency ...................................................................................... 36
      2.2.3 The role and status of skeletal trees in grammar ......................................... 39
      2.2.4 Previous approaches to nonprojective dependency structures ................. 42
   2.3 S-dependency ....................................................................................................... 43
      2.3.1 Linear precedence ......................................................................................... 44
      2.3.2 No promiscuous dependency ....................................................................... 48
      2.3.3 S-dependency and clausals ......................................................................... 49
      2.3.4 Extractability of subjects ............................................................................. 54
      2.3.5 Tails ................................................................................................................. 55
      2.3.6 Landing sites for extraction .......................................................................... 57
      2.3.7 A constraint on prepositional passives ......................................................... 59
   2.4 S-dependency loops ............................................................................................. 61
2.5 Are there principles governing the correspondence between dependency structure and s-dependency structure? .................................................................64

2.6 Proxies and other non-dependency associacies ...................................................67
  2.6.1 Proxies of relative and interrogative pronouns ..........................................68
  2.6.2 That clauses ............................................................................................74
  2.6.3 Extraposition out of NP ..............................................................................75
  2.6.4 Cognate objects .......................................................................................75

3 Coordination ........................................................................................................78
  3.1 Two analyses of coordination .......................................................................78
  3.2 The ‘MWG’ analysis of coordination ..........................................................81
  3.3 Some inter-word relationships involved in simple coordinations ...............84
  3.4 The ‘TDG’ analysis of coordination ............................................................89
  3.5 Complex coordination and gapping, and supplementaries too .....................108
    3.5.1 Complex coordination .........................................................................108
    3.5.2 Problems of gapping ...........................................................................110
    3.5.3 MWG gapping ....................................................................................114
    3.5.4 TDG gapping .....................................................................................117
    3.5.5 Why complex coordination is symmetric rather than gapping ..........124
    3.5.6 Supplementary coordination, and a generalization about conjunctions 125
    3.5.7 Agreement and other quirks of morphology ........................................127
  3.6 Substitutes ....................................................................................................130
    3.6.1 What is substitute of what .................................................................130
    3.6.2 Formalization and application of the notion ‘dependency type’ ..........137
    3.6.3 Wards and dependency sharing .........................................................140
  3.7 Asyndetic coordination ...............................................................................142

[ 8 ]
Introduction to a version of Word Grammar

This thesis outlines and motivates certain developments in Word Grammar’s theory of syntax.\footnote{On Word Grammar, see Hudson (1990, 1994a).} This chapter gives a sketch of where I’m starting from.

1.1 On being a version of Word Grammar

What is Word Grammar? Such a question asked of any theory should be a technical question, the answer to which could be sought in an encyclopaedia, but in actuality it tends towards being more a social question. Is it the totality of ideas set out in the works of its founder, as with, say, Dik’s (1978) Functional Grammar? Or is it the body of ideas given the stamp of approval by the founder, as with, say, Transformational Grammar? I prefer to try to keep the question technical rather than social (leaving the social question to be dealt with by histories of linguistics in the vein of Harris 1993), and to define the essence of WG in terms of some key tenets that are important either in the theory itself or in differentiating it from other theories; these key tenets I’ll discuss shortly. But WG is also characterized by other things besides the key tenets. First, it is characterized by the goals and ethos of work done within the framework. This I discuss a little later in this section. And second, it is in part constituted by the body of analyses done within the framework hitherto.

As I’ll be explaining, (i) I accept some but not all of the key tenets Hudson has suggested for WG, (ii) the present work largely shares the goals and ethos of other work in WG, and (iii) I accept some but not all of the body of analyses done within the framework. Thus, the key tenets of WG, and above all the body of analyses hitherto done
within WG constitute the stage from which the present work is evolved. Hence I consider what I’m doing to be WG in the sense explained here. WG, then, is here construed as a fluid lineage of ideas rather than as a monolithic, circumscribable theory where it is clear what does and does not count as being within the framework.

Rather helpfully, Hudson (1990: 10) provides a list of what Hudson at least circa 1990 considered to be the main tenets of the theory. I’ll begin with this list (in a close paraphrase of Hudson’s own presentation, and using his terms for the characterizing properties).

(a) WG is **lexicalist**. The grammar refers to no syntactic unit larger than the word.

(b) WG is **wholist**. There is no formal distinction between grammar proper and lexicon: the difference, if any, is merely a matter of generality, grammatical facts being the more general ones and lexical facts the more specific ones.

(c) WG is **trans-constructionist**. Many generalizations apply to more than one construction. In other words, the grammar is not a catalogue of constructions but rather is a catalogue of finer-grained and more general properties that combine in different ways to yield constructions epiphenomenally. This is nowadays a fairly mainstream position, so I’ll say no more about it here or below.

(d) WG is **poly-constructionist**. It accommodates a wide variety of constructions other than the canonical ones that every theory feels it must have a story for.

(e) WG is **relationist**. Syntactic structure is based on grammatical relations rather than on constituent structure.

(f) WG is **mono-stratalist**. The grammar generates only one structure for a given syntagm.

(g) WG is **cognitivist**. “[L]inguistic structures are taken as particular instances of general cognitive structures that apply to things other than language.”

(h) WG is **prototypist**. All linguistic and non-linguistic categories are prototypes.

(i) WG is **implementationist** “because there are now some implementations of fragments of English grammar”.^2

Of (a—i), implementationism is not so much a central tenet of WG as a significant ingredient in the ethos of work done within the theory. WG analyses strive to be explic-

---

[^2]: See Hudson (1990: 11) for references to such work. There has since been further work done, including Volino 1990 (on Italian) and Shaumyan 1996.
it, and this might well lend itself to implementations, but implementation does not drive the theory. But implementationism plays no part in this thesis.

Poly-constructionism (at least as construed by me) is also a prominent feature of the ethos of work in WG, rather than a defining characteristic of the theory, but unlike implementationism it does in part characterize the spirit of the present work. WG's 'poly-constructionist' trait means, to generalize, that WG has a goal of rather strict observational adequacy: the method is more to come up with an analysis for all constructions and then refine the result, rather than refine an analysis for a select few constructions and then seek to apply the result to the constructions thus far neglected. Thus, in WG a highly stipulative analysis of a construction hitherto not analysed in WG would count as progress because it forms a basis for further refinements. An example of this would be Hudson (1990)'s analysis of prepositional passives, which establishes that WG can cope with the construction but, it would be fair to say, lacks the elegance of the analysis of other areas of grammar, such as extraction. I emphasize this point (at greater length than Hudson 1990 does) because it contributes to the ethos of if not WG itself then WG's theoreticians, when compared to other major formalist theories. In this thesis there are no analyses I'd be content to describe as highly stipulative, but there is an interest in grammatical detail, which is a reflection of the same ethos.

Cognitivism is a main tenet that has had little effect on the theory, to the extent that one might argue that the absence of such an expected effect serves to repudiate the tenet. Work in WG has found few if any significant connections between linguistic and non-linguistic knowledge, while at the same time being open to the possibility of there being such connections. This is one reason why I would not consider cognitivism a main tenet. The other reason is that in contrast to more incontrovertibly cognitivist approaches (such as Deane 1994), WG is not grounded in ongoing results from cognitive psychology and neuroscience. This is not to say that WG is incompatible with cognitivism, but it is to say that cognitivism is not a main tenet. For my own part, I view a grammar as a body of beliefs, which necessarily must be part of a larger body of beliefs some of which are not about matters grammatical (since we have beliefs about more than just language, of course), but (i) I do not think this implies a cognitive or other mentalist position, and (ii) it does not imply a view that linguistic and non-linguistic beliefs have much commonality in terms of content.
Prototypism is for Hudson a consequence of cognitivism, but the two are not necessarily bound up with one another. If we take prototypes to be categories the membership of which is defined by resemblance to a certain prototype member of that category, then there seems to be no evidence of grammar involving prototypes. For example, instances of the lexeme Cat are a good candidate for prototype noun, judging by the readiness with which it is chosen as an example noun, yet I am not aware of any grammatical rule or phenomenon that involves a contingency upon a word’s resemblance to instances of Cat. So this strong version of prototypism can straightforwardly be rejected. But a weaker form of prototypism does seem useful in a minor way: a form whereby membership of a category is a matter of degree — that is, where everything is to a greater or lesser extent a member and a nonmember of the category. Such a prototype category gives rise to certain predictions about what form the data should take: for a given putative prototype category, C, if we assemble a range of example sentences differing only in the degree to which a certain element in the sentence is a member of C, then we should expect a gradient of acceptability judgements that correlates with the membership gradient of the variable. This we do not generally find, which indicates that most categories are not prototypes even in the weak sense, but occasionally we do find such acceptability gradients. Examples from syntax are heavy-NP-shift, the acceptability of which varies according to the extent to which the NP shifted is heavy, and extraction out of NPs, the acceptability of which varies according to the extent to which the inter-word dependencies composing the extraction path are complement-like. Thus there seems to be a prototype category (in the weak sense) for ‘heaviness’ and for ‘complementyness’. An example from morphology is comparative -er suffixation to adjectives and adverbs, where we again get an acceptability gradient correlating with the degree of phonological and orthographic deviation from canonical -er comparative forms (cf. bigger, huger, heftier, colossaler, massiver, enormouser, giganticker). (Perhaps this last example is a candidate for prototype category in the strong sense, i.e. with membership defined by resemblance to a central member.) In semantics, prototype categories are commoner. For example, the degree to which a benefactive ditransitive such as peel me a grape is acceptable correlates with the extent to which the verb can be construed as denoting an act of creation or preparation (cf. Pinker 1989, Jackendoff 1990). In conclusion, while we need to accept that prototype categories are possible and desirable
even in a non-cognitivist model, it is not the case that all or even many grammatical categories are prototypes.

This leaves us with lexicalism, wholism, relationism and mono-stratalism as the real main tenets in Hudson’s list. Observe that these define a theory of syntax in particular rather than of language in general. WG when incarnated in the flesh of actual analyses is basically a theory of syntax. It is this theory of syntax with these four remaining tenets plus the body of previous actual analyses within the theory that constitutes my starting-off point. I’ll proceed now by saying more about mono-stratalism, lexicalism, wholism and relationism, in the context of a revised list of key features that define WG’s theory of syntax. I will discuss wholism in §1.2.1, relationism in §1.3, and the rest of the central tenets in §1.4.

1.2 Grammatical categories

1.2.1 Wholism

Wholism, we have seen, is Hudson’s term for the absence of formal distinction between grammar proper and lexicon. The distinction is absent in various respects. First, in contrast to a model in which the grammar builds structures and the lexicon is a store from which items are removed and inserted into the structures built by the grammar, in WG all structures arise from the combinatorial properties of lexical items. This is an increasingly mainstream view, so I’ll say no more about it. Second, in contrast to a view in which the lexicon contains only idiosyncratic irregularities, while regularities are taken care of by some other organ of macrogrammar, in WG the difference between exceptionless regularity and idiosyncratic irregularity is just a matter of degree. This feature of WG has its pluses and minuses. The plus is that WG is entirely unfazed by irregularity and exceptions; it is so well able to handle irregularity and exceptions because it is designed with the expectation of exceptions. A concomitant plus is that WG can play fair with awkward data: since a neat analysis is not sunk by the odd exception, there is no need to ignore exceptions. The minus is that when language is kind enough to deliver up an exceptionless rule, WG is unable to show proper gratitude: for WG a rule with no exceptions is no different in status than a rule with a dozen exceptions, the only difference being that the former exceptionless rule is simpler than the latter rule combined
with the rules stating its exceptions. Perhaps the difference in relative complexity is sufficient reflection of the intuitive difference in status between rules with and without exceptions, but the notational practises used in Hudson 1990, and to a considerable extent maintained here, obscure the difference in complexity by not lumping into one complex rule the rule that has exceptions and the rules stating those exceptions.

Remaining with the topic of WG’s wholism, I’ll move on now to sketch what is in the WG lexicogrammar. It contains nothing but a load of categories of various types that are not formally differentiated but which it is helpful to differentiate for expository purposes. First, there are lexeme categories, such as CAT and HAPPY. These categories contain words — that is, they classify terminal nodes. Second, there are the word-classes (i.e. parts of speech), such as Noun and Verb, which also classify terminal nodes. And third, there are classes like Singular and Plural, which once again classify terminal nodes. Thus, lexemes, word classes and morphosyntactic features are all categories of terminal nodes. Generalizations applying to instances of a particular category are stated in rules pertaining to instances of that category, e.g. “If X is instance of BECOME then X has an xcomp”, “If X is instance of Ditransitive then X has an inject (indirect object)”. In summary, lexemes have no special status qua lexemes — they are merely classes of words among many other classes of word (e.g. Adjective, Ditransitive, Preterite) and the grammar does not know the notion ‘lexeme’. This is how things stand, but it is not wholly satisfactory for a couple of reasons. Firstly there is no principled basis for taking Noun to be a class of words rather than a class of lexemes, other than the apparent lack of a need for classes of lexemes. Secondly, while for function/closed-class words the distinction between word class and lexeme is as blurry as WG would predict, for content/open-class words the distinction is clear. But at present WG has no way to formalize the closed/open class distinction.

---

3 I adopt a typographic convention whereby names of lexemes are indicated by an initial large capital letter with the rest of the name in small capitals. When the name is used as a common noun, meaning an instance of the lexeme, the whole word is given in small capitals. Thus, ‘SANDWICH’ denotes the lexeme itself, while ‘a SANDWICH’ denotes an instance (i.e. member) of that lexeme. The distinction between a category and an instance of the category is not made by Hudson (not for any apparent principled reason).

4 Hudson 1990 does make a formal distinction between categories on the one hand and morphosyntactic features on the other, but I find his formal distinction both unnecessary and too hard to make sense of.
1.2.2 Inheritance hierarchy

Hudson (1990 and much subsequent work) lays great emphasis on the idea that categories are organized into an 'isa hierarchy' or 'inheritance hierarchy', wherein more specific categories are stored further down in the hierarchy than more general categories, and properties attributed to a more general category inherit down to the categories below it in the hierarchy. I have no objection to these ideas, but I see no need to bring special attention to them, as it is hard to imagine things working any differently. A rule like Hudson's "cat isa mammal" is merely equivalent to "every instance of Cat is instance of Mammal", a rule with no special formal properties. If we further state a rule like "every instance of Mammal lactates", then since instances of Cat are instances of Mammal it follows that every instance of Cat lactates. We do not need to introduce any special mechanism of inheritance to do this, and nor do we need to foreground the non-special phenomenon of inheritance.

1.2.3 Default-overriding

Related to points raised in §1.2.1-2 is Hudson's use of default overriding. Default overriding says that a set of rules like (i) "Person has 10 fingers", (ii) "Sophy isa person", (iii) "Sophy has 11 fingers" does not involve any contradiction: although given (i−ii) we might expect (iv) "Sophy has 10 fingers", (iii) wins out over it, because while 10-fingeredness is attributed to (members of) the more general category Person, 11-fingeredness is attributed to the more specific Sophy, and rules pertaining to the more specific win out over rules pertaining to the more general.

For convenience, I tend to follow Hudson's practise, in that a statement of a rule does not imply its exceptionlessness (though when a rule is exceptionless and this fact is relevant, I usually indicate this explicitly). Similarly, rules stating exceptions are only sometimes labelled as such. As I said in §1.2.1, this practise tends to obscure the difference in complexity between an exceptionless rule and a default rule plus its exceptions; that is, it tends to obscure the desired (by me) equation of rule regularity with rule simplicity and of rule irregularity with rule complexity.
1.2.4 Derivational ‘processes’ and lexical rules

In this section I outline the positions taken respectively by Hudson (1990) and by me with regard to lexeme creation and lexical rules. The purpose of the section is to contribute to an overall view of the framework, to provide context for the subsequent chapters, but the issues raised in this section, though of intrinsic importance, do not impinge greatly on those that the subsequent chapters are principally concerned with.

There is in WG an unexploited but — perhaps undesirably — exploitable way of capturing the difference between phenomena ‘in the grammar’ and phenomena ‘in the lexicon’: the latter can be equated with lexeme-creation, i.e. creation of categories of terminal nodes. In many theories, such as Lexical-Functional Grammar (Bresnan & Kaplan 1982, Bresnan 1982a), Lexicase (Starosta 1988), and Government Binding Theory, to the extent that it concerns itself with the lexicon (cf. Keyser & Roeper 1984, Fagan 1988, Carrier & Randall 1992), this is done by lexical rules. In WG the counterpart of lexical rules would be relations holding between lexemes. Hudson (1990) allows for such inter-lexeme relations, and introduces the relation ‘Derivative of’ as a cover term for various varieties of inter-lexeme relation, though he does not investigate what more specific varieties of inter-lexeme relation might be necessary.

Other theories, though, make do without lexical rules; Relational Grammar (Blake 1990) is an example. In some versions of RG (i.e. Arc-Pair Grammar, Postal 1982, via Blake 1990), there are neither lexical rules nor indeed a lexicon — everything is ‘done in the syntax’. I align myself with this position: all syntactic properties of a syntagm are determined by well-formedness constraints on the nodes of which it is composed. There are no inter-lexeme relations like Hudson’s ‘Derivative’. Since RG is polystratal, and since lexical rules can simulate polystratality in the area of valency, and since I reject polystratality and lexical rules, and since polystratality (actual or virtual) adds power to the model, the model I’m proposing is prima facie more restrictive. (Of course, if a model is tightened in one place it tends to splurge out in another, so judgements of overall restrictiveness can only be made by comparing models in their entirety. That said, I am not aware of any compensatory splurge in the model I propose.)

---

5 For example, in the LFG passive formation rule an object is turned into a subject and is no longer an object. This gives us two strata, one prior to application of passive formation and one subsequent to it. (See §5.1 for a discussion of stratumhood.)
It will be helpful to discuss some examples. First I'll consider a hypothetical lexical rule of 'middle formation'. This takes the subcategorization frame of a verb, deletes the subject, and changes the object into the subject. As I noted above, this effectively brings the extra power of polystratality into the grammar of valency. Now, let's see how this could translate into WG. First off, there could be an inter-lexeme relation 'Middle of', so that one lexeme can be middle of another. But WG cannot manipulate abstract valency frames; the closest it can get to 'middle formation' is to say that if X is middle of Y, then for every instance of X either (i) its subject is its object, or (ii) its subject has the semantic role usually borne by the object. ((i−ii) are alternative analyses.) Second, we could dispense with the 'middle of' relation, and say simply that for some verbs either (i) their subject is their object, or (ii) their subject bears the secondary semantic role. (These are the same two alternative analyses as before.) If it turns out to be necessary to distinguish middle verbs from other non-middle verbs that have property (i) or (ii), then a word class 'Middle' (a subclass of Verb) can be introduced. In the model I'm advocating only the latter analysis, making no use of the 'middle of' relation is possible (and ideally no use would be made of the word class Middle, either).

For a second example, consider contrasting ways of handling preteriteness and subjunctiveness. On the one hand, we could have inter-lexeme derivational relations 'preterite of' and 'subjunctive of'. This would predict a potential contrast between, on the one hand, instances of lexeme X that is preterite of lexeme Y that is subjunctive of lexeme Z, and, on the other hand, instance of lexeme V that is subjunctive of lexeme W that is preterite of lexeme Z. But no such contrast is evident in English grammar. Thus the lexical derivational approach would have to curtail this excess of power by stipulating that the preterite of the subjunctive of lexeme L is also the subjunctive of the preterite of L. Consequently I use a contrasting method, where preterites and subjunctives are 'formed in the syntax': we can just have word classes Preterite and Subjunctive, with preterites and subjunctives being terminal nodes that are instances of Preterite and Subjunctive.

The rule is hypothetical, but in fact is pretty much that of Grimshaw 1982 and Bresnan 1982b. It is similar to that of Fagan 1988 and Carrier and Randall 1992, except that for them, following Keyser & Roeper 1984, promotion of object to subject happens in the syntax.

Remember that this is only a hypothetical analysis of middles. I do in fact think that some so-called middles are 'unaccusative', i.e. verbs whose subject is also their object, but others are like passives. For an brief account of how these work, see §5.4.2.
None of the above is to say that lexeme-creation does not exist. Lexemes are indirectly created by derivational morphology; the only objects derived in the lexicon are created by derivational morphology. Derivational morphology creates stems, and every stem corresponds to a lexeme, so creating a stem creates a lexeme. I assume that creating a stem also usually involves creating a new sense (i.e. denotation), and the sense of the stem is also the sense of the lexeme corresponding to the stem. I'll now go on to give a very brief indication of how the morphophonology associated with syntactic terminal nodes is arrived at, and then I'll finish by discussing the limits of the role of derivational morphology.

As I explain at length in Chapter 5, in the model I develop in the thesis, an utterance consists of a logical-syntactic form distinct from, but mapped in a very simple way onto, morphophonological form. To every terminal node in logical-syntactic form there corresponds a morphophonological unit (the ‘enunciation’ of the syntactic node), which is a (possibly singleton) sequence of phonological words. The enunciation itself, E, is a realization of a morphophonological ‘blueprint’ composed of (except in cases of suppletion, such as \(\text{went}\)) the stem associated with the lexeme whose instance is the node \(N\) that has \(E\) as its enunciation, plus other morphological material (e.g. inflections, clitic morphemes) that is due to \(N\)’s properties and syntactic environment. The morphophonological blueprint is subject to morphophonological well-formedness constraints, and \(E\) itself is perhaps also subject to similar constraints on realization, though as this takes us deep into the territory of phonology we needn’t enquire here further into such matters.

What I have just described is a version of a word-and-paradigm model of morphology (see, e.g., Matthews 1974, which explains and advocates such a model). From its beginnings WG has taken the word-and-paradigm approach to inflectional morphology, (see in particular Hudson 1984) and the ‘biplanar’ model outlined in Chapter 5, with its complete split between the logical-syntactic plane and the morphophonological plane, provides a fundamental theoretical and logical motivation for the word-and-paradigm model, to supplement earlier data-driven motivations. Besides being supported by the

---

8 Some examples of what I have in mind are:

<table>
<thead>
<tr>
<th>Structure</th>
<th>Realization</th>
</tr>
</thead>
<tbody>
<tr>
<td>([\text{up}_p {\text{grade}}_v])</td>
<td>(\text{to upgrade})</td>
</tr>
<tr>
<td>([\text{up}_p {\text{grade}}_v {\text{an}}_n])</td>
<td>(\text{an upgrade})</td>
</tr>
<tr>
<td>([\text{up}_p {\text{grade}}_v {\text{to}}_n])</td>
<td>(\text{to upgrade})</td>
</tr>
<tr>
<td>([\text{up}_p {\text{grade}}_v {\text{mad}}_n])</td>
<td>(\text{upgrade-mad})</td>
</tr>
</tbody>
</table>

These are morphophonological structures, consisting ultimately of sounds. The subscripts ‘\(N\)’, ‘\(V\)’, etc. indicate not that the stem is a member of the class of nouns, verbs or whatever but that the stem corresponds to a lexeme whose members are members of the class of nouns or verbs or whatever.
data (see Robins 1959, Matthews 1972, 1974, Anderson 1977, Zwicky 1985a) and logical considerations about the architecture of utterances (see Chapter 5), the word-and-paradigm model is supportable by an argument analogous to the one above used in rejecting inter-lexeme relations in the analysis of subjunctive preterites. If preterites and subjunctives arise from the application of a preterite affix and a subjunctive affix, then we should expect (but in fact fail to find) a contrast between a form in which the subjunctive affix applies before the preterite one and a form in which the preterite affix applies before the subjunctive one.

As announced above, I finish this section by discussing the limits of the role of derivational morphology. We'll consider the relationship between dismiss and dismissal. We have already ruled out an analysis in which a word of the form :dismissal belongs to a lexeme that is ‘nominalization of’ the lexeme DISMISS. If instead the two are related by derivational morphology, then a word of the form dismissal would belong to the lexeme that corresponds to the stem \([\text{stem of DISMISS}]_{\text{Noun}}\) or \([\text{stem of DISMISS}]_{\text{AL}}\). It is only fair to admit, though, that this uncontroversial and natural-seeming analysis is not guaranteed to work in the current model if the nominalization keeps some of the unpredictable valency properties of the verb it is derived from. This is because only syntactic nodes have valency; morphophonological blueprints certainly do not. The data for nominalizations is in fact quite messy on the face of things, with some valency properties apparently inheriting to the putatively derived form, and other properties not. But suppose, for example, that the legitimacy and interpretation of the complement of in dismissal of the guard is somehow contingent on the legitimacy and interpretation of the object in dismiss the guard. In this case we would have to take what is tantamount to a word-and-paradigm approach even to dismiss:dismissal: we would employ a single lexeme DISMISS, some of whose instances are also instances of Noun, and other of whose instances are also instances of Verb. The instances of DISMISS and Verb would have enunciations that are realizations of a morphophonological blueprint containing the stem \{dismiss\}, and the instances of DISMISS and Noun would have enunciations that are realizations of a morphophonological blueprint containing the stem \{dismissal\}. It should be emphasized that this is an analysis I’d be forced into reluctantly, due to the current impossibility of treating valency patterns as abstract objects attributable to morphological objects and instantiable by a syntactic node and its complements.
1.3 Relationism

Grammatical relations are basic (rather than derived). That is the principle of relationism. It’s these grammatical relations that the bulk of the syntax component of the grammar is concerned with: most rules state which words can be related to which words, by which relationships. These relationships are, in logical terms, two-place predicates. We are not yet in a position to offer an exhaustive list of what kinds of relationship there are, but a partial list of the relations I recognize can be found by going through the glossary at the end of the thesis, and a list of the relations recognized by Hudson 1990 can be found by going through the index of that opus (relations are there marked with ‘%’) and in particular in its Chapter 9.

As for the question of what structures these grammatical relations form for the multifarious constructions of English, it is clearly beyond the scope of this introduction to outline these, and it is also somewhat incidental to the principal purposes of this thesis. I have followed the analyses of Hudson 1990 where I accept them, and, sometimes without comment, have followed my own analyses where I don’t accept Hudson’s. I’ve tended to rely on diagrams to explicate structures of grammatical relations. The system behind the diagrams is fairly self-explanatory (some supplementary explanation is given in §1.6 and §2.2.1).

In summary, (i) the majority of a WG analysis of a language’s syntax will consist of rules specifying what kinds of grammatical relations there are, and what structures they form, and (ii) only incidentally does this thesis deal with (i).

1.4 Other key tenets of WG syntax

(I) and (II) are important in defining the tenor of the theory.

(I) Grammar is a list of constraints on reading structure into utterances.

(II) Syntax is monostratal.

---

9 One can easily conceive of categories with three or more arguments, such as ‘Parent-Spouse(x,y,z)’, where y is the child of x and z the spouse of x, or, for the phrase Green ideas sleep, a three-place category whose first argument is filled by ideas, whose second is filled by the adjunct of ideas, or the word that precedes ideas, and whose third is the word that ideas is the subject of, or that follows ideas. However, such categories are proscribed by WG, for the sake of restrictiveness. The obligatory binarity of relations is widespread in linguistic theory, if generally unremarked upon; to take a few random examples, grammatical relations in Relational Grammar (Blake 1990) and Lexical Functional Grammar (Bresnan & Kaplan 1982) and conceptual relations in the systems of de Beaugrande (1980) and Sowa (1983) are binary.

10 And also what kinds of word classes there are.
(I) means that grammar states the properties one must believe an utterance to have in order for it to be deemed a manifestation or symptom of a grammatical structure. To express this idea another way, WG models the grammar as a body of beliefs about utterance events. Although this view of the nature of grammar has some important consequences, e.g. in the analysis of deixis and grammatically-encoded illocutionary properties (Hudson 1990, Rosta 1995b), for syntax its most important consequence is in the way it constrains the formalization. Specifically, it constrains, by means of a gradient of plausibility, the ontological variety of the sorts of thing that are involved in constituting a licit syntactic structure: the sorts must be ones that are plausibly part of our beliefs about utterance events. Sorts that are plausibly part of our beliefs about utterances are those that are found in our beliefs about the world in general. In the version of WG that I'm using, there are just three sorts: events, classes of events, and collectivities of events. Terminal nodes are events, word classes and lexemes are classes of events, and phrases are events that are collectivities of events (i.e. they are events composed of a sequence of temporally contiguous events). Those who are concerned only with the formal properties of the model of grammar, may not feel inclined to care about these issues of what we might call ‘ontological parsimony’ or ‘ontological realism’, but even then, ontological parsimony at least provides some criteria for preferring one formalization over another.

As for (II), monostratalism, I say a bit more about what I mean by this in §5.1, where I also suggest that Relational Grammar is the only major necessarily polystratal theory. Thus, although WG is resolutely monostratal, this characteristic does not go far in differentiating WG from other theories. But in informal, nontechnical use, “monostratal” gets used to mean “without transformations”, and it’s also true that WG

---

11 These beliefs are those held by the knower of the language, not those held by the linguist who investigates language.
12 Hudson 1990 also recognizes feature values as a distinct sort, but from a WG perspective this has always looked suspicious, since feature values constitute an ontologically novel sort not attested in beliefs about things other than utterance events. His proposal can be remedied by analysing feature values as classes of events, just as word classes and lexemes are. But I myself prefer to do without the notion of attribute and value altogether, for the sake of ontological parsimony; see §7.3.2.
13 A justification of treating words as, more specifically, actions, performed by the speaker, is given by Hudson 1990 and Rosta 1995b. Irrespective of the grounds for treating words as actions, the overarching reason for treating words as events is that they are temporally sequenced, a property possessed only by events. See further §5.2.
14 Hudson 1990 does not recognize phrases, but (unlike me) does recognize ‘word strings’, which are also collectivities of events.
15 I take it that it is not necessarily the case that two generatively equivalent formalizations are of equal merit. But I don’t think that a long excursus into a defence of that view would help the general description of WG in which we are currently engaged.
is transformationless, being without chains of coindexed empty categories. However, as I explain in §7.3, a transformational model is merely a (slightly inferior) notational variant of the model assumed by WG, particularly the version developed in Chapter 2 (especially if linear precedence is taken out of the picture so that only immediate dominance is left). Therefore I don’t claim transformationlessness as a key tenet of WG.

(III) (a statement of what Hudson termed ‘lexicalism’) is a constraint whose particular importance is in defining WG’s salient differences from other theories, since most other theories still reject (III) or at least have not volubly embraced it.

(III) All nodes are terminal.

Hudson’s (1984, 1988a, 1990) version of WG allows non-terminal nodes solely to deal with the grammar of coordination. In Chapter 3 I show that this is unnecessary. I argue in Chapter 2 that it is worthwhile deriving a rudimentary constituent structure from the structure formed of grammatical relations, but, as I explain then, the grammar does not need to refer to the resulting nonterminal nodes, so they constitute at most a trivial exception to (III). However, as I speculate at the end of §4.6.4, it could well be that the difference between adjuncts and non-adjunct dependents is that adjuncts modify non-terminal nodes and non-adjunct dependents are subordinated to terminal nodes. It would therefore not be the case that all nodes are terminal. In summary, then, (III) is a default assumption but is abandonable in the face of good evidence without bringing the rest of the theory down in ruins.

In Chapters 5—6 I argue for (IV), which is fundamental enough to amount to another key tenet.

(IV) Words (i.e. syntactic terminal nodes) are not made up of morphemes or sounds.

1.5 WG and universalism
Related to WG’s poly-constructionism is the fact that the goal of most work done in WG to date has been to investigate the grammars of particular languages (chiefly, but
by no means exclusively, English). The goal has not been to ascertain the nature of Universal Grammar, and indeed work in WG has not assumed that UG exists; the method of investigating the existence of UG would be to discover by inductive comparison of the grammars of individual languages the properties that they share. While much of what WG has to say about one language tends to be applicable to others, an analysis of language X does not entail any particular analysis of language Y, and similarly if an analysis of language X fails to work for language Y, this does not invalidate the analysis for X. This failure to treat the existence of UG as axiomatic is independent of the WG model itself, but to adopt the axiom would very likely have an effect on the analyses offered for particular languages, in that a suboptimal analysis of a particular language might be preferred to a better one, on the grounds that it comports better with an analysis of UG. In comparison with other theories, WG's poly-constructionism — its descriptive comprehensiveness — and its rejection, in practise, of axiomatic UG, and its concomitant tendency not to take a cross-linguistic perspective, go a long way towards characterizing work in WG.

This point bears repeating, because unless one knows what exactly my proposals are a theory of, one cannot properly evaluate them. My position is that I am proposing a theory of English. Since the grammar of an individual language is an autonomous system of rules (or so I would argue), it is legitimate to study it in its own right, and a theory of an individual language is no less valid than a theory of language in general. This theory of English may (and, I believe, would) prove applicable to other languages, but data from other languages is strictly relevant only if it shows that a given phenomenon is possible in language, but it cannot be counterevidence to an analysis of English. No claims are made — except in passing — about languages other than English, but it is nonetheless my view that notions of at least (i) dependency, (ii) s-dependency (see Chapter 2), and (iii) biplanarity (see Chapter 5) are universal, as is the s-dependency loop solution for coordination and pied piping (see Chapters 2—4), to the extent that coordination and pied-piping are all universal, though I doubt that many of the details are. In sum, the phenomena constituting the main points of the thesis are sufficiently fundamental to the theory of grammar for me to believe them universal properties of language design.

---

\[25\]

\[16\] It may be less awesome, though.
Hudson's view on the universalism question is somewhat different from mine. For him, WG is at one and the same time a theory of the properties that are universal to language and a theory of the properties that are particular to individual languages. The fact that WG has so far not had much to say about which of its proposals are universals and which aren't is a reflection of nothing more than a reflection of (i) an insufficiency of the quantity of work so far done within the theory, and (ii) a preference for making relatively concrete claims about phenomena for which there is a relative abundance of reliable evidence, as opposed to relatively abstract claims about phenomena for which there is a relative scarcity of reliable evidence.

1.6 Notation: rules and diagrams
Grammatical rules are in Hudson's work on WG given in a theory-particular notational system. The very complex system of Hudson 1984 is completely overhauled and replaced by a new system in Hudson 1990, where the notation is closer to normal English. The latter system has its problems, though, in that it hovers in a limbo of semantic uncertainty between being written in normal English and being written in a special formal metalanguage. It does not seem an important enough matter to devote space here to an account of some specifics of the problem with that notation: instead, my attempts in §3.5.2 to make sense of the formalization of a rule from Hudson 1989 will illustrate the kinds of problem that arise.

I have written rules as far as possible in normal English that will translate in an obvious and straightforward way into an unambiguous notation like that of predicate logic. Sometimes I've used predicate logic notation for quantification and connectives, when sticking purely to normal English would result in cumbersomeness or dangerous ambiguity. Note that predicate logical notation is therefore being used as an augment to normal English, and not as something specific or peculiar to WG. Note also that the method of formalizing rules does not constrain or determine what is a possible rule in WG. In fact, there are so far no constraints on what is a possible rule: there are only constraints on what is a probable rule (which can be gauged, roughly, by the amount of complexity and new machinery that must be introduced in the grammar for it to work).

For purely practical reasons, I also use a different notation in diagrams from that used by Hudson in all his work on WG, and by other dependency grammarians, such as
Mel'čuk 1988. Where Hudson would have (A) (but with curved arcs rather than rectilinear lines), I have (B).

\[
\begin{array}{c}
  s \\
  \downarrow \\
  \text{(A) She is.}
\end{array}
\]

\[
\begin{array}{c}
  s \\
  \downarrow \\
  \text{(B) She is.}
\end{array}
\]

Both diagrams show *She* is subject of *is*. I have used (B) because experience shows that people find it easier to understand, since (A) sometimes gets misunderstood (by beginners, at least) as indicating that *is* is subject of *she*. Even if this reason did not exist, exigencies of a graphics-impoverished word processor very much favour (B). If impressionistically my diagrams sometimes look more complicated than Hudson’s, this is not because of the diagram style but because I believe there is more syntactic structure present than Hudson does.
2

Skeletal dependency

2.1 Associacy and dependency

In this section I introduce three kinds of relationship, each of such a generality that every word (in a syntagm of more than one word) is an argument of each kind of relationship. They are ‘associacy’, ‘dependency’ and ‘s-dependency’: they are relationships such that one of their arguments is, respectively, ‘associate’, ‘dependent’ and ‘s-dependent’ of the other. \(^{18}\) Associacy and dependency will be defined in this section. \(^{19}\) S-dependency will be discussed at length in the rest of the chapter.

In any collection of things we are at liberty to discern relationships between members of the collection. Thus, in a group of people we may conceive of Person X as friend of Person Y, of B as disciple of J, of Y as secret admirer of B, and so on. Similarly, in a sequence of words we may conceive of various relationships holding between them. A subset of these relationships holding between words are recognized by the grammar:

\(^{17}\) Rosta (1996) is a condensed and presentationally simplified version of §§2.1—4.

\(^{18}\) The converse of ‘Dependent’ is ‘Head’ and converse of ‘S-dependent’ is ‘S-head’. Some recent work, e.g. Hudson (1995b, 1995c), has established ‘Parent’ and ‘Anchor’ as synonyms of ‘Head’, the term ‘Head’ being disfavoured because of its significantly different meaning in Phrase Structure Grammar. The present ‘Dependent’ corresponds in practise quite closely to Hudson (1990)’s ‘Dependent’, while the ‘Dependent’ of Dependency Grammar is more like a conflation of Dependent and ‘Branch-dependent’/‘Niece’ (on which, see below). The ‘Dependent’ of Rosta (1994) has bifurcated into S-dependent and Branch-dependent/Niece. Rosta (1994)’s ‘Grammatical Relation’ (‘GR’) has bifurcated into Associate and Dependent. Rosta (1994)’s motive for allocating the terms in this way was a suspicion that dependency (in the traditional WG sense) will prove to be a theoretically vacuous notion, amounting to little more than “relationship between words in a sentence”, or, more probably, to the disjunction “subject or complement or adjunct or extractee or …”, a disjunction to which the grammar may ultimately not need to refer. But, as I show in this section, that suspicion was incorrect. My choice had some unfortunate consequences. First, it disagreed with Hudson’s usage and precedent: he has always used ‘dependency’ and ‘grammatical relation’ pretty much synonymously. Second, my use of ‘GR’ to denote not the relationship but the words related by it was a source of mild confusion. It is for these reasons that I have executed a terminological about turn. I have not always scrupulously distinguished Associate from Dependent, sometimes using the latter when the former would be more appropriate. My excuse for this is a wish to use the more familiar term rather than the less familiar term.
these grammatically significant syntagmatic relationships between words, I shall term 'associacies'. When an associacy obtains, one word is 'associate' of another.20

Words are arguments of other kinds of relationship besides associacy. There are relationships between words and their meanings, such as 'Sense of' and 'Referent of'. There are relationships between words and their morphological and phonological form, such as 'Stem of' and 'Enunciation of'. And there is the Instance relationship, which is not specific to language; it relates a word to the classes the word belongs to, such as Tetchy (the lexeme) and Adjective.

Dependency, being a grammatically relevant syntagmatic relationship between words, is a kind of associacy: if X is dependent of Y then X is associate of Y. It is defined by two properties. The first is, roughly, that, except for adjunct dependencies, if the same kind of dependency holds between word X and its dependent Y as holds between X and its dependent Z, Y and Z must be conjoined. Thus, for example, *She he would is ungrammatical because she and he are both subject of would, but aren't conjoined, as they are in the accordingly grammatical She and he would. A rough formalization of this is given in (1—6) (the behaviour of adjuncts is ignored here, for simplicity's sake). Every word has a subject-set that contains all its subjects, an object-set that contains all its objects, and so on, for each kind of dependency. ((2a—c) are rules of terminological convenience: the relationship 'member of the subject-set of' is made equivalent to 'subject of' because the latter is shorter.)

(1) a. Every word has a subject-set.
b. Every word has an object-set.
c. etc.

(2) a. X is subject-set of Y and Z is member of X iff Z is subject of Y.
b. X is object-set of Y and Z is member of X iff Z is object of Y.
c. etc.

Every word also has a dependent-set, which contains its subject-set, its object-set, and so on.

---

20 Associacy may or may not be the same as Hudson (1990)'s Companion relationship. If X is head of Y or X is dependent of Y then X is companion of Y, which is compatible with Companionship being Associacy, but it is not clear whether if X is companion of Y then X is head of Y or X is dependent of Y, which is not compatible with Companionship being Associacy.
(3) Every word has a dependent-set.

(4) a. If X is subject-set of Y and Z is dependent-set of Y then X is member of Z.
b. If X is object-set of Y and Z is dependent-set of Y then X is member of Z.
c. etc.

(5) is the rule that states the coordination requirement on multiple dependents of the same type.

(5) If X is dependent-set of Y, and Z is member of X, and V and W are members of Z, then V and W are conjoined (or are identical).

(6) is a rule of terminological convenience that interdefines 'dependent' and 'dependent-set'.

(6) X is dependent-set of Y and Z is member of X and W is member of Z iff W is dependent of Y.

The other defining property of dependency is (7).

(7) If X is s-dependent of Y then X is dependent of Y.

This says in effect that dependencies are potential s-dependencies. To make sense of this, we must proceed to investigate s-dependency.

Whereas associacy is a nebulous cover-all notion, and dependency is given coherence as a notion by only two properties, s-dependency appears to have a greater range of properties. A good way to get an initial intuitive notion of s-dependency is to imagine that WG was polystratal: dependencies that are s-dependencies would obtain at 'surface structure' and dependencies that are not s-dependencies would not. Accordingly, the 's' of 's-dependency' can be understood as standing for 'skeletal', 'surface', 'scaffold', 'structural', 'subordinating'... From s-dependencies (only from s-dependencies but not from all s-dependencies) is built the skeletal tree responsible for continuity (of
phrases), and in part for processing complexity and for intonation phrasing. This is discussed in §2.2 and again in §§2.4–5. S-dependency also provides the means of stipulating (i) linear precedence restrictions (§2.3.1), (ii) that dependencies cannot unconstrainedly hold between any pair of words that in themselves are eligible to be linked by a dependency (§2.3.2), (iii) when words must 'remain in situ' and when they are 'movable' (§2.3.3–5), (iv) permitted 'landing sites' for extraction (§2.3.6), and (v) that 'movement' is 'upward' (§2.3.7).

2.2 Skeletal trees

§2.2.1 explains how every sentence must contain a skeletal tree. §2.2.2 explains the correspondence between the skeletal tree and s-dependency structure and dependency structure. §2.2.3 considers the role of skeletal trees in grammar. §2.2.4 sees how WG used to fare without skeletal trees.

2.2.1 Projectivity and skeletal constituency

In (8a), the precedes went, as subjects should, and children follows the, as complements should. In (8b), book properly follows the, and the and tomorrow properly follow read. Yet the word order of (8a–b) is ungrammatical. Why?

\[
\text{(8) a. * The went children.} \\
\text{of O of C} \\
\text{b. * Read the tomorrow books.} \\
\text{of O of A} \\
\text{of O of C}
\]

The solution, put forward in Hudson (1994b) and Rosta (1994), is that for every sentence it must be possible to construct a 'skeletal tree' such that to every branch in the tree there corresponds a dependency, with the lower node being dependent of the higher. The nature of this correspondence is described in §2.2.2. In this section we'll look at other properties of skeletal trees, namely (9a–d). (9a) is merely an assumption; it could be abandoned without consequence to the other restrictions on skeletal trees.

\[\text{[31]}\]
There is one node per word. In the rectilinear diagrams I'm currently using, a vertical line is equivalent to a node, and taller verticals are superordinate to shorter verticals.

There is exactly one tree per sentence.\textsuperscript{22}

Branches don't converge.

Branches don't cross.

(9c–d) are a statement of the principle of projectivity, which most versions of Dependency Grammar adopt in some form or other. (9d) is a requirement of continuity; it is equivalent to a prohibition against discontinuous constituency – a point which will be made use of shortly.

\textsuperscript{22}This constraint is stated without an adequate account of parentheticals having been worked out. (i), for example, might contain two trees, as indicated.

But while I do think (i) may well involve two trees, I think it is two sentences rather than one; that is, there are two entirely independent macrosyntagms. Weight is lent to this view by the fact that in frescoes by Tiepolo, the TIEPOLO refers to Tiepolo, while in Tiepolo, I think his name is, Tiepolo signals reference to the name Tiepolo. Two different lexemes seem to be involved, and therefore two different words, each enounced with the morphological form Tiepolo; one lexeme means 'Tiepolo' (the person) and the other lexeme means 'Tiepolo' (the morph). The structure of the sentence in (i) might therefore actually be as in (ii), with words indicated by blobs, and the morphs that enounce them represented in italics. Syntactically, the structure of (ii) is identical to (iii). See Chapters 5–6 for fuller exposition of the nature of the syntax–morphology interface.

(iii) She used to like frescoes by Tiepolo ('Tiepolo', I think his name was) that are full of putti.

If the analysis in (ii) is correct, then such parentheticals do not constitute an exception to the one-tree-per-sentence rule, because they in fact consist of two sentences, that happen to share some morphology.
(8a—b) yield the skeletal trees shown in (10a—b). The branches of these trees cross, and hence the sentences are ungrammatical.


b. *Read the tomorrow books.

For the same reason, (11a) is grammatical but (11b) isn’t. The skeletal tree (or, in the case of (11b), the closest approximation to it) is shown by stemma, and the dependencies the branches correspond to are shown in italic boldface. The interesting thing about these examples (originally from Dick Hudson, and discussed in Hudson 1994b and Rosta 1994) is that all the dependencies that in (11a) are sufficient to provide a legitimate skeletal tree are present in (11b), yet (11b) cannot provide a legitimate skeletal tree.

(11) a. Tomorrow who will come?

b. *I wonder tomorrow who will come.

Let us proceed, then, to formalize projectivity. Constraint (9b) is formalizable as rule (12), the effect of which is that for any two words linked by a dependency, one is subordinate of the other, or both are subordinates of a third; it is assumed that every word in a sentence is a head or dependent of another word in the sentence, and not a head or dependent of any word not in the same sentence.

---

23 Provisionally, we can understand a subordinate of W as a word dominated by W in the skeletal tree. Formally, a subordinate of W is an s-dependent of W or a subordinate of an s-dependent of W.
One tree per sentence.

∀x ∀y if x is dependent of y then ∃z z is unsubordinate of x and z is unsubordinate of y

And then we need rules (13a–b), which I shall henceforth assume obtain, along with (12). (13b) ensures that for every word except the root of the tree, there is exactly one word that is its aunt; this will be made use of in §2.2.2. (13a) gives us the ban against convergence.

(13) a. Every word is daughter of exactly one phrase.
    b. Every phrase contains exactly one word and zero or more phrases

As for the ban on crossing, i.e. the requirement of continuity, I would argue that the use of constituency means we need no stipulation of continuity specifically for syntax. That skeletal constituent structures can be derived from a projective relational structure, and vice versa, is hardly news. It has long been held24 that a projective dependency grammar and a phrase structure grammar are equivalent, in the sense that each can be derived from the other.25 However, that constituent structures and projective dependency structures are each derivable from the other does not mean they are equivalent in all respects, and nor does it mean there is no point in building constituent structures off dependency structures. With constituency, continuity comes for free: continuity may reasonably be considered inherent (at least by default) in phrase structure. This is because the rule that parts and wholes, of which phrases are a subtype, are continuous applies throughout the grammar,26 and in fact could be inherited from extragrammatical knowledge — a leg, for example, could not contain a thigh and a shin but not a knee, and a

---

24 An unsubordinate of W is W or a superordinate of W. A superordinate of W is an s-head of W or a superordinate of an s-head of W.

25 At least since Gaifman (1965), Robinson (1970), but probably earlier by anyone who gave the matter a moment’s thought. See Fraser (1989).

26 The equivalence holds only if the dependency structures are projective. The dependency structures of WG certainly aren’t equivalent to phrase structures. Projective dependency structure is equivalent to unlabelled bracketing if a phrase can’t be mother of more than one word, and is equivalent to headed phrase structure if a phrase can contain more than one word. See the discussion in Fraser (1989, 1990). It is the mistaken assumption that dependency grammars necessarily impose the requirement of projectivity, or at least prohibit bicipitality (double headedness) and interdependency that is responsible for the myth that DGs and PSGs are “notational variants”.

27 There are possible exceptions, such as complex Semitic-type nonconcatenative morphology, where morphemes are intercalated. But all that shows is that constituency (understood as based on part-whole relationships rather than member-sequence relationships) is not the proper basis for analysing such phenomena.
collection of a finger, a nose and a buttock would not be conceptualized as a single body part. Continuity of wholes must ultimately be stipulated, but only at a broader level of generality than syntax. Thus, if one wished to permit discontinuity (as, for example, McCawley (1982, 1989) does), then either one must stipulate construction-specific possibility of discontinuity, or one must reinterpret constituency as based not on the part—whole relationship but instead on the member—sequence relationship and accordingly be obliged to stipulate default continuity. Without recourse to constituency, continuity of projective dependency structure must be stipulated. This can be confirmed by considering extralinguistic analogues. Imagine a gathering of men, where each man is either father or son of another man present. There would therefore be a non-looping convergenceless structure of relationships, but it would certainly not necessarily be the case that they form a continuous structure such that each man is physically separated from his father by none but his brothers or his sons.

A disadvantage of constituency is that it increases the number of nodes present in the structure of a sentence, but at least the rules in (13a—b) only double the number of nodes, by allowing exactly one phrase per word, and these phrasal nodes bear no features — they are classified as nothing but phrases, not as noun phrases or suchlike. Furthermore, our goal is not to make the representations of the syntactic structure of sentences maximally simple: rather, our goal is to make the grammar maximally simple, though it's an added benefit if sentences come out simple too. We seek a simple grammar wholly or partly because in general the simpler it is the more learnable, or knowable, it is.28 If we seek simple sentences it is presumably because we believe it will be the more easily parsable, or processable, but work such as Groefsema (1992) suggests that syntax isn't all that important for comprehension, and in production it seems quite plausible that the simplicity of the grammar is at least as important as, if not more important than, the simplicity of the sentences themselves. So all in all I do not give much weight to this objection to constituency.

Rules (13a—b) are the only constraints on what constituent structures are permitted (though see also §3.2). The constituent structures we end up getting are of a very skeletal or minimal kind. Phrases have no features. Every word projects exactly one phrase. Rules (13a—b) are the only rules that refer to phrases, and but the merest handful (all of them mentioned in this chapter) refer to nieces. It is the interaction with these rules

28 That at least is ostensibly our reason. Issues of aesthetics, such as the beauty and pleasure of the analysis are also important.
with the rest of the grammar that generates the proper constituent structure. For example, the sequence of words in *mouldy cheese reeks* could, by (13a–b), get any of the bracketings in (14a–e), but (14b–e) get ruled out because in each structure there is a word that isn’t a dependent of its aunt (*mouldy* isn’t a dependent of *reeks*, *reeks* isn’t a dependent of *mouldy* or *cheese*). As already stated, every niece of W must be s-dependent of and dependent of W.

(14) a. [[[mouldy] cheese] reeks]  
b. [[mouldy] [cheese] reeks]  
c. [mouldy [[cheese] reeks]]  
d. [mouldy [cheese] [reeks]]  
e. [[[mouldy] cheese [reeks]]

### 2.2.2 Branch-dependency

In this section we will see how the skeletal tree — i.e. the rudimentary constituent structure — corresponds to dependency structure. As a start, in (15) we define ‘branch-dependency’, a niece relationship defined ‘configurationally’ over a constituency structure. ‘Skeletal branch’ and ‘Branch-dependency’ are synonymous, as are ‘skeletal tree’, ‘branch-dependency structure’, ‘constituency structure’ and ‘phrase structure’. Branch-dependencies hold between two words/nodes linked by a branch of the skeletal tree; the lower word/node is branch-dependent of the higher.

(15) X is branch-dependent of Y iff X is niece of Y and X is instance of Word and Y is instance of Word.

S-dependency mediates between branch-dependency structure and dependency structure. The exceptionless rules (16a–c) define the basic correspondence between branch-dependency structure (i.e. the skeletal tree), s-dependency structure and dependency structure: dependency structure contains s-dependency structure, and s-dependency structure contains branch-dependency structure.
without exception, if X is 'branch-dependent' of Y then X is s-dependent of Y.

b. Without exception, if X is s-dependent of Y then X is dependent of Y.
c. Therefore: without exception, if X is 'branch-dependent' of Y then X is dependent of Y.

The default rules are (17a—c), which require dependency structure, s-dependency structure and branch-dependency structure to be the same. §2.5 asks whether we need anything in addition to (17a) to govern the correspondence between dependency structure and s-dependency structure, and concludes that we don't.

a. By default, if X is dependent of Y then X is s-dependent of Y.
b. By default, if X is s-dependent of Y then X is 'branch-dependent' of Y.
c. Therefore: by default, if X is dependent of Y then X is 'branch-dependent' of Y.

The exceptions to (17a—b) are construction-specific and licensed by rule. The exceptions to (17b) are discussed in §2.4, and are stated there in just one rule. S-dependency is characterized on the one hand by (16—17) and on the other hand by the properties outlined in §2.3. (At this point it might seem to the reader that s-dependency could be eliminated from (16—17) and from the theory entirely, with the properties from §2.3 being attributed to branch-dependency rather than to s-dependency, but in §2.4 the necessity of s-dependency becomes clear when we see that there are s-dependencies that have the properties from §2.3 but that aren't branch-dependencies.)

The exceptions to (17a) are greater in number, though still few. They allow dependencies that exceptionally aren't s-dependencies to occur in constructions involving adjuncts, as in (18a—b), raising, (18c—d), passive, (18e), extraction, (18f), and extrapos-
Dependencies that correspond to branch-dependencies/s-dependencies are capitalized and dependencies that don’t correspond to s-dependencies are uncapitalized. The constituency structures are shown by stemma and by bracketing.

(18) a. [[[Smiling] students] waved].

b. [Anyone [smiling]] will be reprimanded.

c. [[It] keeps [raining]].

d. [Make [it] [rain]].

e. [[Doors] were [marked [on [the [top [of [the [left [of]]]]]]]]].

f. [[Sordello], [he] knows [[she] read]].

---

30 This list is not necessarily complete, though for English there aren’t any obvious additions to be made. A possible candidate is preposed adjuncts, as in On Tuesday she may have been visiting him. Outside English, there are further constructions, such as clitic climbing, as in Italian Io lo voglio leggere (“I want to read it”), and, in partial VP-fronting in German, what Hudson (1995b) has called ‘universal raising’.
2.2.3 The role and status of skeletal trees in grammar

Hudson (1994b) takes the view that projectivity is effected by processing. The processor seeks dependencies, and happens to work in such a way that it can find them only if the dependency structure contains a skeletal tree. For Hudson, the grammar knows nothing of skeletal trees or s-dependency. For Rosta (1994), the grammar does know about what amount to branch-dependencies, and the processor seeks branch-dependencies and, it is asserted, also works in such a way that it can find them only if the dependency structure contains a well-formed skeletal tree. Sentences like (8a–b)/(10a–b) would on this view be grammatical but unprocessable.

§2.3 provides many reasons the grammar needs to be aware of s-dependency. Thus, Hudson cannot be right that the grammar itself knows only dependency. But the question remains as to whether the grammar knows about branch-dependency. In this section I will argue that whatever the facts of processing are, the grammar very much appears to have knowledge of branch-dependency. Therefore processing, and thus the proposals of Hudson (1994b) and Rosta (1994), turns out to in this matter be irrelevant to the grammar; branch-dependency is not an extragrammatical epiphenomenon of processing.

Skeletal trees do indeed appear to play an important role in processing. Whether or not the processor seeks skeletal trees, it certainly seems to find them. Hudson (1995c) has shown that two primary determinants of the processing difficulty of a sentence are (i) the number of words that a branch passes above (Hudson’s ‘dependency distance’), and (ii) the number of branches that pass above a word (Hudson’s ‘dependency density’). The more words that a branch crosses, and the more branches that cross over a word, the harder the sentence is to process. (So, whether or not skeletal trees figure in the grammar, they do matter to usage — to matters of style and processing.)

From the point of view of grammar, though, the issue of whether projectivity can arise from processing is relevant only if it can be shown that the grammar need have no knowledge either of skeletal trees in general or of the requirement of continuity in particular. Let us suppose that processing could give us continuity. Does the grammar
nonetheless need to know about branch-dependency? My answer to this is Yes, and the reason for this answer is that it is possible to state in a single rule exactly when an s-dependent of $W$ needn’t be branch-dependent of $W$. (This rule is given at the end of §2.4.) I do not see how the exemptions from the rule that s-dependents of $W$ be branch-dependents of $W$ could be so simply statable and tightly constrained and contingent on specific associacies (Legate, in particular — see §2.4) if branch-dependency were a mere epiphenomenon of processing.

There are only two other areas I’m aware of where rules make any reference to skeletal trees (not counting be rules defining geometric properties of skeletal trees (some or all of (9a—d)). First, it appears that to some significant extent it is the skeletal tree that determines intonation phrasing (cf. Taglicht 1994, 1995), but the nature of these formal rules for this is not clear. Second, it may be that skeletal branches figure in rules for the form of instances of the lexemes ME, US, HIM, HER and THEM. For example, instances of the lexeme ME, must be pronounced I when sole subject of a finite: I will, *ME will. But when the me is one of two or more subjects, it can have the form me: Me and him will, Him and me will. There are several ways to formally describe these facts. For example, the key factor could be whether the pronoun is a dependent of a conjunction: the rule would be that by default every instance of ME must be pronounced I, but, exceptionally, not every me that is dependent of a conjunction need be pronounced I. But there is a much tidier description available, so long as we can refer to dependencies that are branches in the skeletal tree, and so long as the skeletal tree shown in (19) obtains (as indeed I assume it does — see §2.4.1 and Chapter 3): a ME can be pronounced I when subject of a finite, and must be pronounced I when it is subject of and branch-dependent of a finite.

(19) Me and him will.

But it seems strange that there should be just this one small construction that is sensitive to branch-dependency, so in the end it might be cleaner overall to go for the analysis where the morphophonological form of the personal pronoun is sensitive to whether the

---

31 See Hudson (1992, 1995a) on these lexemes and English caselessness.
32 There is of course tons of idiolectal and dialectal variation in the grammar of pronoun forms. I’m here reporting a bit of my own lect, which is not unusual in this respect.
pronoun depends on a conjunction, especially given the mess that the rules governing pronoun forms are synchronically apparently in.

In summary, if we set aside the formal rules of intonation phrasing as too ill-understood, then the rules we’re left with that invoke the notion ‘branch-dependent’ are the rules of correspondence between s-dependencies and skeletal trees, and perhaps the rules pertaining to pronoun forms. These need to refer to skeletal branches, but not to other properties of the tree. So the grammar knows about skeletal branches, but needn’t know about some or all of the skeletal tree properties in (9a—d).

This means that in principle the grammar can be left ignorant of (9a—d) provided that for sentences to be processable, they must conform to (9a—d). Nonetheless, for several reasons I prefer that the grammar should know of skeletal trees. My principal reason is that if we have the rudimentary constituent structure that I’ve proposed, then from this we derive both the skeletal tree properties and branch-dependency. But without constituency we must have branch-dependency as a primitive. Thus, getting rid of constituency does not constitute a real simplification of the grammar, since we’d need branch-dependency as a primitive instead. Then, on grounds of plausibility, the constituency system is to be preferred over the constituency-free, branch-dependency-as-primitive system, for why should the grammar know of skeletal branches but not skeletal trees? Apart from constituting skeletal trees, the only role for skeletal branches

---

33 Some subsidiary reasons are the following.
- First, for standard examples of unprocessable but grammatical sentences, e.g. self-embeddings, we have the intuition that they would be acceptable if only our brains were nimbler. But for nonprojective sentences we don’t have such intuitions.
- Second, if the grammar is defined as consisting of rules determining what are and aren’t potential sentences of English, then we want the grammar to know about skeletal trees.
- Third, relying on a processing explanation commits us to a view of language (or at any rate grammar) as inherently neurologically implemented. If one takes the view that grammar in the broad sense, just alluded to, of specification of potential utterances may be neurologically implemented but not inherently, then grammar must know of skeletal trees. (This view that language is not necessarily mental, even though it indubitably is mental, would be vindicated if (but not only if) (a) it proves to be possible to accurately model a grammar as a set of rules, but (b) nothing remotely resembling these rules exists in the mind or brain. (An analogous circumstance would arise if it were found that nothing resembling the rules of chess exist in the minds/brains of chess players; despite this finding, the rules of chess would still have a virtual existence.) This is more an ideological than an empirical question: is the goal of cognitive science, including theoretical linguistics, to discover what, if anything, we know, or is it to model what we intuit that we know? If faced with this choice, most linguists would claim they would choose the former goal, but I would choose the latter. In practise, of course, most linguists assume the goals amount to the same thing and proceed to pursue the latter.)
- Fourth, relying on a processing explanation commits us to a particular theory of processing (i.e. one that gives us the effect of skeletal trees). We’d be giving a hostage to fortune: if processing turns out not to work in the way the theory of grammar requires it to, the theory of grammar collapses. Naturally one would rather the theory of grammar didn’t depend too much on assumptions about phenomena outside the scope of the theory.

[41]
is in determining pronoun forms. It would be most strange for a key function of the processor to be to seek out certain relationships whose only job is to in certain lects occasionally play a marginal rule in determining the morphological form of some instances of five lexemes.

To summarize my position, processing appears to be particularly sensitive to skeletal constituency. But nonetheless the grammar knows about s-dependency, skeletal constituency, and the rules that relate the two. However, in all other respects the grammar is oblivious to skeletal constituency, and it is dependency, s-dependency and other kinds of associacy that syntactic rules are sensitive to and syntactic constructions are built from.

2.2.4 Previous approaches to nonprojective dependency structures

Faced with constructions like (18a–g), given above, where the dependency structure does not conform to projectivity, there are various approaches one can take. At the most extreme, one could abandon it altogether, but to my knowledge, the only person who has advocated abandoning it altogether is me (DG e-list, 1993), and since the present discussion constitutes both a recantation and refutation of those views, there’s no point in presenting them here. Slightly less extreme is Mel’čuk (1988)’s treatment of projectivity. Mel’čuk keeps nonconvergence as an absolute requirement, but continuity — the ban against crossing — is merely the unmarked condition of syntactic structures, and violations are permitted in constructions that are “marked emphatically, stylistically, communicatively” (1988: 36). He is therefore unable to explain why all or most discontinuous orderings of words in a given sentence are utterly and irredeemably unacceptable rather than merely slightly marked in the way (18a–g) are. Another method of coping with (18a–g) is a Transformational approach: allow multiple transformationally related strata at each of which projectivity holds.

Hudson’s solution in WG (prior to Hudson (1994b)) has been to relax the requirements of the projectivity principle to the more laxly formulated ‘Adjacency Principle’.34 The unicipitality requirement (no convergence) is dropped, and the no tangling requirement is accordingly revised in Hudson (1990) to (20).

---

34 Hudson follows Robinson (1970) in his use of this term rather than ‘projectivity’.
D is adjacent to H provided that every word between D and H is a subordinate either of H, or of a mutual head of D and H.\textsuperscript{35}

But this runs into problems. It is hard to formulate a weakened version of the principle that allows the tangled dependency structures in (18a—e) yet still rules out projectivity violations like (8b—c). Violations of (20) are not hard to find. (21) is an example.

\begin{center}
\begin{tabular}{c}
\text{e} \\
\text{of} \\
\text{e} \\
\text{of} \\
\text{.} \\
\text{s} \\
\text{s} \\
\text{of} \\
\text{of} \\
\text{of} \\
\text{of} \\
\text{of} \\
\text{of} \\
\text{x} \\
\text{x} \\
\end{tabular}
\end{center}

Choc{a}late, she loves to eat.

Each line in the graphs indicates what is, in Hudson’s analysis, a dependency. In this example \textit{loves} is between \textit{chocolate} and \textit{eat} (a head of \textit{chocolate}), but it is neither a subordinate of \textit{eat} nor a subordinate of a mutual head of \textit{chocolate} and \textit{eat}. In the light of (21), one might consider relaxing the principle to (22).

D is adjacent to H provided that every word between D and H is a subordinate either of H, or of a mutual superordinate of D and H.

But this would fail to prohibit (23).

\begin{center}
\begin{tabular}{c}
\text{s} \\
\text{of} \\
\text{of} \\
\text{of} \\
\text{of} \\
\text{c} \\
\text{c} \\
\end{tabular}
\end{center}

* All the went people.

\textit{The} and \textit{people} are separated by a mutual superordinate, \textit{went}, so the ungrammatical word order is allowed by (22).

\subsection{2.3 S-dependency}

The function of s-dependency that we’ve seen so far is that it plays a crucial mediating role in effecting projectivity — it determines constituent structure. This has been described in §2.2. This section outlines further functions of s-dependency.

\textsuperscript{35} A subordinate of W is a dependent of W or a subordinate of a dependent of W.
§2.3.1 describes s-dependency's function in linear precedence rules.

§2.3.2 explains the role of s-dependency in limiting which pairs of words a dependency can potentially hold between.

A further function of s-dependency is that it constitutes a grammatical phenomenon, ‘remaining in situ’, on a par with extraction, extraposition, promotion, and so on. If X is dependent of Y, then potentially there are a range of further relationships that may also hold between X and Y, and in general one of these relationships must apply. Simplifying, the options are as follows. X can also be extractee of Y — i.e. it is extracted. Or X can be extraposee of Y — i.e. it is extraposed. Or X can also be ‘promotee’ of Y — i.e. it is raised or passivized. Or, if none of these other options apply, X can be s-dependent of Y. Depending on the word class of X and Y, and the type of dependency that holds between them, only a subset of these options are available. For example, if X is adjunct of Y, and Y is a verb, then X will be extractee or s-dependent of Y, while if Y is not a verb, X will be its s-dependent. §2.3.3 is a brief case study of the options available to various clausal dependents.

§2.3.4 applies the same mechanism to a statement of the rules determining when subjects in finite clauses can extract.

§2.3.5 motivates the relationship ‘Tail’: tails are s-dependents with special word order properties, and some words appear to subcategorize for them.

§2.3.6 discusses the role of s-dependency in statements of rules concerning where extracted words end up slotting into the skeletal tree.

§2.3.7 shows, in the context of a constraint on prepositional passives, how s-dependency allows us to capture notions that we might describe, in transformational metaphor, as ‘movement must be upwards’.

2.3.1 Linear precedence

The sorts of word order restrictions that continuity accounts for are holistic in nature. Even though syntax consists mostly of pairwise relationships between words, no sentence that is unicipital (i.e. in which every word but one has exactly one head) can be discontinuous solely because of the relative position of a head—dependent pair. In a unicipital structure, discontinuity is due to the relative position of at least two head—dependent pairs. This is of course not all there is to the grammar of word order. Besides continuity, we need rules governing the relative order of codependents and
rules governing the relative order of head—dependent pairs. Examples of codependents
whose order is fixed are the italicized words in (24—25) (on examples like (25), cf.
Crystal (1971: 126 ff.)).

(24)  a. He bet *her five pounds that it would.
     b. *He bet her five pounds that it would.

(25)  a. all the very same *favourite great big old red Dutch cardboard picture
        books
     b. *all the very same Dutch red old big great favourite picture cardboard
        books
     c. *all the very same cardboard picture favourite great big old red Dutch books

It remains a matter of some uncertainty what sort of rules we need in order to effect
these orderings, and whether we need to stipulate the orderings explicitly. The other
sort of ordering rules, pertaining to head—dependent pairs, will be the focus of this sec­
tion.

The grammaticality difference between (26a—b) cannot be explained in terms of
continuity — both are continuous, and it makes no difference to the structure’s continu­
ity whether stunningly precedes or follows beautiful.

(26)  a. the stunningly beautiful picture
     b. *the beautiful stunningly picture

In order to account for the ungrammaticality of (26b), we need a rule requiring stun­ningly to precede beautiful, rather than vice versa. There are many such rules in
Hudson (1990), stating, for example, that a complement of W follows W, a subject of
W precedes W, and so on.

The problem with such rules is that they clash. Consider various exceptions, as in
(27a—c), to the generalization that the subject of W precedes W. (27a) obeys the rule
that attributive adjectives precede the noun; (27b—c) obey the rule that complements of
W follow W.

(27)  a. merry people
In WG hitherto, one would have had to override the default subject order rule for each of these constructions. That strategy amounts merely to stipulating the ordering for every possible construction. But note that these exceptional subjects in (27a–c) are also not s-dependents of the word they’re subject of. This is the key to the problem.

S-dependency, I maintain, is crucially involved in the grammar of head-dependent orderings. The grammar is sensitive to the relative order of two words only if one is s-dependent of the other. This is stated in (28), using ‘Precessor’ and ‘Successor’ to denote relationships of precedence and succession. (Note that Precessor and Successor are not converses: it is not possible for X to be precessor of Y and Y to be successor of X, since that would require each to be s-dependent of the other, which is prohibited by constraints on s-dependency loops stated in §2.4.1.)

(28) Without exception, X is s-dependent of Y iff X is precessor of Y or X is successor of Y.

The intention is that Precessor and Successor are the only relationships of sequence that the grammar has access to, at least in so far as their arguments are words.

---

36 I am aware of two exceptions to this, one bona fide and the other suspicious. The bona fide exception is that words can be coenounced (see Chapters 5–6) with their ‘preneighbour’ or ‘postneighbour’ (the immediately preceding or following word), which needn’t be their s-dependent or s-head. An example is *the king of Bulgaria’s nose* where the possessive pronoun is coenounced with its prneighbour, *Bulgaria*, even though neither is s-dependent of the other. The suspicious exception is that no word can be separated from any of its conjunct-mates (see Chapter 3) by anything other than another of its conjunct-mates. In other words, conjuncts must be continuous. The formulation I give is rather counterintuitive, and if there were a means of having headless nontermal nodes, conjuncts would be the prime candidate for such an analysis.

37 This conflicts with the WG view that grammar is part of general knowledge, since general knowledge surely can tell whether one word follows or precedes another, even when neither is s-dependent of the other. However, that view is more a null hypothesis than a central tenet of the theory, and to the extent that my contentions about word order hold good, the null hypothesis is thereby disproved: the grammar appears at least to some extent to be informationally encapsulated. But anyway, virtually all aspects of WG syntax call that null hypothesis into question: for example, if the grammar has access to the full panoply of intragrammatical and extragrammatical knowledge, why does no word subcategorize for a complement that is slang, or that starts with a B? — Presumably
(27a–c) turn out not to be exceptions to the word order rules pertaining to subjects, since the actual rule is (29): the reason (27a–c) are not exceptions is that although *people, the* and *everyone* are subject of, respectively, *merry, on* and *happy*, they are not s-dependent of the word that they're subject of.

(29) If X is subject of Y and X is s-dependent of Y then X is precessor of Y.

The word order in (27a–c) will therefore follow from the rules that pertain to the s-dependencies in these sentences; in (27a), for instance, *merry* precedes *people* because *merry*, as an attributive adjunct and s-dependent of *people*, is precessor of *people*.

Thus, under my proposals the ungrammaticality of *Smiling him, (she entered)* and *Entered she* is due to a subject of W being a successor of W rather than a precessor, as in *Him smiling, (she entered)* and *She entered*. Hudson attempts to account for this data by saying that the subject of W must precede W, but in *Smiling, she entered*, the subject of *smiling* (i.e. *she*) does not precede *smiling*. So Hudson would be forced into some cumbersome rule saying something like that the subject of W precedes W unless the subject is also subject of a word that W is adjunct of, with sundry other such rules for the many other constructions where the subject of W does not precede W.

I am not aware of any exception to (29), save perhaps for *the more fool she* and a handful of similar examples. There is not, however, a general presumption that such rules will be exceptionless. Indeed, by stating (30) as a default rule and accepting that there will be exceptions to it, we actually end up requiring very very few word order rules.

(30) By default, if X is s-dependent of Y then X is successor of Y.

The only word order rules (for head—dependent pairs) we'll need are those that are exceptions to (30) (or those that are exceptions to the exceptions, if there are any). We therefore need just a small handful of rules to cover orderings where dependents are precessors: subjects, extractees, preposed adjuncts, the relative sequence of *the* and 's in *the boy’s*, and the position of ‘quantitatives’ (a.k.a. ‘measure modifiers’) of non-

---

because syntax has no access to stylistic or phonological information. The grammar really does look like it is modular.

38 In inversions like *Did she* and *Am I*, I don’t take *she* and *I* to be subjects. See next footnote.
verbs, as in *three miles wide, three feet under, three years ago*. A further consequence of (30) is that virtually all clashes between word order requirements are resolved automatically by default overriding (i.e. the elsewhere principle). For example, in *be here, here* follows *be* because it is s-dependent and therefore, by default, also successor of *be*, while in *here am I* the rule requiring an extractee and s-dependent of *W* to be precessor of *W* automatically overrides (30) by virtue of its conditions being a subset of (30)'s. Any need for further explicit rules to resolve conflicts between word-order rules would arise only in a construction in which an s-dependent that is normally a precessor is actually a successor. At present I'm not aware of any clear examples of that happening.\(^{39}\)

2.3.2 No promiscuous dependency

What prohibits the asterisked dependencies in (31a–b)?\(^{40}\) (I take it for granted that prohibit them we indeed wish to do.)

(31) a. The naked minstrel endeavoured to play a saraband.

b. The minstrel endeavoured to play a saraband naked.

To the rescue comes rule (17a), which says that by default, every dependent of *W* must be s-dependent of *W*. Exceptions are permitted for certain constructions, such as extraction (for examples, see §2.3.3–4), but there are no rules stating exceptions that would

\(^{39}\) So-called subject-auxiliary inversion is the absence of object-to-subject raising, not postponing of subjects or preposing of auxiliaries. The structure is as in (i) (syntactic words in capitals, morphological in italics).

\[\text{DO HER SMILE} \quad \text{Did she smile?}\]

For locative inversion, the structure is as in (ii).

\[\text{HERE DO COME SOPHY} \quad \text{Here comes Sophy.}\]

In neither of (i–ii) is there a word whose subject is its successor. There is insufficient space in this thesis to justify (i–ii), but it is fairly clear how they might be plausible. The only really controversial element is the mismatch between syntactic words and morphological words, and this is justified at length in Chapters 5–6.

\(^{40}\) This question was brought to my attention by Hazel Clark.
allow examples like (31a—b). The effect of (17a) is to prevent dependencies from rambling freely throughout the sentence, engineering liaisons between any old pair of words. Instead, all dependencies must be s-dependencies (and therefore, by default, branch-dependencies) unless permission is expressly granted for a dependency to be unaccompanied by an s-dependency.

2.3.3 S-dependency and clausals

The grammar contains rules like (32a—c), which specify requisite and permissible combinations of the different kinds of interword associacy. These are the kinds of rules that state exceptions to the default rule given above in (17a), which states that dependents of W are s-dependents of W.

(32)  a. If X is object of Y then X is s-dependent of Y or X is extractee of Y.\(^{41}\)
    b. If X is clausal of Y and Y is instance of Verb then X is s-dependent of Y.
    c. If X is clausal of Y and Y is instance of Preposition then X is extractee of Y.

(32a) allows for both (33a—b); in (33a) it is object and s-dependent of read and in (33b) it is object and extractee of read.

(33)  a. She endeavoured to read it.
    b. What did she endeavour to read?

(32b) allows only for (34a—b), in which that is clausal and s-dependent of convince and inform, and not for (35a—b), in which that is clausal and extractee of convince and inform.\(^{42}\)

---

\(^{41}\) This rule is somewhat simplified.

\(^{42}\) On these judgements, see below.
(34) a. She managed to convince him that she loved him.
   b. She sought to casually inform him that she loved him.

(35) a. * That she loved him, she managed to convince him.
   b. * That she loved him, she sought to casually inform him.

(36) a. That she loved him, she managed to convince him of.
   b. That she loved him, she sought to casually inform him of.

(37) a. * Of that she loved him, she managed to convince him.
   b. * Of that she loved him, she sought to casually inform him.

(38) a. * She managed to convince him of that she loved him.
   b. * She sought to casually inform him of that she loved him.

(Note, incidentally, that (32a–c) could all be learnt from positive evidence if the learner adopts a conservative strategy that avoids assuming generalizations for which there is not overwhelming evidence.)

For (35a–b) it is the judgements of Kaplan & Bresnan (1982: 242) and Postal (1994: 70) that I am reporting. I actually find (35a–b) reasonably okay, and find (39)
impeccable. So in my lect, clausal of verbs can be either s-dependent or extractee of the verb.

(39) That she had cheated, she vehemently denied __.

Kaplan & Bresnan’s example is (40), which I find much odder than (39). Perhaps this is because when interpreting (40) (which would be still be grammatical) we cannot help gardenpathing by considering a parse in which think is detransitive. This might also apply to judgements of (35a), since She convinced him is okay; but since She informed him is less normal, one wouldn’t expect it to apply to judgements of (35b).

(40) “*” That he might be wrong he didn’t think.

It may be relevant (though I’m not sure exactly how) that Postal (1994: 109) also judges (41a) to be ungrammatical, whereas I find it reasonably okay, and find the less gardenpathy (41b) to be wholly unexceptionable.

(41) a. That traces are lexical is held by many scholars.
    b. That traces are lexical is asseverated by many scholars.

If (39) were grammatical and (40) ungrammatical then perhaps an explanation could be developed whereby the ability for a clausal to be absent and its ability to be an extractee are mutually exclusive. However, if (40) is ungrammatical then some way must be found to account for (42).\(^{44}\)

(42) He might, he thinks, be wrong.

There are a number of exceptions to (32b–c). One is (43).

(43) a. If X is clausal of Y and Y is instance of Verb and X is instance of As then X is extractee of Y and Y is subordinate of X.

\(^{44}\) Postal (1994) does indeed suggest an explanation for this. Conventional that-clause extraction leaves an NP trace but parenthetical extraction leaves a non-NP trace. The position of complement of preposition is an NP position, and the slot occupied by a non-extracted clausal complement is not an NP position. As this won’t translate into the present model, I won’t consider it further.
This allows (44a) but not (44b–c). ('Ap' denotes appositive adjuncts; dependencies paralleled by s-dependencies are capitalized.)

(44)  
a. She loved him, as he had suspected.

b. * She loved him he had suspected as.

c. * He had suspected as she loved him.

The contrast between (45a–c) and (46a–b) is explained by a rule along the lines of (47) along with a revision of (32c) to one of (48a–b).45

(45)  
a. She loved him, as he had convinced her (*of).

b. She loved him, as she had informed him (*of).

c. She loved him, as he had thought (*of).

(46)  
a. She loved him, as had often been thought (of).

b. She loved him, as has often been spoken about.

(47)  
Every as is complement of a verb.

(48)  
a. If X is clausal of Y and Y is instance of Preposition then X is not s-dependent of Y.

b. If X is clausal of Y and Y is instance of Preposition then X is extractee of Y or X is 'promotee' of Y.

---

45 I assume that prepositional passives don't involve extraction, e.g. because they don't 'license' parasitic gaps.

(i) * The painting has been gazed at ___ without us getting bored of ___.
As can be ascertained from the diagrams, in (44a) and (46a–b) as is complement of a verb, in conformity with (47), while in (45a–c) if as is complement of of then, in violation of (47), it would be extractee but not complement of the verbs in the subordinate clause.

Another exception to (32b–c) involves parenthetical clauses. As Postal (1994) observes, these give inverse grammaticality results to (34–38), as can be seen from (49–50). (‘P’ denotes a parenthetical adjunct.)

(49)  
\[
\begin{array}{c}
\text{OF} \\
\text{of} \\
\text{of} \\
\text{of} \\
\text{of} \\
\text{of k}
\end{array}
\]

a. She, he had convinced her, loved him dearly.

b. She, she had informed him, loved him dearly.

(50)  
\[
\begin{array}{c}
\text{OF} \\
\text{of} \\
\text{of} \\
\text{of} \\
\text{of of e} \\
\text{of k}
\end{array}
\]

a. * She, he had convinced her of, loved him dearly.

b. * She, she had informed him of, loved him dearly.

Given the grammaticality of (49a–b) we must override (32b) for parentheticals, by means of (51). But note that the necessity of (32b) is contingent on the ungrammaticality of (35a–b), which I’ve disputed. If I am right to suspect that (35a–b) are in actual fact grammatical for all speakers then both (32b) and (51) are unnecessary.

(51)  
\[
\text{If X is clausal of Y and Y is instance of Verb then X is s-dependent of Y or [X is extractee of Y and Y is unsuperordinate of Z and Z is parenthetical of X].}^{46}
\]

As for (50a–b), these are ruled out by rule (52), which is somewhat like (47).

---

46 X is unsuperordinate of Y iff X is subordinate of Y or X is Y. X is subordinate of Y if X is s-dependent of Y, or X is subordinate of Z and Z is s-dependent of Y.
(52) a. If X is parenthetical of Y then Y is extractee of X and Y is clausal of Z and Z is instance of Verb (and Z is subordinate of X).

These extractions are unusual in that the extractee can actually end up to the right of the words it is extractee of, as in (49a—b). This is possible because (53), the general rule which gives extractees their position, does not apply to parentheticals, for the extractee of the parenthetical is not s-dependent of any of the words it is extractee of.

(53) If X is extractee of Y and X is s-dependent of Y then X is precessor of Y.

2.3.4 Extractability of subjects

Given the alternation know it is/know that it is, it is tempting to offer an analysis in which the complement of know must be a finite and that is an optional dependent of a finite that is a complement. (54) represents this.

(54) 

However, such an analysis not only fails to explain why that must precede all other elements in the clause: it also fails to provide a basis for explaining the well-known contrast (55a—b).

(55) a. Who do you reckon went?
   b. * Who do you reckon that went?

The basis for developing an explanation for this data is to follow Hudson (1984, 1990) in taking that in (54) to be complement of know, and is to be complement of that, as shown in (56).

(56) 

According to Hudson's analysis of extraction, a word can only be extractable by virtue of its subjecthood if it is the subject of a complement of a verb. In (55a) who is a
subject of a complement (went) of a verb (reckon), so who is extractable. Assuming the
dependency analysis in (56), in (55b) who is a subject of a complement (went) of a non-
verb (that), so it is not extractable.

Hudson's explanation is in fact correct in its essence but wrong in its details. In
cases where the subject in a finite clause is also a complement of a word within the
clause Hudson wrongly (and inadvertently) predicts that the subject should extract. For
example, she in He knows that she will is both subject and object of will and by virtue
of being an object it should be able to extract. Even if we restrict ourselves to comple-
mentation patterns countenanced by Hudson (1990) we have examples like He knows
that she was seen, where she is object of seen and so should be extractable.

The solution to this problen follows from rules such as (32a—c). Now, what are the
rules of this sort that apply to subjects? The default rule is (57a), but it is overridden by
the more specific rule (57b) (which is more specific because finites that are comple­
ments of verbs are a subset of finites in general).

\[(57)\]

\[\begin{align*}
\text{a. } & \text{If } X \text{ is subject of } Y \text{ and } Y \text{ is instance of Finite then } X \text{ is } s\text{-dependent of } Y. \\
\text{b. } & \text{If } X \text{ is subject of } Y \text{ and } Y \text{ is instance of Finite and } Y \text{ is complement of } Z \\
& \text{and } Z \text{ is instance of Verb then } X \text{ is } s\text{-dependent of } Y \text{ or } X \text{ is extractee of } Y.
\end{align*}\]

(57b) allows Who do you doubt went, because went is complement of a verb, doubt, so
its subject can extract, but does not allow Who do you doubt that went, because if we
assume the dependency structure in (56) then here went is not complement of a verb so
its subject must be its s-dependent.

2.3.5 Tails

There seems to be a slight correlation between word order and extractability: certain
unexpectedly unextractable words are required to be the nearest complement of the
word they’re complement of. Examples are:

\[(58)\]

\[\begin{align*}
\text{a. } & \text{She read the book.} \\
\text{b. } & \text{She read Sophy 's book.}
\end{align*}\]
c. She wished of c him to go.

d. She gave of i him flowers

None of these complements extract: *Book she read the, *Sophy she read's book, *Him, she wished for to go, %Him, she gave flowers. Another thing they have in common is that no other dependent of the word they're complement of can intervene between them: for example, in (58d) no other dependent of gave could come between gave and him, as (59a—c) show (cf. Hudson 1992).

(59) a. * She gave flowers everyone who'd ever craved a bunch of tulips.

b. * She gave yesterday everyone attending the lecture an impromptu disquisition on transitivity.

c. % She gave back Sophy the draft.

Suppose that these complements are examples of a relation, 'tail', such that the tail of W (i) must be complement of W, (ii) must be s-dependent of W, and (iii) mustn't be separated from W by anything other than an s-dependent of the tail. Though Hudson's (1988b, 1990) analysis of extraction predicts that these complements ought to extract, if they are also 'tails' then extraction will result in a nonprojective s-dependency structure, as shown in (60a—d). The structures are nonprojective because the tail's s-head is not a superordinate of all the words intervening between the s-head and its tail.

(60) a. * Book, she wrote on the back of the.


c. * Who did she wish for to go?

d. % Who did she give flowers?

Tails would also provide a link between unextraposability and unextractability. Nouns' OF complements, for instance, do not happily extrapose or pied-pipe. If the OF is the
noun's tail, then the extraposition and extraction loses projectivity, as shown in (61a–d).

(61)  

a. * Of whom did she invite a few?  
b. * A few arrived of her friends.  
c. * Of what did she drink a pint?  
d. * A pint was drunk of milk.

For complements of determiners, for partitive of complements of nouns, and for the object of infinitival for, we need a rule requiring these complements to be tails. For injects of ditransitive verbs, we need a rule requiring the inject to be a tail or a promotee of the verb, since although a ditransitive verb's inject can't extract, it can passivize. Given our definition of Tail, this analysis of the constraints on injects both generalizes and remedies Hudson's (1992) proposed rule that injects must be next to the word they're injects of. (The rule is intended to account for data like (59a–c).) As it stands, the rule makes wrong predictions that Hudson did not intend: in She gave Sophy's brother flowers, 's, the inject of gave, is not next to gave, yet the sentence is grammatical.

2.3.6 Landing sites for extraction

Word Grammar has a stepping-stone analysis of extraction. In (62), for example, chocolate is extractee not only of on and of has but also of all the words in the s-dependency chain between them, as shown in the diagram. (The decision to treat on as dependent of cut rather than down is gratuitous.)

---

47 According to Hudson (1994b) only adjuncts extrapose out of NPs, in which case we would of course expect (21b) and (21d) to be ungrammatical. However, certain dependents that are not obviously adjuncts do extrapose:

(i) The realization dawned on her that her hypothesis was wrong.
(ii) A picture was commissioned of each member of the family.
Having *chocolate* be extractee of all these words has various benefits, notably the benefit of accounting for a number of island constraints (see Hudson (1990)), but it runs into a problem: how come *chocolate* can’t come to rest at some point between *has* and *on*? What rules out either of (63a–b) as possible structures?

(62) Chocolate, she has been meaning to try to cut down on.

(63) a. *She has been meaning to chocolate try to cut down on.*
She has been meaning to chocolate try to cut down on.

The solution to this problem is rule (64), which says that if X is extractee of its s-head, its s-head must be finite. The effect of the rule is that an extractee can come to rest at only a subset of the words it can extract across. Note that this rule cannot be formulated without reference to s-dependency.

\[(64) \text{ If } X \text{ is extractee of } Y \text{ and } X \text{ is s-dependent of } Y \text{ then } Y \text{ is instance of Finite.}\]

2.3.7 A constraint on prepositional passives
In this section we see how s-dependency allows us to capture the constraint that a word can promote (NP-move) only upwards.

(65a—b) are of course acceptable, as are, if suitably intoned, their counterparts (66a—b), with the preposition extracted. (67a—b), prepositional passive counterparts of (65a—b) are fine too. But (68a—b), counterparts of (67a—b), with the preposition extracted, are ungrammatical.

(65) a. She slept in the bed.
   b. She wrote on one side of the paper.

---

48 There are exceptions to the rule, e.g. to accommodate Where to?, Who with?, etc. All such exceptions arguably involve 'extraction to subject' – constructions where a word's subject and extractee merge. The more accurate rule would therefore be (i).

(i) If X is extractee of Y then either X is extractee of and s-dependent of a finite or X is extractee of and subject of a word.
(66) a. In the bed, she slept.
b. On one side of the paper she wrote.

(67) a. The bed was slept in.
b. The paper was written on one side of.

(68) a. * In, the bed was slept.
b. * On which side of was the paper written?

How come the complement of the prepositions in (67a–b) can promote, but the complement of the prepositions in (68a–b) can't?

Prepositional passives seem to work by means of a severely constrained recursive mechanism the details and full nature of which remain unclear. But I will sketch an analysis that is probably broadly along the right lines. There seems to be no limit in principle to the length of the promotion path, as (69) suggests, though each step in the promotion path is tightly restricted — it's hard to come up with examples that differ very much from (69). I therefore assume that (69) involves something akin to the partial associacy structure shown.

(69) The box was marked on the inside of the upper half of the left side of.

One of the constraints on 'promotees' is, I suggest, something like (70), the effect of which is to allow promoted complements to proceed only upwards.

(70) If X is promotee of Y then Y is unsuperordinate of Z and Z is instance of Passive.

This has the effect of allowing (71a), because the can get to be promotee of written and once it is promotee of written it can be subject of written. But (71b) is excluded, because although the could get to be promotee of on, it cannot go the further step of be-
coming promotee of, and hence subject of, written, because on is not unsuperordinate of written. (‘Pm’ = ‘promotee’.)

(71) a. The paper was written on one side of.

b. * On which side of was the paper written?

2.4 S-dependency loops

This section explains how some s-dependencies do not form branches of the skeletal tree. This happens in coordination of various sorts and in pied piping. In this section we have an initial look at how this happens in standard coordination. Coordination is treated in greater depth in Chapters 3–4, and the s-dependency loops in pied piping are discussed in §4.2, especially §4.2.3.

Hudson’s (1984, 1988, 1989, 1990) WG analyses of coordination have always taken the view that coordinate structures, and no others, are built by constituency rather than dependency. It would be better if coordination were by dependency, firstly because this would reduce the range of structural devices employed in syntax (from dependency plus constituency, to just dependency), and secondly because, as any moderately comprehensive survey of coordination data will reveal, there is no sharp divide between coordinate and non-coordinate structures — there is a gradient of constructions with incrementally fewer and fewer properties of canonical coordination. Surprisingly, Hudson has never really justified his rejection of coordination by dependency. In Hudson (1984: 212) he says “for simplicity I shall simply assume that they are not dependency structures, and leave it to others to prove the contrary”. Since dependency is in principle preferable to constituency, the job for these others is merely to show that dependency does the job. This I have done (in Chapters 3–4), and in this section I will report on the results. In Hudson (1990: 405) the objection to coordination by dependency is fleshed out
only to “no one word stands out as the root of the whole coordination. Any one of the conjuncts could in general occur in place of the coordination [...] so each of their heads would qualify equally well as the head of the whole coordination. Nor could we take the conjunction AND as the head of the coordination because its distributional properties have nothing at all to do with those of the whole coordination (which follows the ordinary rules for nouns, verbs and adwords rather than those for conjunctions).” It’s curious that he takes this as necessitating a rejection of dependency, since it is an argument not so much against dependency as in favour of dependency loops, which WG, exceptionally among theories of Dependency Grammar, allows.

Considerations of continuity lead us to conclude that the conjunction is the root of the coordination. Conjuncts are branch-dependents of the conjunction. So are words that are dependents of each of the conjuncts: this can be seen from contrasting (72a—b). *Who (was French) is adjunct only of another and is branch-dependent of another, while who (were embracing) is adjunct of both one and another and is branch-dependent of and. If the order is as in (72b), branches cross, and the phrase is ungrammatical.

(72) a.

```
  of
  of another
  of another
  of a
  one who was Japanese and another who was French who were embracing
```

b. *one who was Japanese and another who were embracing who was French

(73) therefore has the skeletal tree that is shown in the diagram. To each of these branches there corresponds an s-dependency and a dependency; these are shown too. Coordinatee dependencies hold between conjuncts and conjunction. An ‘advener’ dependency holds between and and knows, and a ‘ward’ dependency holds between she and and and between apples and and.

(73) He knows she washed, cored and peeled apples.
But we also require the additional s-dependencies shown in (74). She and apples must be s-dependent of washed, cored and peeled because that is what by default they must be, unless they are extracted (or otherwise moved), which they are not, and because she must precede — i.e. be predecessor of — washed, cored and peeled and apples must follow them — it must be their successor.

(74) He knows she washed, cored and peeled apples.

(75) gives most of the associacy and dependency structure and shows the branch-dependency structure by single-lined stemma and the rest of the s-dependency structure by double-lined stemma. As we can see, she, washed, cored, peeled and apples are all s-dependent of more than one word.

(75) He knows she washed, cored and peeled apples.

We can now replace the default rule (17b), repeated here as (76), by the exceptionless (77), which also includes in it the earlier exceptionless (16a).
By default, if $X$ is s-dependent of $Y$ then $X$ is 'branch-dependent' of $Y$.

Without exception, $\forall x \forall y \ x$ is s-dependent of $y$ iff

- $x$ is 'branch-dependent' of $y$, or
- $\exists z \ z$ is legate of $x$ and $z$ is advener of $y$, or
- $\exists z \ z$ is legate of $y$ and $x$ is ward of $z$.

The Legate associacy has the special property of licensing s-dependency loops — it allows s-dependencies to tangle in the ways specified by (77). Any legate can be an advener and have a ward.

2.5 Are there principles governing the correspondence between dependency structure and s-dependency structure?

The short answer is No: the only principle governing the correspondence between dependency structure and s-dependency structure is the one stating that by default if $X$ is dependent of $Y$ then $X$ is s-dependent of $Y$. But nonetheless there are some candidate principles worth considering before we reject them.

There are some generalizations about the correspondence between s-dependency structures and dependency structures or about other global properties of s-dependency structure that appear to apply without exception, without their being stated by rule. At present it remains to some extent unclear (i) whether the generalizations are correct, (ii) whether they needn’t be stated by rule, and (iii) why, if the generalizations are correct but not stated by rule, the generalizations obtain. Further work remains to be done on these questions, but so far it looks as if the candidate principles are either wrong or redundant.

The first generalization is that, other factors permitting, the s-head of $W$ precedes all other heads of $W$. The main kinds of exception to this are structures similar to (78c—d): if $X$ is head of $W$, and $Y$ is s-head of $W$, and $X$ precedes $Y$ (as in (78a—b)), then either $W$ will be superordinate of $X$ (as, for example, in (78c) and (78e)) or $X$ will be a subordinate of $Y$ and $Y$ will precede $W$ (as, for example, in (78d) and (78f)). So, more precisely, the generalization is that either the s-head of $W$ precedes all other heads of $W$, or one of the patterns (78c—d) obtains (where $X$ need only be subordinate and not s-dependent of $W$, in (78c), and of $Y$ in (78d)).

49 'Legate' and Chapters 3—4's 'Supervener' are synonyms.
This generalization is what we get under Hudson’s (1994b) treatment of skeletal dependency, in which the parser, going incrementally left to right, links a word into the skeleton at the earliest opportunity. We can call this The Early Linkage Principle.

Further generalizations emerge from (79a–d), which are taken from Volino (1990) who is quoting Hudson (1986 — an opus no copies of which are known to remain extant).

(79) a. A word may have more than one head.

b. A word may be separated from one of its heads by the other, provided that the head furthest away from it depends on the other head.

c. A word which has two heads need have no semantic relations to the higher of them; i.e. a word takes some or all of its semantic relations from its lowest head.

d. A word takes its position from its highest head.

(79b) is instructive, because it looks right, so long as “depends on” is changed to “is subordinate to” — one thinks of subject-raising constructions. But in fact it is falsified by relatively uncommon constructions, such as locative inversion, where (speaking in conventional metaphors) a word is raised to subject and then extraposed back to where it started, as in In this drawer are going to be kept her socks. Here are, going, to, be and kept are heads of her but the first four aren’t subordinate to kept. (79c) looks true, though is arguably falsified by Remember when we met: the lowest head of when is met but when, being an adjunct, doesn’t get its semantic relations from the word it is an adjunct of.

---

50 This at least is how I analyse locative inversion in unpublished work.
I shall interpret as (80), on the grounds that words take their position from their s-head.

The Highest Head Principle.

No word is a dependent of a superordinate of its s-head; in other words, if X is s-dependent of Y and Y is subordinate of Z, then X is not dependent of Z.

Highest Head and Early Linkage agree in requiring (81a) instead of (81b). Highest Head requires (82a) instead of (82b); Early Linkage would allow either. Highest Head requires (83a) and (84a), while Early Linkage requires (83b) and (84b).

The structures required by Highest Head are intuitively more fitting than those required by Early Linkage, which is evidence against Hudson’s (1994b) method of building the skeleton. However, each of these structures required by Highest Head is anyway required by the default requirement that if X is dependent of Y then X is s-dependent of Y. For (81—82) there is a rule that requires the object of W to be extractee, raisee or s-dependent of W. Since it is neither extracted nor raised, it must be s-dependent. For (83—84) there is a rule that requires the subject of a finite to be its extractee or s-dependent. Since the subject is not extracted, it must be s-dependent of the finite.
2.6 Proxies and other non-dependency associacies

In this chapter so far it's been my goal to motivate the relations of Associacy, Dependency, S-dependency and Branch-dependency. By definition, every grammatically significant syntagmatic relationship between words is an associacy, in the sense that for every such relationship type, R, every R of W is associate of W. Most varieties of associacy are also varieties of dependency, and it is beyond the scope of this thesis to motivate any further elaborations to the taxonomy of dependency types. But it is feasible to note here the few associacy types that are not dependency types. Most are introduced properly in Chapter 3, and Proxy, the remaining associacy type that needs special motivation, is given it in §§2.6.1–4.

- **Subordinate, Superordinate, Unsubordinate, Unsuperordinate.** These are defined in a single statement each:
  - X is subordinate of Y iff X is s-dependent of Y or X is s-dependent of a subordinate of Y.
  - X is superordinate of Y iff X is s-head of Y or X is s-head of a superordinate of Y.
  - X is unsubordinate of Y iff X is superordinate Y or X is Y.
  - X is unsuperordinate of Y iff X is subordinate Y or X is Y.

- **Preneighbour, Postneighbour.** These relate a word to the immediately preceding and immediately following word, irrespective of whether any other associacy holds between them. Preneighbour and Postneighbour play a role only at the syntax—morphology interface, where various rules can permit a word to be pronounced at the same time as its preneighbour or postneighbour. See Chapters 5—6 for more on this (though the discussion there is not formalized enough for Preneighbour and Postneighbour to be used.)

- **Conjunct-mate, Substitute, Legate.** These are associacies involved in coordination and related constructions. See Chapters 3–4. If X is conjunct-mate of Y, then there is a word of which X and Y are both subordinates. In general, the same goes for substitutes. In general (though not in asyndeton), if X is legate of Y then X is superordinate of Y.

---

51 This is not to say that in relevant respects my analyses necessarily follow Hudson (1990). Here and there I depart from his analyses in various ways, but discuss such departures only when they are pertinent.
• **Proxy.** X is proxy of Y when X is unsubordinate of Y and X has distributional properties characteristic of Y. §§2.6.1—4 motivate the Proxy associacy.

### 2.6.1 Proxies of relative and interrogative pronouns

I'll begin with a brief outline of the structure of free (or ‘independent’) relatives, as in (85a—b):

```
(85)  a. She bought what entertained her.
     of O
     . of K
     . s of

   b. she bought what she wished to enjoy.
     of O
     . of K
     e of
```

First, it is the relative pronoun that is dependent of a word outside the relative clause; in these examples it is the object of *bought*. Second, the pronoun has a ‘clausal’ dependent, *entertained* in (85a) and *wished* in (85b). And third, the pronoun is either subject or extractee of its clausal, as in (85a—b) respectively. (The extractee of W is extracted via W.)

---

52 For a discussion of the range of reasons for treating the root verb in the free relative clause as a dependent of the relative pronoun, see Hudson (1990). Other (possible) examples of clausals are given in (i—x).

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(i)</td>
<td>the book she read</td>
<td>(vi)</td>
</tr>
<tr>
<td>(ii)</td>
<td>the book to read</td>
<td>(vii)</td>
</tr>
<tr>
<td>(iii)</td>
<td>the book for her to read</td>
<td>(viii)</td>
</tr>
<tr>
<td>(iv)</td>
<td>the book that she read</td>
<td>(ix)</td>
</tr>
<tr>
<td>(v)</td>
<td>wonder what she read</td>
<td>(x)</td>
</tr>
</tbody>
</table>

I use ‘clausal’ in a different sense from Hudson (1990), who uses it for clausal complements. In the present use of the term, a clausal of W is sometimes also a complement of W, but sometimes is neither complement nor adjunct of W. It is not a typical adjunct, because no word can have more than one clausal (*He bought what entertained him he wished to enjoy*), and sometimes is not a typical complement, because there is no necessary semantic relation between a word and its clausal (e.g. between *what* and *wished* in *He bought what he wished to enjoy*). Hudson (1990) treats the pronoun’s clausal as a complement (of the pronoun) of no specific variety (such as ‘object’), but this is not possible, since in (x) *what* would have two complements (*books* and *entertained*), and if a word can have more than one complement each must be a distinct variety of complement (e.g. ‘indirect object’ vs. ‘direct object’ vs. ‘xcomp’).

(xi) He bought what books entertained him.

If *books* is a complement of *what* and *entertained* is a clausal of *what* but not a complement of *what*, then the general rule that complements precede other dependents explains the ungrammaticality of *He bought what entertained him books.*

[68]
The structure of interrogative and ordinary relative clauses like (86a–b) is like the structure in (85a–b). In (86a) what is the object of wondered; WONDERS subcategorize for interrogative pronoun objects. In (86b) which is an appositive adjunct of book.\footnote{It is appositive because, loosely speaking, the meaning of which is equated with the meaning of book, though it is not 'appositive' in the fairly standard sense defined in Meyer (1992). Interrogative pronouns can also be appositive adjuncts, as in (i), where which is the appositive adjunct of idea.}

(86) a. She wondered what entertained her.

\[
\begin{array}{c}
\text{of} \\
\text{of K} \\
\text{of}
\end{array}
\]

b. She bought the book which she wished to enjoy.

\[
\begin{array}{c}
\text{of A} \\
\text{of K} \\
\text{of}
\end{array}
\]

However, in pied-piping constructions like (87a–b), it is not the wh-pronoun that is subject or extractee of the wh-pronoun’s clausal.

(87) a. He wondered ON what she sat.

\[
\begin{array}{c}
\text{of} \\
\text{of K} \\
\text{of}
\end{array}
\]

b. a book THE height of the lettering on the cover of which is excessive

The essence of pied-piping is that a PP or NP/DP has, with respect to extractability or subjecthood, the distribution of the wh-pronoun it contains. To put it another way, the root of the phrase (capitalized and italicized in (87a–b)) is in the relationship of ‘having certain aspects of the distribution of’ with respect to the wh-pronoun; the wh-pronoun confers upon the italicized word the right to be subject or extractee of the wh-pronoun’s clausal, so in this respect the italicized root acts as the wh-pronoun’s proxy. By incorporating this observation into the grammar, and establishing an associacy type ‘Proxy’, we can state the rule for which word is subject or extractee of the pronoun’s clausal.

(88) If P is a relative or interrogative pronoun, the subject or extractee of the clausal of P is a proxy of P.
If \( P \) is its own proxy, then (88) covers relative and interrogative clauses with and without pied piping.

To give an indication of the lines along which proxies of relative and interrogative pronouns are defined, I'll now discuss the two constructions in which the conditions on proxyhood are most and least restrictive. The most restricted is the proxy of wh-pronouns with \( \text{TO}_{\text{infinite}} \) clausal, and the proxy of interrogative pronouns with finite clausal in subordinate clauses. Assuming the grammaticality judgements given with (89a–o), a proxy of these wh-pronouns is defined so that (a) it must be a preposition, and (b) its complement must be the wh-pronoun or a proxy of the wh-pronoun. This means that in (89b,f,j,n), for instance, \( \text{up} \) and \( \text{to} \) are proxies of \textit{which}; it is \textit{up} that is an extractee of \textit{count}.

(89)  
\begin{align*}  
a & \text{ a book in which to find the truth} 
\quad b & \text{ a number up to which to count} 
\quad c & \text{ * a number up to the square of which to count} 
\quad d & \text{ * a number the square of which to count up to} 
\quad e & \text{ She wondered in which book to find the truth.} 
\quad f & \text{ She wondered up to which number to count.} 
\quad g & \text{ * She wondered up to the square of which number to count.} 
\quad h & \text{ * She wondered the square of which number to count up to.} 
\quad i & \text{ In which book to find the truth?} 
\quad j & \text{ Up to which number to count?} 
\quad k & \text{ * Up to the square of which number to count?} 
\quad l & \text{ * The square of which number to count up to?} 
\quad m & \text{ She wondered in which book he'd found the truth.} 
\quad n & \text{ She wondered up to which number he'd counted.} 
\quad o & \text{ * She wondered up to the square of which number he'd counted.} 
\end{align*}

The least restricted proxies are proxies of relative wh-pronouns that have finite (i.e. tensed) clausal. Since it is not my purpose here to debate the precise restrictions on what is eligible to be proxy of a wh-pronoun, I'll just assume that (90a–l) are grammatical; nothing hinges on this assumption.
(90)  a. The book, which a label is attached to the cover of, has been despatched.
b. The book, the cover of which a label is attached to, has been despatched.
c. The book, to the cover of which a label is attached, has been despatched.
d. The book, which she attached a label to the cover of, has been despatched.
e. The book, the cover of which she attached a label to, has been despatched.
f. The book, to the cover of which she attached a label, has been despatched.
g. The book, the paper in which she made notes on, has been despatched.
h. The book, thirteen bibliophiles desperate to gain possession of which she sent catalogues to, has been despatched.
i. The book, fond of which she'd always been, was despatched.
j. She described the film on first seeing which many people had fainted.
k. The shop was open, noticing which, she decided to go in.
l. The shop was open, that she entered which, he noticed.

(91a—b), which Hudson (1990) is willing to accept, I find unacceptable, but this, I suggest, is because of complements of nouns resist extraction, not because they’re ineligible to be proxies. Some of the speakers I’ve consulted will accept (91a—b) as “possible, but not something I’d ever say” and find them much worse than (90a—f). Judgements on (91a—b) correlate with judgements on *Of which book did she tear the cover?,* which supports my explanation.

(91)  a. * The book, of which a label is attached to the cover, has been despatched.
b. * The book, of which she attached a label to the cover, has been despatched.

Hudson’s analysis of pied-piping covers only (90a—f), and so undergenerates, noting which (p.c.), he has pointed (90j—k) out to me.

I offer (92) as a first stab at defining the range of proxies in (90a—l).
X is proxy of Y and Y is a relative wh-pronoun that has a finite clausal iff
X is Y or X is s-head of Y or X is s-head of a proxy of Y.

The effect of (92) is to treat every unsubordinate of W, a wh-pronoun, as a proxy of W.
(93) shows part of the relational structure of a pied piping construction of this sort. (See
§4.2 for a discussion of the s-dependency structure of pied piping.)

The restrictions on being a relative wh-pronoun’s proxy are thus hardly stringent. It
might be that these restrictions are too lax and that extra ones are needed. For instance,
if (90l) is ungrammatical then the proxy can’t be a finite (i.e. can’t be entered), and if
(90h–i) are ungrammatical then the proxy can’t be an adjective (desperate, fond), and if
(90g–h) are ungrammatical then proxies are defined in terms of complements rather
than dependents in general (because in is an adjunct of paper, and desperate an adjunct
of bibliophiles).

The proxy of W is an associate of W, but it is not a dependent of W. This is firstly
because a word can be proxy of itself but not dependent of itself, secondly because a
word can have multiple unconjoined proxies, but not multiple unconjoined non-adjunct
dependents of the same type, and thirdly because (contra Rosta (1994)) the proxy of W
cannot be s-dependent of W. Were proxies of W allowed, without constraint, to be s-
dependent of W, problems would arise. First, and less seriously, the associacy structure
would underdetermine the s-dependency structure to an extraordinary extent: for all the
proxies except the (ingenuity), the one associacy structure in (93) would permit several
alternative s-dependency structures, since, subject to projectivity, each proxy could be
s-dependent of the word it is complement of or the word it is proxy of (unless, as noted
earlier, a requirement for s-dependencies between certain pairs of words is stipulated).
For a sequence of $n$ proxies beginning with the one that is subject or extractee of the wh-pronoun's clausal and ending with the last one before the wh-pronoun, the number of possible s-dependency structures would be $2^{n-1}$ — for the ingenuity of the design of the cover of which there would be $2^8$ (256). For on a leg of whom there would be 8 s-dependency structures:

(94) a. on a leg of whom b. on a leg of whom
c. on a leg of whom d. on a leg of whom
e. on a leg of whom f. on a leg of whom
g. on a leg of whom h. on a leg of whom

A more serious problem is that in, say, *She brought a picture of the person who went every word bar she and went is proxy of who, since they are all superordinate of who. If proxies of W could be s-dependent of W, the s-dependency structure in (95) would therefore be possible.

(95) She brought the picture of the person who went.

Now the ungrammaticality of *She brought the immediately picture ought to be explicable as a projectivity violation, as in (96a), yet the ungrammatical (96b) would be able to get a projective s-dependency structure.

(96) a. *She brought the immediately picture
b. *She brought the immediately picture of the person who went.

The conclusion, then, is that if X is proxy of Y then X is unsubordinate of Y, which rules out X being s-dependent of Y.
2.6.2 *That* clauses

A problem with the analysis in which *that* is the root of the clause is that subcategorization rules for clausal complements must contain a disjunction: “The complement of a *know* can be a finite or a *that*”. The prevalence of this disjunction in the grammar, in oodles of subcategorization rules, would argue that we are missing some property that *thats* and finites have in common. I suggest that this shared property is that both *thats* and finites are proxies of finites: a *that* is a proxy of its finite complement and a finite is a proxy of itself. With the single rule (97), the relevant subcategorization rules can be simplified to: “The complement of a *know* can be a proxy of a finite”.

(97) If W is a finite, every proxy of W is a *that* or W.

In the same way, (98a—b) can be accounted for by rules which allow the proxy of a finite to be the (extraposed) complement of an *it* or an appositive adjunct of a noun.

(98) a. It's true (that) she went.
   b. the notion (that) pigs might fly

The associacy and s-dependency structure of *know (that) it is* is shown in (99a—b).

(99) a. know that it is
   b. know it is

A clinching piece of evidence in favour of this analysis comes from Payne's (1993) problematic example *They require that he be there by noon* (see also Cormack & Breheny (1994)). If a *require* subcategorizes for a CP, or, in Dependency Grammar terms, if its complement is a *that*, how can we state the rule that the finite verb within the subordinate clause may or must be subjunctive? Under the present proposals, a *require* can subcategorize for a proxy of a subjunctive, i.e. either for a subjunctive or for a *that* whose complement is a subjunctive.
2.6.3 Extraposition out of NP

(100a–b) are examples of extraposition out of a subject NP. The extrapositionless counterparts are (101a–b).

(100) a. The word had passed around that the colt from Old Regret had got away.
    b. The time has come to speak of many things.

![Diagram]

(101) a. The word that the colt from Old Regret had got away had passed around.
    b. The time to speak of many things has come.

Whereas extraction is recursive and unbounded, this extraposition is not unbounded. For example, while (102a–b) have the grammatical corresponding extrapositions in (103a–b), (104a–b)’s corresponding extrapositions in (105a–b) are not grammatical.

(102) a. The news that Sophy was to inherit was announced.
    b. The requirement that Sophy be pardoned was stipulated.

(103) a. The news was announced that Sophy was to inherit.
    b. The requirement was stipulated that Sophy be pardoned.

(104) a. Reports of the news that Sophy was to inherit were coming through.
    b. Rumours of the requirement that Sophy be pardoned were circulating.

(105) a. * Reports of the news were coming through that Sophy was to inherit.
    b. * Rumours of the requirement were circulating that Sophy be pardoned.

Proxies provide a mechanism for accounting for this data, if we assume rule (106).

(106) If W is subject of X, and Y is dependent (or complement, or whatever the actual restrictions are) of Z, and W is proxy of Z, then Y can be extraposee of X.
If we assume that in (107a) the is proxy of news and in (107b) dozens is proxy of rumours, but in (107c) reports is not proxy of news, then (106) correctly dictates that (107a—b) are grammatical and (107c) isn’t.

2.6.4 Cognate objects

According to Levin (1993: 95), “some basically intransitive verbs take as their object a noun zero-related to the verb — a so-called ‘cognate object’” (I assume ‘zero-related’ means ‘lexically related with no morphological difference’). If this observation is saying that cognate objects may be zero-related, it is not of any particular interest, but if the rule is that the cognate object of some verbs must be zero-related, this presents a very interesting problem. The rule would make the indicated grammaticality predictions about (108a—d).[^54]

[^54]: I’m assuming the rule just for the purposes of discussion. The requirement that the cognate object be zero-related is surely too strong, given examples like live a happy life, sing a song, think a strange thought or She cerebrated her usual cerebrations, and so on. But anyway, as it happens, examples like She lived a peaceful existence suggest that cognate objects are subject not to lexical constraints but to semantic constraints. The constraint possibly requires the object of the verb to be coreferential with the verb. (In WG, virtually all words, verbs included, have referents; the referent of a verb is a situation.) In support of the view that the constraint is semantic rather than lexical/syntactic, I offer the observation that I find acceptable all of the examples given.
(108)  a. She slept the profoundest sleep she'd ever slept.

b. * She slumbered the profoundest sleep she'd ever slept.

c. She smiled a most charming smile.

d. * She smiled a most charming grin.

The interesting problem that this assumed rule presents is that the word that must be lexically related to the verb needn't be the object of the verb. For example, in *She smiled a conspiratorial smile*, the object of *smiled* is *a*, not *smile*. If the rule were shown to contain some truth, it would be possible to offer a solution: if a determiner is a proxy of its complement, then the cognate object rule can be formulated so as to require the object of the verb to be a proxy of a word that is lexically related to the verb. (109) illustrates the relational structure this solution involves.

(109) of O
    . of C
    . p of
    She smiled a smile.

Other similar examples are (110a–b).

(110)  a. She craned a gracile neck.

b. She bears him the same grudge she's always borne.

Although it is more likely that the object of a *crane* is required to refer to the neck of referent of its subject, and the object of an instance of this idiomatic *bear* must refer to a grudge, in which case (111a–b) should be acceptable, the rule could be that the object of a *crane* must be a proxy of a *neck* and that the object of an idiomatic *bear* must be a proxy of a *grudge.*

(111)  a. (As for her neck,) she craned it.

b. (As for the grudge,) she no longer bore it.

55 A similar sort of problem concerning idioms like *kick the bucket* has been discussed from a WG perspective by Fraser (1989a).
3

Coordination

3.1 Two analyses of coordination

We are seeking an analysis of coordination that, essentially, does not conflict with the rest of the grammar. Specifically, we want the analysis to ensure firstly that coordination doesn't violate projectivity, and secondly that rules, especially word order rules, that pertain to X such that X is dependent of Y and s-dependent of Y still pertain to X and Y when X or Y is conjoined. Of course we also want the analysis to cope with all the grammatical phenomena of coordination, but that is largely, though not entirely, beyond the scope of the present study.

If projectivity is stated by rule, then we can add extra exception rules allowing a word to have two s-heads if the s-heads are conjoined, and allowing a word, X, to be separated from its s-dependent by a word, Y, that is not a subordinate of X, so long as X and Y are conjoined. But this of course is getting ever more arbitrary, and is therefore placing ever more of the burden of explanation on metagrammatical factors, such as functional pressure.

Similarly, if projectivity arises from the way utterances are processed, we should have also to assume (as Dick Hudson (p.c.) does) that there also exist procedures for processing coordination, and procedures for modifying the basic procedure responsible for projectivity in order to make allowances for coordination. It is possible to construct such a procedure, but to assume it would be to give a yet greater hostage to fortune, while allowing processing to swallow an increasingly large chunk of the grammar. (My stance on that matter is mentioned in §2.2.)

I shall assume that dependency structure maps onto s-dependency structure and that s-dependency structure maps onto constituent structure. Projectivity results from the continuity of constituents. This has been explained in §2.2. As we have seen, without coordination constituent structure is extremely simple, being equivalent to unlabelled
bracketing, and the definition of s-dependency is equally simple, s-dependenthood being equivalent to niecehood. But coordination disrupts this simple state of affairs, and consequently the constituent structures and the definition of s-dependency cannot both remain so simple. At least one must become more elaborate. The benefit will be that coordination doesn’t lead to projectivity violations. In this chapter I’ll discuss two alternative analyses of coordination. The first, which is a modified version of the traditional WG analysis (Hudson 1984, 1988, 1989, 1990), yields constituent structures like (112a); it employs extra types of constituent (as indicated by the kinds of brackets). I’ll call it ‘MWG’, for ‘Modified Word Grammar’. The second, which is like a typical Dependency Grammar analysis, in that the conjunction is the root of the coordination, yields constituent structures like (112b). I’ll call it ‘TDG’, for ‘Typical Dependency Grammar’. The chief difference between the MWG and TDG analyses is that TDG allows only constituent structure that is representable by unlabelled bracketing (with ‘unlabelled’ also implying that there be only one kind of bracket) but has a disjunctive definition of s-dependency, while MWG has a non-disjunctive (though not necessarily all that simple) definition of s-dependency, but requires different kinds of brackets.

(112) a. $[[<\text{my}> \text{ and } <\text{your}>] \text{ [friends]} \text{ met}]$

b. $[[\text{my} \text{ and } \text{your} \text{ [friends]} \text{ met}]$

I am inclined to favour the TDG analysis, (112b), so my exposition of the MWG will be a bit less explicit and articulated. The MWG and TDG analyses require much the same range of dependency types, while the MWG requires a greater range of constituent types, and it is highly unlikely that a fully elaborated MWG analysis would require fewer rules than the TDG. TDG is therefore to be preferred on the grounds of simplicity. Overall, TDG handles the data better, especially in the case of gapping, and though it is nothing like the traditional WG treatment of coordination, it is more in the spirit of WG than either the traditional WG or the MWG analyses, for it places the burden of representation on interword dependencies.

In the process of developing the TDG analysis I have considered in some depth some alternative dependency-based analyses of coordination, in which it is the root of a conjunct, rather than the conjunction, that is root of the whole coordination. Of these, the least unsuccessful yields constituent structures like (113).
This attempts to keep both to the simplest constituent structures and to the simplest definition of s-dependency, i.e. niecehood. Among its other attractions, it derives the order of elements within coordinations in a very simple way. But as well as being counterintuitive it falls down on explaining the word order of shared dependents of conjuncts; the rules become impossibly complex. I shall therefore discuss it no further.

The next section looks at the MWG analysis. §3.3 introduces various interword relationships that obtain in both the MWG and TDG analyses. §3.4 presents the TDG analysis. §3.5 shows how constructions like *Eat this here and that there* and various other forms of coordination are accommodated.

### 3.2 The ‘MWG’ analysis of coordination

(114a—f) are the rules for generating the constituent structure of coordinations. These rules assume there are four sorts of constituent in syntax: words, phrases, coordinations and conjuncts.

(114) a. Every phrase contains one part that is a word or a coordination, plus zero or more phrases.

   b. Every word is daughter of a phrase.

   c. Every coordination contains two or more conjuncts, plus, usually, one conjunction, and, sometimes, one ‘correlative’.

   d. Every conjunct contains one or more phrases.

   e. Every conjunct is daughter of a coordination.

   f. Every coordination is daughter of a phrase.

In addition, we need a set of rules that stipulate the ordering of elements within a coordination (i.e. correlative < one or more conjuncts < conjunction < one conjunct).

These rules give the bracketing shown in (115). The brackets indicate: [phrase], <conjunct>, {coordination}.

(115) *[[He] knows [[she] {<[washed]>, <[cored]> and <[peeled]>} [apples]]].*
The various bracketings will continue to ensure projectivity, but we need to alter (116), our earlier definition of s-dependency.

(116) If X is s-dependent of Y then X is niece of Y, and X is dependent of Y.

The purpose of this rule is to tie constituent structure to dependency structure, so that the latter determines the former, but it won't work for (115), since she, washed, cored, peeled and grapes have no head, because there is no word that is their aunt. If this headlessness were allowed, we would of course lose our ability to rule out a structure like (117).

(117) * [[He] knows [a] { <[washed]> , <[cored]> and <[peeled]> } [chef] [apples]].

"He knows a chef washed, cored and peeled apples."

(116) must therefore be reformulated along the lines of (118a−b). (118b) defines s-dependencies in terms of constituent structure. Constituent structure is derived only by (114a−f). (114a−f) permit any possible bracketing of words in a sentence, so clearly something is needed to constrain possible bracketings to only those that actually occur. This job is done by (118a), which obliges constituent structure to conform to dependency structure.

(118) a. If X is s-dependent of Y, then X is dependent of Y.
   b. X is s-dependent of Y iff
      (i) X and Y are descendants of a constituent Z, and
      (ii) the sequences of nodes from Z to X contains exactly one sequence in which a phrase is daughter of a phrase, and
      (iii) the sequences of nodes from Z to Y does not contain a sequence in which a phrase is daughter of a phrase.

Note that (ii−iii) in (118b) define relationships comprising not only those whereby X is niece of Y. (I consistently use 'niece' in the standard sense of 'daughter of sister of'.)

Let's see how (118a−b) applies to (115) and (117), which are given again using tree notation in (119a−b). (For purposes of exposition some nodes are numbered.)
(119) a. He knows she washed, cored and peeled apples.

b. * He knows a washed, cored and peeled chef apples.

In (119a), knows is the s-head of washed, of cored and of peeled, because (i) the sequence of nodes from node 1 to knows does not contain a phrase—phrase sequence, and (ii) the sequences from node 1 to washed, cored and peeled contain exactly one phrase—phrase sequence (node 1, node 2). Knows can't be the s-head of she because the sequence from node 1 to she contains two phrase—phrase sequences: node 1 and node 2, and node 2 and node 3. She is s-dependent of washed, cored and peeled because the sequence from node 2 to she contains exactly one phrase—phrase sequence (node 2, node 3), and the sequences from node 2 to washed, cored and peeled do not contain a phrase—phrase sequence. In (119b), chef must be s-dependent of washed, cored and peeled, because the sequence from the latter three to node 1 contains no phrase—phrase, while the sequence from node 1 to chef contains exactly one (node 1, node 2). For the constituent structure to be licit, the would therefore have to be dependent of washed, cored and peeled, yet no rule licences this; chef is dependent only of a, but it is not dependent of a. The constituent structure is therefore prohibited, and the sentence is ungrammatical.

The MWG analysis of a more complicated coordination is shown in the tree diagram in (120a). The s-dependencies that rule (118b) would derive from (120a) are shown by
stemma in (120b). In the stemma single lines are used for words with one s-head and double lines for words with more than one s-head.

(120) a.

According to (120b), the either, the or, both boths and both ands are s-dependents of that and s-heads of probably and yesterday. She, he and piteously are each s-dependents of a both and an and. None of these s-dependencies are paralleled by an existing dependency, so, since every s-dependency must be paralleled by a dependency, some new dependency types have to be created to do this job. These are Advener and Ward, which are introduced in the next section.

Since projectivity, in the sense of prohibited discontinuity, is derived from constituency, rather than from a (possibly extragrammatical) ban on tangling s-dependencies (as in Hudson (1994b) and Rosta (1994)), coordinate structures remain projective even though they can contain tangling s-dependencies. Also, it doesn't matter that shared s-dependents of conjoined words have more than one s-head: a word can have any number of s-heads so long as it is in an appropriate structural relationship to each of its s-heads. The rules that generate constituent structures are such that it is only in coordinate structures that two words can have the same s-dependent.
3.3 Some inter-word relationships involved in simple coordinations

A certain number of inter-word relationships, some but not all of them dependency types, appear to be needed in any analyses of coordination, irrespective of whether these relationships are basic or instead are derived from constituent structure. In this section I’ll consider three of these relationships: (i) Substitute, (ii) Coordinatee, (iii) Advener, (iv) Ward, and (v) Conjunct-mate. Only those of their properties that hold for each competing analysis of coordination are mentioned in this section. Subsequent sections will describe additional properties these relationships have in the TDG analysis, and will introduce other relevant inter-word relationships involved in coordination.

Substitutes. Substitute is not a dependency type, for a substitute of W can never be an s-dependent of W.56 This relationship takes the place of Hudson’s (1989, 1990) ‘Replacer’ and ‘Replacee’, which he uses for comparative constructions and for gapping. I see substitutes, unlike replacers/replacees, as occurring in all coordinate structures, not just gapping. If a word is linked to more than one of its dependents by the same type of dependency (with the probable exception of adjunct dependencies), then either each of these dependents will be substitute of the others, or all but one of the dependents will be substitutes of another of the dependents. We find substitutes in coordination and in comparative constructions, though if my tentative suggestions in §3.7 about than and as being conjunctions are correct then we can say that a sentence will contain a substitute if and only if it contains a conjunction. (The rules stating this are given in §3.6.3.) Thus, if a word has two subjects, then either its subjects will be conjoined or one of them will be in a comparative phrase. (121) illustrates a substitute in a comparative constructions.

(121) She has read more books than he/him.

56 Hopefully it will prove unnecessary to state rules such as (i–ii) (which state correct generalizations), if the conditions under which a word could be head or dependent of its substitute could never arise.

(i) If X is substitute of Y then X is not s-dependent of Y and X is not s-head of Y.
(ii) If X is substitute of Y then X is not subordinate of Y and X is not superordinate of Y.
In coordinate structures in which the conjuncts have just one root, the root of a conjunct is a substitute of roots of the other conjuncts, as illustrated in (122). Here Sophy and Edgar are substitutes of each other.

(122) Sophy and Edgar sat on a drawing pin.

The principal syntactic characteristic of substitutes is that if X is a substitute of Y then the dependencies holding between Y and Z, where Z is either s-head of Y or a subordinate of the s-head of Y but not a subordinate of Y also hold between X and Z. A formalization of this property and a discussion of other rules pertaining to substitutes are presented in §3.6. To simplify a little, the roots of conjuncts are substitutes of each other. This is formalized in §3.6.2.

The chief semantic characteristic of substitutes concerns the distributive interpretation of coordination. Most semantically conjunctive coordinations can receive either a distributive interpretation or a collective interpretation: Sophy and Edgar weigh twenty stone can mean either that they weigh 20 stone between them, which is the collective reading, or that they weigh 20 stone each, which is the distributive reading. The distributive but not the collective reading survives if the coordination is expanded into Sophy weighs twenty stone and Edgar weighs twenty stone. Only and with distributive reading is altered to or in non-assertive contexts: this is shown by (123a—d).

(123) a. I don’t think Sophy and Edgar weigh 20 stone each.
   [= “It’s not the case that I think Sophy and Edgar weigh 20 stone each.”]

b. I don’t think Sophy and Edgar weigh 20 stone together.
   [= “It’s not the case that I think Sophy and Edgar weigh 20 stone together.”]

c. I don’t think Sophy or Edgar weigh(s) 20 stone each.
   [= “I think Sophy and Edgar each don’t weigh 20 stone.”]

d. !* I don’t think Sophy or Edgar weigh(s) 20 stone together.
   [= “I think Sophy and Edgar don’t weigh 20 stone together.”]
Now consider (122), the most plausible interpretation of which is distributive, involving two different events: an event of Sophy sitting on a drawing pin and an event of Edgar sitting on a drawing pin. I shall informally call each event an 'expansion'. When a word, W, has conjoined dependents, and there is a distributive reading, W is associated with a range of expansions such that each conjoined dependent links semantically to a different expansion from its substitutes. This is the job the Substitute relationship plays in the semantics of coordination: in distributive interpretations a root of a conjunct links semantically to a different expansion from the substitutes of the conjunct root. In (122) the expansions are events of sitting on a drawing pin; *Sophy* links semantically to a different event of sitting on a drawing pin from *Edgar*, the substitute of Sophy. I shall make no attempt to investigate the semantics of expansions. And note that it is not just coordination that involves multiple expansions — plurals do too: in both *Sophy and Edgar smiled* and *The children smiled* we have multiple smilings.

**Coordinatees.** Each root word in a conjunct is coordinatee of the conjunction. This gives us the partial dependency structure in (124). 'Coordinatee' is abbreviated 'co'.

\[
\begin{array}{c}
\begin{array}{cccc}
s & \text{of} & c & c \\
\text{of} & c & c \\
\text{of} & c & c \\
\text{of} & c & c \\
\text{of} & c & c \\
\text{of} & c & c \\
\text{of} & c & c \\
\text{of} & c & c \\
\text{of} & c & c \\
\end{array}
\end{array}
\]

(124) I know she likes tea, eschews coffee and dislikes wine.

Some characteristic properties of coordinatees are as follows.

(i) Only conjunctions have coordinatees.

(ii) The rules for the semantics of conjunctions can refer to coordinatees: ORs express the disjunction of the meanings of their coordinatees; ANDs and BUTS express the conjunction of their coordinatees' meanings, and BUTS express contrast between them.

(iii) If one coordinatee of C, a conjunction, is linked by a dependency to a word, W, that is outside the coordination then all other coordinatees of C are also linked to W. This is stated more formally and with greater depth in §3.4 and §3.6.

---

57 For Hudson's treatment of the phenomenon, see Hudson (1990: §7.4).
Adveners. The rules that define s-dependency under the MWG analysis have the effect that both either and or in She'll have either coffee or tea are s-dependents of have. Since s-dependents of W must also be dependents of W, either and or must be dependents of have. What sort of dependent is it? We'll invent one — ‘Adveners’, and allow any word to have an advener. Adveners must be correlatives or conjunctions. The TDG analysis of coordination, introduced in the next section, makes greater use of adveners.

Wards. According to the MWG s-dependency definition, in [[She] {<[smiled]> and <[sipped [her [tea]]]>}], she is s-dependent not only of smiled and sipped but also of and. She must therefore be dependent of and. What sort of dependent is it? Again, just as we did with Advener, we'll invent one: ‘Ward’. In the next section I'll say more about wards in the TDG analysis. I shan't bother with a more formal treatment of wards in the MWG analysis, but it is clear that in MWG rules would have to refer to wards: in the MWG structure {<[men]> and <[[old] women]>}, what prevents old from being an adjunct of men as well as of women? The rules discussed so far simply require men and women to be substitutes of each other; they haven't said anything about the structural or sequential position of old. If old is dependent of both men and women and is not their s-head then we need to label old the ‘ward’, and require wards to occur outside the coordination.

Conjunct-mates. Conjunct-mate is a type of relationship holding between two roots of the same conjunct, as in, for example, (125a), a partial dependency structure for which is shown in (125b) ('d1, d2, d3' indicate different types of complement or adjunct, ‘ss’ stands for ‘substitute’, and ‘cm’ stands for ‘conjunct-mate’).

(125) a. She gave Simon snowdrops on Sundays, Myron marigolds on Mondays, and Timon tulips on Tuesdays.

---

58 Advene = “to accede or come to; to be superadded, as part of something, though not essential” (O.E.D.) Latin aduena = “stranger, foreigner”. The idea is that whereas X is a valent (subject or complement) of Y by requirement of Y, and X is adjunct of Y by permission of Y, the advener of W is dependent of W without any license or invitation from W. The term is also mnemonic in that it is proposed in §3.7 that only superveners (introduced in §3.4) can be adveners.

59 The motivation for this term is as follows. If X is ward of Y then Y is supervener of Z and X is s-dependent of Z. X would ordinarily (were Y not in the picture) be s-dependent only of Z, but in constituent structure X is in the care of (i.e. is niece of) the supervener (hence X is the ‘ward’ of Y, and Y is in loco parentis, or, more accurately, in loco materterae, for Z). Y holds X ‘in trust’ for Z (on which grounds I considered using ‘fiduciary’ rather than ‘ward’).
For extra clarity I re-present this coordination in tabular form in (126). Each row contains a conjunct: the italicized words in each row are conjunct-mates of each other. In each column, the italicized words are substitutes of each other.

(126)

<table>
<thead>
<tr>
<th>Inject</th>
<th>Object</th>
<th>Adjunct</th>
</tr>
</thead>
<tbody>
<tr>
<td>first conjunct:</td>
<td>Simon</td>
<td>snowdrops</td>
</tr>
<tr>
<td>second conjunct:</td>
<td>Myron</td>
<td>marigolds</td>
</tr>
<tr>
<td>third conjunct:</td>
<td>Timon</td>
<td>tulips</td>
</tr>
</tbody>
</table>

Whereas mutual substitutes are associated semantically with different expansions, mutual conjunct-mates are associated with the same expansion; that is, a word links semantically to the same expansion as its conjunct-mate does, and to a different expansion from the expansion its fellow coordinatees that aren’t its conjunct-mates link to. It is for this reason that there are no pairs of words such that one word is both substitute and conjunct-mate of the other. So (125) has one expansion per conjunct: (i) giving snowdrops to Simon on Sundays; (ii) giving marigolds to Myron on Mondays; (iii) giving tulips to Timon on Tuesdays. In (127), to (Rupert) and to (Daphne) are mutual substit-
utes, so we have an initial bifurcation into two sets of expansions — givings to Rupert, and givings to Daphne. The first and is a conjunct-mate of to (Rupert), so the set that expresses the conjunction of its members will be associated with the same expansions as to (Rupert) — i.e. the givings to Rupert. Rubies and roses are mutual substitutes, so they're associated with different expansions — givings of rubies to Rupert, and givings of roses to Rupert.

Some syntactic characteristics of conjunct-mates are given in (128). More are given in §3.5.1.

(128) \( \forall x \forall y \ [x \text{ is conjunct-mate of } y] \leftrightarrow \)

\[ [y \text{ is conjunct-mate of } x, \text{ and } \exists z \ x \text{ is coordinatee of } z, \text{ and } y \text{ is coordinatee of } z] \]

3.4 The ‘TDG’ analysis of coordination

The constituent structure of a simple coordination under the TDG analysis is illustrated in (129a–b); a partial dependency structure is also shown.

(129) a. [[[She] drinks [either [[tea], [coffee] or [cocoa]]]]].
I call this a ‘typical Dependency Grammar’ analysis first because it strikes me as obvious and intuitive (though not necessarily the more correct for being so), and second because it resembles to some degree the structures given to coordination by, e.g., Link Grammar (Sleator & Temperly 1991) and Lexicase (Starosta 1988 and postings to Dependency Grammar e-list). Of course some dependency theories handle coordination in different ways. WG has traditionally not used dependency at all. Mel’čuk (1988: 27) offers the structure in (130). (I’d have thought this would be unable to cope with old men and old women, but Mel’čuk doesn’t discuss such structures. Though he allows tangling, he doesn’t allow bicipitality, so he couldn’t allow old to be dependent of both men and women.)

(130) John and Mary

In the TDG analysis there are no conjunct or coordination constituents; phrases are the only sort of constituent there is. (114a–f), the rules that generate constituent structure under MWG can be simplified to (131a–b) (which correspond to (114a–b)).

(131) a. Every phrase has exactly one daughter that is a word, and zero or more daughters that are phrases.
   b. Every word is daughter of a phrase.
We continue to assume (132). Note that this rule constrains s-dependencies; it doesn’t license dependencies. By default no word can be a dependent of another word; each permissible dependency must be licensed by rule, and (132) is not such a rule.

(132) If X is s-dependent of Y then X is dependent of Y

In other respects, the TDG analysis, like the MWG, requires a special definition of s-dependency. We come to this below.

**Correlatives.** I begin the TDG analysis of coordination proper with ‘correlatives’ — both and either. As Hudson (1988) shows, these occur either before a verb, external to a coordination, as in (133a—b), or at the start of the coordination, as in (134a—b). I give (133a), which is in fact pretty ghastly, only because Hudson (1988) appears to allow for it; there is no problem in allowing only either to occupy the preverbal position. I have an account only of the correlatives’ position in (134a—b), but not of its position in (133a—b).  

(133) a. ? She both drinks tea and coffee.  
   b. She either drinks tea or coffee.  

(134) a. She drinks both tea and coffee.  
   b. She drinks either tea, coffee or cocoa.

Rules (135a—b) allow ands and ors to be complements of, respectively, boths and eithers. Since it is clear that some boths and some eithers are pronouns, and since it is not clear that any boths or eithers are ever not pronouns, I shall presume that correlat-

---

60 I assume a rule “∀x ∀y ∀z z is a type of dependency → x is not a z-type dependent of y” (see §3.6.2 for the formal means for stating this rule). This state of affairs described by this rule is without doubt the default, and, desirable though it would be to do so, it has yet to be shown that factors other than this default rule can rule out every dependency that is conceivable but doesn’t obtain. (For a sentence of n words there are 2(2(n−1)!) ordered pairs of words. The members of each of these ordered pairs could, in the absence of other constraints, be related by dependencies of each of however many types there are. With 20 dependency types and a sentence of 8 words, there’d be 20 × 2 × (8−1)! = 201,600 conceivable dependencies, only a tiny proportion of which actually obtain.

61 Hopefully any such account would show it to be part of a more general pattern, as (i—iv) suggest.

(i) She’ll have either tea or coffee.
(ii) She’ll either have tea or coffee.
(iii) She’ll have only tea.
(iv) She’ll only have tea.
ive BOTHs and EITHERs are pronouns. I also assume that like other transitive pronouns, they share their meaning with their complement.

(135) a. Some instances of AND are complement of a BOTH.
    b. Some instances of OR are complement of an EITHER.

The position of the correlative relative to the conjunction follows from two very general rules, which apply to other constructions besides coordination: (136a) states the apparent generalization that complements of pronouns (a.k.a. ‘determiners’) are always s-dependents of the pronoun, and (136b) states the even more general rule that if a word is a complement of its s-head then it follows its s-head. The upshot is that the conjunction must follow the correlative, and be its niece.

(136) a. If W is a complement of a pronoun then W is an s-dependent of that pronoun.
b. If X is a complement of Y and an s-dependent of Y, then X follows Y.

For the correlative’s position in (133), I (like Hudson in his earlier accounts of coordination) have no analysis to offer.

Superveners. The Supervener relationship is defined in (137). The effect of this rule is that correlatives are superveners of the conjunction’s coordinatees, but a conjunction is supervener of its coordinatees only when it is not complement of a correlative. So, in meet Sophy and Edgar, and is supervener of Sophy and of Edgar, while in meet both Sophy and Edgar, it is both that is supervener of Sophy and of Edgar.

(137) a. X is supervener of Y iff
    either [Y is coordinatee of X and X is not complement of Z]
or [Y is coordinatee of Z and Z is complement of X]

---

62 Supervene = “to come on or occur as something additional or extraneous” (O.E.D.). Latin superuenio = “come upon, rise above, to arrive unexpectedly”. The idea here is that when X is advene of Y and supervener of Z, Z is s-dependent of Y. X is superordinate of Z, and subordinate of Y, but is extraneous to the relationship between Y and Z. The sense “follow close on the heels of” is not relevant here.

63 ‘Supervener’ and §2.4’s ‘Legate’ are synonyms.
Whereas Coordinatee is a dependency type, for, as we'll see later in this section, nieces must be dependents, and coordinatees of conjunctions are nieces of the conjunction, it may be that Supervener is not, strictly speaking, a dependency type, since no word is ever a supervener of its aunt (or of its s-head, however s-heads be defined). In this respect, Substitute and Conjunct-mate are like Supervener.

**Coordinatees.** Conjunctions have two or more coordinatees. The relative ordering of the conjunction and its coordinatees can perhaps be explained on the basis of a suggestion made by Stan Starosta (on the Dependency Grammar e-list): one coordinatee is a complement of the conjunction, and, like complements in general, follows its s-head, while all other coordinatees are adjuncts of the conjunction, and their position is determined by a rule that makes adjuncts of a conjunction precede their s-head. In the context of WG, this suggestion makes the prediction that if the conjunction has more than one adjunct (i.e. preceding coordinatee) they can be conjoined (but needn't be), and that the conjunction can have more than one complement (i.e. following coordinatee) so long as they are conjoined. In the TDG analysis that prediction is wrong, since it would mean that in *Ann or Beth and Carl or Dave*, each of *Ann, Beth, Carl* and *Dave* could be coordinatee of *and*, whereas in the TDG analysis each *or*, but not *Ann, Beth, Carl or Dave*, is coordinatee of *and*.

---

64 ETCETERAS, which look rather like they're conjunctions, would be an exception in not having a complement — that is, in not having a following coordinatee. They just have an obligatory preceding adjunct coordinatee, or a preposed complement coordinatee.

65 I am assuming that there is a rule that requires multiple complements of the same type to be conjoined, and that allows multiple adjuncts of the same type to optionally be conjoined. That assumption may be incorrect. It may be that the rule in fact requires that when a word has multiple complements of the same type, each complement must be substitute of another, except for one which needn't be a substitute as long as it has a substitute. Adjuncts would be exempt from this requirement. But whichever version of the rule is correct, it doesn't explain why a conjunction can have only one complement coordinatee.

66 [Since writing this, I have come to the conclusion that the prediction is in fact correct, for it is far simpler to have an analysis in which the coordinatee of a conjunction is conjungible. In this case, the complement needn't be a niece of the conjunction. This simpler structure had not occurred to me previously. The structure I now favour is illustrated in (i). (Only advener and coordinatee dependencies are indicated; on advener dependencies, see below in this section).

```
<table>
<thead>
<tr>
<th>av</th>
<th>of</th>
</tr>
</thead>
<tbody>
<tr>
<td>co</td>
<td>of</td>
</tr>
<tr>
<td>co</td>
<td>of</td>
</tr>
<tr>
<td>of</td>
<td>co</td>
</tr>
<tr>
<td>of</td>
<td>co</td>
</tr>
<tr>
<td>of</td>
<td>co</td>
</tr>
<tr>
<td>of</td>
<td>co</td>
</tr>
</tbody>
</table>
```

(i) *Amy and Beatrice or Caroline and Diana.*

I therefore repudiate the remainder of this paragraph and the entirety of the next paragraph.]
However, other factors will account for the ungrammaticality of the predicted conjungibility, and in this case we can preserve Starosta's proposal. These other factors work as follows. First, the coordinatees of a conjunction are always nieces of the conjunction. This requirement can be achieved by either of (138a–b).

(138) a. If X is coordinatee of Y then X is niece of Y.
    b. No word is coordinatee of more than one word.

Second, superveners can themselves be coordinatees. We thus get the constituent structures in (139).

(139) a. [either [both [[Ann] and [Beth]] or [both [[Carl] and [Dave]]]]]  
    b. [either [[[Ann] and [Beth]] or [[Carl] and [Dave]]]]

Ann and Beth can't be coordinatees of both and and the first or, since a word can't have two aunts that are themselves words. It is for the same reason that Carl and Dave can't be coordinatees of the second or.

We have seen that taking the following coordinatee of a conjunction to be complement of the conjunction wrongly predicts that the complement of the conjunction should be conjungible. This is accounted for by, one way or another, requiring coordinatees to be nieces of the conjunction.67 Similarly, we find that if we take the conjunction to be complement of the correlative, then since complements of pronouns can normally be conjoined, as in a man or woman, we wrongly predict that (140) is grammatical. If we continue to take the conjunction to be complement of the correlative, then one of (141a–c) will serve to rule out (140).68 (141a–b) amount to much the same thing, but

---

67 [See last footnote for a repudiation of this point.]
68 A rule stating the generalization that no word has more than one supervener would not do the job here, since if each OR were complement of either, then either would be supervener of only Ann, Beth, Carl, and Dave, and and would be supervener only of each OR; that is, none of the words in this ungrammatical phrase has more than one supervener. Additionally, if such a rule were stated we
(141c) is completely different, though it achieves the same effect — that conjoined conjunctions can't be complement of a pronoun. What (141c) actually says is that a conjoined conjunction is supervener of its coordinatees, which means that it can't be complement of a correlative, since if it were, the correlative rather than the conjunction would be supervener of the conjunction's coordinatees, as stated by (137).

(140) * either both [Ann or Beth] and [Carl or Dave]

(141) a. If X is complement of Y, and Z is coordinatee of X, then X is niece of Y.
    b. If X is complement of Y, and X is instance of Conjunction, then X is niece of Y.
    c. If X is coordinatee of Y, and Y is coordinatee of Z, then Y is supervener of X.

At present I have little grounds for choosing between the alternatives, but it is worth noting that if (141c) and (138b) are chosen over their alternatives, then we have no rules that stipulate that X must be a niece of Y.\textsuperscript{69} That is, no rules stipulate what constituent structure must correspond to a given dependency structure. As that seems a desirable weakening of the power of the grammar, perhaps (141c) and (138b) should be preferred to their rivals.\textsuperscript{70}

In the TDG analysis, polysyndetic coordinations like \textit{Ann and Beth and Carl and Dave} are treated as embedded coordination. (MWG is the same in this respect.) Any of the structures in (142a–e) are possible. (Intuitions about the correspondence between syntax and intonation suggest to me that (142a) is the neutral default structure.) It might be objected that this gives rise to vacuous ambiguity, but, first, as I averred in §2.7,

\textsuperscript{69} With the possible exception of the quirky rules determining when the forms \textit{I/we/he/she/they} must occur (see §3.5.7 and later in this section), and with the exception of the very general default rule (i) and its only two exceptions, (ii–iii).

\begin{itemize}
\item (i) If X is s-dependent of Y then X is niece of Y.
\item (ii) If X is s-dependent of Y and Z is advener of Y and Z is supervener of X then X is not niece of Y.
\item (iii) If X is s-dependent of Y and X is ward of Z and Z is supervener of Y then X is not niece of Y.
\end{itemize}

See §2.2.3 for more discussion of this.

\textsuperscript{70} [(138b) is repudiated in a footnote earlier in this section. Even though in the simpler revised analysis proposed there, conjunctions are not coordinatees, (141c) should still work, for irrespective of whether conjunctions can be coordinatees, it still rules out conjunctions that are both coordinatee of a conjunction and complement of a correlative.]
vacuous ambiguity is not all that objectionable, and, second, it could be said that what we have here is not vacuous ambiguity, and (142a–e) happen to be synonymous simply because the same conjunction is used in each coordination; if some ands were changed to ors the meaning would of course change.

(142)  

a. [[[Ann] and [Beth]] and [Carl]] and [Dave]]

b. [[Ann] and [[Beth] and [Carl]]] and [Dave]]

c. [[[Ann] and [Beth]]] and [[Carl] and [Dave]]]

d. [Ann] and [[[Beth] and [Carl]]] and [Dave]]]

e. [Ann] and [[[Beth] and [Carl]]] and [Dave]]]

This strategy for dealing with polysyndeton helps with (143a–b), pointed out to me by Dick Hudson.

(143)  

a. Ann and Beth and Carl were playing a duet and a solo respectively.

b. Ann and Beth and Carl were playing a trio.

I observed in §3.3 that there are two kinds of and, one that gets a distributive interpretation and one that gets a collective interpretation. As I also pointed out in §3.3, only distributive ands change to or in nonassertive contexts; cf. (144a–d) (given earlier as (123a–d)).

(144)  

a. I don’t think Sophy and Edgar weigh 20 stone each.

[= “It’s not the case that I think Sophy and Edgar weigh 20 stone each.”]

b. I don’t think Sophy and Edgar weigh 20 stone together.

[= “It’s not the case that I think Sophy and Edgar weigh 20 stone togeth-
er.”]
c. I don’t think Sophy or Edgar weigh(s) 20 stone each.
   [= “I think Sophy and Edgar each don’t weigh 20 stone.”]

d. ! I don’t think Sophy or Edgar weigh(s) 20 stone together.
   [= “I think Sophy and Edgar don’t weigh 20 stone together.”]

The grammaticality of are in (145b) suggests perhaps that sometimes the morphological-phonological form or is not associated with an instance of the lexeme Or, but rather is associated with an instance of distributive AND occurring with an allomorphic form conditioned by the non-assertive context. A distributive AND with morphology or would therefore be a negative-polarity item.

(145) a. Sophy or Edgar is/*are visiting.
   b. I don’t think Sophy or Edgar is/are visiting.

We need more evidence before committing ourselves to whether on the one hand collective ANDs and distributive ANDs belong to different lexemes or to different sublexemes of AND, or, on the other hand, the difference is purely semantic. But for our present purposes, it is sufficient to recognize that the collective and distributive versions are distinguishable. Notating the collective version with a subscripted ‘C’ and the distributive version with a subscripted ‘D’, we can represent (143a) as (146).

(146) \([\text{Ann}_{c} \text{and}_{c} \text{Beth}] \text{ and}_{d} \text{Carl}] \text{ were playing a duet and a solo respectively.}\)

As for (143b), it is reasonable to hope that one of (147a – b) will prove to be an appropriate analysis.

(147) a. \([\text{Ann}_{c} \text{and}_{c} \text{Beth}] \text{ and}_{c} \text{Carl}] \text{ were playing a trio.}\)
   b. \([\text{Ann}_{c} \text{[Beth and}_{c} \text{Carl]}] \text{ were playing a trio.}\)

S-dependents. We now come to the definition of s-dependency under the TDG analysis. Suppose that we went for the simplest definition, treating ‘s-dependent’ as a synonym of ‘niece’. That works for everything except coordination. It would predict that
all rules that pertain to dependents of W when the dependents are also s-dependents of W do not apply when the dependents are coordinated. In general this will subvert the intent of the rules, showing up mainly in subversions of word order rules, since ideally word order rules always pertain to orderings between s-heads and s-dependents. For example, when X is the s-head of Y, X follows Y if Y is X’s subject and X precedes Y if Y is X’s object. Therefore nothing would rule out (148a—b) if Sophy and Edgar, which aren’t niece of arrived, and both OFs, which aren’t nieces of brother, were not s-dependents of arrived or brother.

    b. * He met the of Sophy and of Thalia brother.

Also, since the complement of a pronoun is always its s-dependent, it follows that if the s-dependent of W must be niece of W then complements of pronouns shouldn’t be conjungible. But they are: *her friend and colleague.

The solution is to define s-dependents not simply as nieces, but rather as dependents of W that either (i) are nieces of W or (ii) would have been nieces of W were they not conjoined. (149) is this definition.

(149) X is s-dependent of Y if either
     [X is instance of Word, and Y is instance of Word, and X is niece of Y],
     or [Z is supervener of X, and Z is s-dependent of Y]

The effect of this rule is that in Sophy and Edgar met, both Sophy and Edgar are s-dependent (i) of and, because they’re niece of and, and (ii) of met, because and is their supervener and, because and is niece of met, and is s-dependent of met. In (150) each of Ann, Beth, Carl and Dave is s-dependent of met, for the following reason. Either is niece of met and therefore s-dependent of met. Either is s-dependent of met and supervener of each BOTH, so therefore each BOTH is s-dependent of met. Each BOTH is s-dependent of met and supervener of Ann and Beth or of Carl and Dave, so Ann, Beth, Carl and Dave are s-dependent of met.

(150) [[either [(both [Ann] and [Beth])] or [both [(Carl) and Dave]})] met]
Since if X is subject and s-dependent of Y X precedes Y, (151b–e) are all going to be ruled out, now that we’ve ensured that coordinatees are s-dependents of their supervener’s s-head (which makes Ann, Beth and Carl s-dependents not only of and but also of arrived).

(151) a. Ann, Beth and Carl arrived.
    b. * Ann, Beth and arrived Carl.
    c. * Ann, Beth arrived and Carl.
    d. * Ann, arrived Beth and Carl.
    e. * Arrived Ann, Beth and Carl.

Adveners. Since in (150) the either and both boths are s-dependents of met, they must be some kind of dependent of met. What sort? We’ll use the Advener dependency introduced in §3.3, where it was seen to be needed for the MWG analysis too. Any word whose s-dependents can be coordinated can have an advener. Adveners are subject to the restriction in (152) (a slightly revised version of which is given in §3.5.4).

(152) If X is advener of Y then X is s-dependent of Y and X is supervener of Z.

Since (152) says that all adveners are s-dependents, this means that though in (153) where and when are adjuncts of visit and extractees of does, plan, to and visit, and is advener only of does.

(153) Where and when does she plan to visit?

Since all s-dependents are dependents, we could say that s-dependency is a variety of dependency, and since all adveners are s-dependents, we could say that the Advener relationship is a variety of s-dependency. For the same reason, the Ward relationship is also a variety of s-dependency. To wards, we now proceed.
Wards. If there is a ward, a dependent of each coordinatee, it must occur at the outer edge of the coordination. (154a–c). In these examples, young is the ward.

(154) a. a young English men and Italian women of 
   a young English men and Italian women of 
   b. * English men and young Italian women of 
   a young English men and Italian women of 
   c. * both young English men and Italian women of 

How do we account for this? Note first that the sequential position of the ward relative to the coordinatees is the same as it would be were the coordinatees not conjoined; cf. (155a–c), with wards in italics.

(155) a. * Met ten years ago and married four years later Sophy and Edgar.
   b. * He met the of Sophy brother and sister.
   c. * Find friend my and your.

The ungrammaticality of (155a–c) is explained if we assume that the ward is (somehow) s-dependent of each coordinatee, and therefore has its normal preceding or following position relative to its s-head. This also allows us to preserve without exception such rules as that which requires the complement of a pronoun to be s-dependent of the pronoun.

Taking wards to be s-dependents of coordinatees rules out (154b), and (156b–e), where the verbs are all conjoined.

(156) a. She sings, dances and juggles.
   b. * Dances, she sings and juggles.
   c. * Dances, sings she and juggles.
   d. * Dances, sings and she juggles.
   e. * Dances, sings and juggles she.

It will also rule out (157b–e), but there remains a problem, for it fails to rule out (154c) or (157f).
There also remains the problem that non-ward s-dependents occur nearer to their s-dependents than ward s-dependents. Thus (158a) is structurally ambiguous between readings (158b–c), while (158d) is unambiguously equivalent to (158e).

(158) a. She sometimes dances and sings.
   b. She sings and sometimes dances.
   c. She sometimes dances and sometimes sings.
   d. Sometimes she dances and sings.
   e. Sometimes she dances and sometimes she sings.

Our solution to both problems is to assume that Ward is a dependency holding between the ward and a supervener: any supervener can have a ward, and the ward is always s-dependent of the supervener (by the definition of s-dependency in (149)). (159a–b) have the constituent structure shown, by bracketing and by stemma.

(159) \[
[[\text{She} \ [\text{danced}], \ [\text{sang}] \ & \ [\text{juggled}]].
\]

b. \[
[[\text{She} \ either \ [(\text{danced}), \ (\text{sang}) \ or \ (\text{juggled})]].
\]

(158d) cannot mean (158b), with sometimes being dependent only of sings, because that would then involve the discontinuity shown in (160).

(160) * Sometimes she sings and dances.

The ban against discontinuity would also rule out (157b–d) (which are also ungrammatical for the reasons stated above, since the structure would be (161) (for (157b)).
But (157f) is not ruled out by discontinuity, since it has the projective structure shown in (162).

What rules this out is the general rule that if X is a complement and s-dependent of Y then X precedes other following non-complement s-dependents of Y. Or is a complement and s-dependent of either, so if she follows either then or should precede she. But if or does precede she as in Either danced, sang or juggled she, this breaks the rule that she should precede the verbs by virtue of being their subject and s-dependent. So the only structure possible is She either danced, sang or juggled.

It should be noted that we are predicting all of (163a—d) to be grammatical.

It is hard to see how these can be accounted for other than by stipulating that if X is ward of Y and Y is not a conjunction then either X must be a subject of each word Y is

---

71 It's also ruled out by the fact that the complement of a pronoun is tail of the pronoun (§2.3.5).

72 Curious datum: observe the agreement in (i).

(i) Them all both dancing and singing delights her.

The analysis developed in Chapters 5—6 affords us a representation where this agreement makes sense: the syntactic structure is as in (ii). Here there is only one subject of the finite rather than two, so the singular agreement follows.

(ii) THEM ALL BOTH DANCE AND SING DO DELIGHT HER

I don't yet know why (iii) is bad, or what if anything is wrong with the syntactic structure in (iv).

(iii) * Them all both dancing and singing delight her.

(iv) THEM ALL BOTH ING DANCE AND ING SING DO DELIGHT HER

(Incidentally, (ii) is in certain respects in conflict with the analysis of gerunds given in §5.4.3, wherein INGS are said to coenounce with the verb that is their xcomp, and the enunciation of the ing and its xcomp consists of the stem of the verb plus the suffix -ing. In contrast, (ii) is based on an assumption that the enunciation of the xcomp of INGS consists of the stem of the verb plus the suffix -ing, while the ING itself coenounces with the word that follows it, which is, in (ii), the BOTH.)
supervener of or each word Y is a supervener of is finite and has X as its adjunct. In the case of (163d–e) it is perhaps significant that either is in the preverbal position that it turns up in in examples like She'll either get a grade A or a grade B. Could it be that at least for (163d–e), if not for (163c), either can have a ward only if it is also itself an adjunct of the words it is a supervener of? The structure would be (164a), but conceivably it could be (164b).

(164) a. \([\text{She} \text{ either } [[\text{danced}] \text{ or } [\text{sang}]]].\)

b. \([\text{She} [\text{either}] [\text{danced}] \text{ or } [\text{sang}]].\)

At present I've nothing more to offer on these constraints on wards.

To finish off the discussion of wards, we need to account for the ungrammaticality of (165a–c).

(165) a. *Give young men magnolia and women wisteria.

["Give young men magnolia and young women wisteria"]

b. *Give new friends books and inventions names.

["Give new friends new books and new inventions new names"]

[103]
There is probably some principled reason for the ungrammaticality of these, but at present I can do no better than to define Ward in such a way as to rule them out. The rules are (166). One effect of this rule is that wards can’t be in a coordination in which there are conjunct-mates: by forcing coordinatees to be substitutes (line 8), we prevent them from being conjunct-mates, since mutual substitutes must be associated with different expansions while mutual conjunct-mates must be associated with the same expansion.

(166) a. \[\forall x \forall y \ [x \text{ is ward of } y] \leftrightarrow \]
\[ [x \text{ is s-dependent of } y, \text{ and} \]
\[ \exists z \ [y \text{ is supervener of } z, \text{ and} \]
\[ [\forall v \forall w \ [y \text{ is supervener of } v, \text{ and} \]
\[ y \text{ is supervener of } w] \rightarrow \]
\[ [x \text{ is s-dependent of } v, \text{ and} \]
\[ [\text{either } v = w \text{ or } v \text{ is substitute of } w]]]]\]

In core and peel apples, and is supervener of core and peel, and apples is s-dependent of core and peel, so apples is ward of and. In either core and peel or revere and worship apples, or is supervener of both ands, and apples is s-dependent of both ands — specifically, apples is ward of both ands and is therefore their s-dependent. Because apples is ward of both ands, it is s-dependent of the words each and is supervener of (core, peel, revere, worship).

I have stated that wards are s-dependents of coordinatees, but according to rule (149) they’re not, since they’re not nieces of the coordinatees and they do not have a supervener that is s-dependent of the coordinatees. So (149), the earlier definition of s-dependency, is going to have to be revised, to (167a). This, coupled with (132), repeated as (167b), is our final definition of s-dependency.
(167) a. X is s-dependent of Y if either
X is instance of Word, and Y is instance of Word, and X is niece of Y,
or Z is supervener of X, and Z is s-dependent of Y
or Z is supervener of Y and X is ward of Z
b. If X is s-dependent of Y then X is dependent of Y

Note that words can now have more than one s-head. For example, in (150) both is s-dependent (and coordinatee) of or and s-dependent (and advener) of met, and Ann is s-dependent (and coordinatee) of and and s-dependent (and subject) of met. This, however, is not in the least problematic, for continuity, which I’ve also been calling projectivity or banned tangling, is achieved by constituent structure rather than by not allowing words to have more than one head and banning tangling of s-dependencies. The grammar, however, says very little about constituency, referring instead to s-dependency. It is s-dependency that effects the correspondence between constituent structure and dependency structure.

Though the grammar does not generally refer to nieces, sometimes the distinction between between niecehood and s-dependency can be useful. As has been observed by Starosta (DG e-list) with respect to a Lexicase analysis similar to TDG, it offers us a nice explanation for the forms pronouns take in coordinations, as in (168a—d).

(168) a. Me and him are going.
    b. He and I are going.
    c. She wanted to invite Sophy and I.
    d. She wanted to invite Sophy and me.

According to the TDG analysis, the pronouns in (168) are s-dependents but not nieces of the verbs. For speakers who would always say (168a), rather than (168b), the rule would be that pronouns take subjective forms only when they are both subjects and nieces (and therefore also s-dependents) of finite verbs. Speakers who reject (168a) would have a rule requiring pronouns to have their subjective form whenever they are subjects and s-dependents of finite verbs, irrespective of whether they are nieces of the verb. For speakers who accept either of (168a—b), the subjective form is obligatory when the pronoun is a subject and a niece of the verb, and optional when it is subject

[105]
but not niece. For speakers who accept (168c), there seems to be a rule that requires a ME that is s-dependent of a conjunction to be a complement of the conjunction and to take its subjective form. The potential conditioning factors for subjective forms for pronouns are even more complicated, for it is also possible for a word to be subject but not s-dependent (and therefore not niece) of a finite verb, as in She reads more than he/him and as in the gapping constructions considered in §3.5, and it is possible for a word to be subject and s-dependent of something that is not a finite verb, e.g. She/Her being French, she likes to dine well. It would take an elaborate variational study to work out which factors condition subjective forms for which lects.

(169) presents a comprehensive diagram of the interword relationships involved in coordination under the TDG analysis. Constituency is shown by stemma, and t\(t^\}: labels of nieces are capitalized. The labels of s-dependents are italicized. (An uncapitalized italicized label therefore indicates an s-dependent of W that isn’t a niece of W.) The dependencies are labelled as follows: ‘a’ = adjunct, ‘c’ = complement, ‘co’ = coordinatee, ‘av’ = adverner, ‘w’ = ward, ‘s’ = subject. The non-dependency relationships are: ‘sv’ = supervener, ‘ss’ = substitute.

\footnote{It is commonly regarded as a hypercorrection, but it is so widespread and so frequent, especially among younger middle class speakers (according to my observations), that we must conclude that it is grammatical and, incidentally, is not stigmatized. Indeed, forth even from the mouths of aristocrats have I heard such forms as saw Sophy and I issuing.}
that probably either she both laughed and smiled or he both wept and loudly wailed piteously yesterday
3.5 Complex coordination and gapping, and supplementaries too

In this section I’ll discuss ways to analyse (i) what Quirk et al. (1985) call ‘complex coordination’, e.g. (170a), (ii) gapping, e.g. (170b), and (iii) ‘supplementary’ coordination, e.g. (170c). I’ll deal with them in this order.

(170) a. Many envy Sophy her composure and Edgar his poise.
    b. Sophy is envied for her composure, and Edgar for his poise.
    c. Sophy is envied for her composure, and Edgar too.

Once we’ve dealt with these varieties of coordination we’ll have covered most of the major types except for right node raising, covered in §5.3, and asyndetic coordination, contemplated wistfully in §3.7.

3.5.1 Complex coordination

Complex coordination is not gapping; rather, it is a form of what I’ll call ‘symmetric’ coordination, which is a variety of coordination that we’ve already looked at in the previous section. The justification for this conclusion will be given in §3.5.5, after we’ve dealt with gapping. Until then, we will simply take it for granted that complex coordination is symmetric coordination. The essential analytical ingredient that we will add to accommodate complex coordination is the Conjunct-mate relationship which holds between coordinatees that are, notionally, ‘in the same conjunct’. This was introduced in §3.3.

The TDG analysis outlined so far runs into certain problems with a coordination like (171).

(171) Give Sophy tulips and Edgar roses.

There are two problems with this. First, the position of the conjunction relative to its coordinatees has been explained by saying that all coordinatees are preceding adjuncts of the conjunction, except for one, which is a following complement. But we don’t get structures like (172).

---

74 Quirk et al. call this ‘appended coordination’, which is apt, but because of its inflectional suffix doesn’t convert from adjective to noun as satisfactorily as ‘supplementary’.
Second, nothing so far prevents a coordinatee from being separated from its conjunct-mates by fellow coordinatees that aren’t its conjunct-mates, as in (173).

Each problem must be fixed by stipulation. The second problem can be fixed by one of (174). Each of these rules is eccentric, especially the former. Both pertain to word order, yet neither needs to refer to s-dependents.

The first problem can be fixed by (175). This is an exception to the general rule that multiple complements of the same complement-type must be substitutes of each another.

In the MWG analysis, these problems don’t arise. The first problem doesn’t arise because it is stipulated that the daughters of a coordination comprise an optional correlative, followed by one or more conjuncts, followed by a conjunction, followed by a conjunct. The root of the final conjunct is not taken to be a complement of the conjunction. The second problem doesn’t arise because it is stipulated that X and Y are conjunct-mates of each other if and only if X and Y are granddaughters of the same conjunct, as in give \{
\langle [Sophy] [sapphire]\rangle \text{ and } \langle [Edgar] [amber]\rangle \}.

Complex coordination is never matrix coordination; the coordination is always subordinated to other words. For example, in give \{
\langle [Sophy] [sapphire]\rangle \text{ and } \langle [Amy] [amethyst]\rangle \} the coordination is subordinated to give. An example of impossible mat-
rix coordination is \{<\text{Sophy sang}] [Edgar danced]} \text{ or } <\text{Rupert rang}] [Thalia answered]}\}. I'm not sure whether this restriction on complex coordination follows from general properties of the grammar, but it definitely doesn’t follow from any of the other rules pertaining to coordination. If each of (i) Sophy, (ii) roses, (iii) Edgar and (iv) tulips is potentially a grammatical macrosyntagm but Sophy roses and Edgar tulips isn't, then we’d therefore need to stipulate the ban on conjunct-mates occurring in matrix coordination, by means of rule (176).

\begin{equation}
\text{If } X \text{ is conjunct-mate of } Y \text{ and } Z \text{ is supervener of } X \text{ then } Z \text{ is advener of } W.
\end{equation}

3.5.2 Problems of gapping

We now proceed to gapping. We’ll take as our prime example (177) because firstly the ‘gap’ corresponds to more than one word (will try to draw) and secondly because the position of the correlative suggests pretty strongly that for those speakers who accept the correlative the first conjunct is Ann will try to draw Beth. As it is hard to know how to punctuate this sentence perspicuously I’ve added some tonetic marks for a likely tune.

\begin{equation}
\text{Remember that } (%\text{either}) \text{ 'Ann will try to draw 'Beth, 'Carol 'Daphne or 'Eleanor 'Fred. ["Remember that either Ann will try to draw Beth, Carol will try to draw Daphne or Eleanor will try to draw Fred."]}
\end{equation}

I take for granted Hudson's (1989) summary of the basic facts of the gapping construction. I agree with Hudson that in no sense is a gap in any way syntactically present in this construction, but arguably we could consider it to be technically ungrammatical. More pre-

\footnote{Sophy roses and Edgar tulips is a normal answer to the quizshow question You gave who what?, but arguably we could consider it to be technically ungrammatical.}

\footnote{Not least among my reasons for thinking this is that I cannot even in principle think of a viable way within WG for the grammar to get there to be a syntactically present gap. One idea is that the words missing from the non-initial conjuncts last a long time and are simultaneous with other words. That way the gaps in the second conjuncts are filled by the same words (i.e. word tokens) that are not missing from the gapless conjunct: will, try, to and draw would begin at the point where they’re enounced, and end somewhere after the point where Eleanor is enounced. This idea is so weird I shan’t consider it further; in no other circumstances do we allow that words may be simultaneous or overlapping. Another idea is that copies of the gapped words in the first conjunct are syntactically present in the subsequent conjuncts but are either unpronounced or are pronounced at the same time.}
cishly, I find his analysis incoherent but easily remediable. I'll first explain his analysis, then why it’s incoherent, and then how it’s remedied. In Hudson’s analysis, (177) has the constituent structure shown by brackets in (178a), and, of crucial importance, the grammatical relations Replacer and Replacee relate words in the way indicated in the graph. The complete dependency structure under Hudson’s analysis is represented by stemma in (178b).

Hudson (1989: 76–77) requires that the “possible-function” of a replacer of W be the “possible-function of W” (his rule (41)). It is the notion of “possible-function” that I have problems with. It contrasts with “actual-function”. If X is the subject of Y then the actual-function of X is “subject of Y”. This clearly doesn’t mean that the actual-function of X is X, where “subject of Y” is understood to denote X. Rather, I interpret it to mean that the actual-function of X is the class of subjects of Y, where “subject of Y” is understood to denote the class of subjects of Y. Presumably there are a myriad of rules saying “X is subject of Y iff Z is function of X (i.e., X is a member of Z) and every member of Z is a subject of Y”, and so on for each different type of dependency.77 “Possible-function” is much harder to make sense of in anything better than an

---

77 In fairness, while the relationship ‘Function’ is terminologically redundant, given that we have its partial converse, Member (a.k.a. ‘Instance’), there are definite virtues to rules such as these, since they may help to formalize rules pertaining to the necessary properties of substitutes, such as the requirement that if a word has two subjects then one must be substitute of the other, and the requirement that if X is substitute of Y and Y is s-dependent of Z then the ‘actual-functions’ of X include the ‘actual-function’ of Y with respect to Z. See §§3.6.2–3 for the actual rules.
intuitive way. The idea is that in gapping the replacer does not actually share the
function of its replacee, but were the sentence paraphrased it could. There are two
claims here: first, that the replacer in some sense could perform the function of its re­
placee, and, second, that the replacer does not actually perform the same function of its
replacee. I disagree with the second claim, so naturally I agree with the first. Let’s see
some evidence for the first claim, and then Hudson’s reasons for the second.

(179a—b) (Hudson’s (15a—b)) show that it is not merely the meaning of the verb in
the gapless conjunct that determines the properties of the ‘remnants’ in the gapped con­
junct but also the verb’s syntactic properties — in this instance, it’s .ubcategorizational
properties. (180a—c) show the same thing: ceased requires its xcomp to be a to infiniti­
ve and stopped requires its xcomp to be an ing-participle; these requirents are also
imposed upon the replacers. (181a—b) show that idioms can survive gapping.

(179) a. Fred gave a present to Mary and Bill, to Jane.
   b. * Fred gave Mary a present and Bill, to Jane.

(180) a. Sophy ceased to dance and Edgar, to sing.
   b. Sophy stopped dancing, and Edgar, singing.
   c. * Sophy stopped dancing, and Edgar, to sing.

(181) a. Ann made a mess, Beth a mistake and Chloe a fuss.
   b. Ann brought her son up to be an anarchist, and Beth her daughter up to be
   a botanist.

The most obvious explanation for this compelling data is that the replacers are de­
pendents of the same words that their replacees are dependents of. So what is Hudson’s
reason for rejecting this? “This”, he says, “would indeed be the correct analysis for a
simpler example like John and Bill love Mary, in which the two nouns share the same

78 More accurately, I can’t really understand it in anything other than its normal English sense. The
definition is given as “relatum of possible-function of word is definition of relatum of itself”. Ex­
plaining “relatum”, Hudson says “the relatum of ‘subject of loves’ is loves”. So if the possible-func­
tion of W is ‘subject of loves’ then the relatum of the possible-function of W is loves. Then accord­
ing to Hudson’s rule, loves is definition of the relatum of W. As W has no relatum, since W is a
word, not a formula of the metalanguage, and as Hudson doesn’t say what a “definition of” means,
I conclude that the formal definition of Possible-function is unintelligible. Hudson says “the sent­
ence-word John has both an actual-function (‘subject of loves’) and also a possible-function, namely
‘subject of verb’”. Perhaps the intent is that not only words but also classes of words, such as the
class of verbs, have subjects; the actual-function of John would be the class of subjects of love,
while the possible-function of John would be the class of subjects of the class of verbs. How that
would help, I don’t know.
actual-function; but it cannot be right for our gapped example [John loves Jane and Bill, Beth] — otherwise why should the verb agree only with John, for example?” He gives no further examples, so the chief reason seems to be that the verb agreement cannot be explained if a replacer is subject of the verb. I shall show later that the verb agreement can be explained even if a replacer is its subject (though it could not have been so explained in WG circa 1989), so we can go ahead and conclude that replacers share the ‘actual-function’ of their replacees. This gives us the partial relational structure shown in (182). (As I explained in §3.3, I have replaced Replacer by Substitute and have done away with Replacee.)

(182)

I prefer (182) to (178a) not just because it will more effectively account for data like (179—181) but also because it will allow us to maintain the generalization that every word in a macrosyntagm except for the root word has an s-head. It’s not obvious, however, what the s-dependency structure of (177) should be. To be going on with, note that on the basis of the dependency structure shown in (182) it is impossible to build a structure of non-tangling s-dependencies, as (183a) shows, even though regular coordination can in principle yield a non-tangling s-dependency structure, as in (183b—c).  

79 In practise s-dependency structures like these, where external s-dependencies of the conjuncts are asymmetrical, shouldn’t be allowed, since they break various rules. For example, (i) breaks the rule requiring complements of a determiner to be s-dependents of the determiner, since in (i) friends is not s-dependent of his. It is because apostle in (iii) is complement but not s-dependent of his that (iii) is ungrammatical (in the analysis I’ll be ending up advocating, apostle is s-dependent only of and).

(i) 

meet his and her friends

(ii) She is his inspiration, and he her apostle.

(iii) *Amy is his inspiration, and Beth __Ault apostle.

[113]
3.5.3 MWG gapping

Next, let’s try (177) with MWG coordination. (177) rendered with an MWG treatment of the coordination gives us the constituent structure shown by the tree in (184). The problem here is that according to the tree, Carol, Daphne, Eleanor and Fred come out as s-dependents of that, yet none of them is a dependent of that.

I do not really think MWG can handle gapping, and therefore it should be abandoned, at least if TDG can do better, which it can, as I’ll be showing later. However, in fairness, I’ll briefly present my best shot at getting it to work. The goal is to make the remnants in the gapped conjuncts s-dependents only of the conjunction, since the conjunction is the only candidate for s-head of each of the remnants. The strategy for achieving this goal is to keep the definition of s-dependent constant and to vary the constituent structure of gapped coordinations. One could keep the constituent structure constant and vary the definition of s-dependency, but this would undermine the essential spirit of the MWG analysis: the MWG analysis pays for a single, non-disjunctive definition of s-dependency with enriched possibilities in constituent structure, while the TDG analysis gets simpler constituent structure by allowing a disjunctive definition of s-dependency.
I shall now proceed to list a number of possible constituent structures for gapped and nongapped coordination. I see few if any grounds for choosing between the alternatives; they all seem equally ill-motivated. The suggested structures assume that there is a distinction between conjunctions that allow gapped conjuncts (subscripted 'gap') and those that don't (subscripted 'nongap'). We start with (185), the structure we've so far been assuming for non-gapping coordination. In the representations for possible constituent structures for coordination the numerical subscripts indicate the number of elements possible. The s-dependency structure that would result from the constituent structure is shown by stemma; links to external s-dependents are shown by a double line. (The lines to external s-heads and s-dependents are shown following the coordination; this doesn't indicate that they must follow the coordination.)

(185) NON-GAPPING (I):

This structure can alternate with (186), in the sense that the necessary differences between the two structures are kept to a minimum. In this instance, (185) differs from (186) in that three additional phrasal nodes are added to the latter: the first one contains the final conjunct, the second contains the medial conjuncts, and the third contains the other two along with the conjunction. The effect is that given the way s-dependency structure corresponds to constituent structure in MWG (rule (118) in §3.2), the roots of the conjuncts dominated by these additional phrasal nodes will be s-dependents of the conjunction, which is exactly the result we need to achieve.

(186) GAPPING (I):

A second alternant for (186) is (187), which has the advantage that the alternation would look more natural: a coordination contains a 'conjunction phrase', and the internal structure of the conjunction phrase varies according to whether the conjunction is a gapping-conjunction. On the minus side, we must allow a conjunct to be daughter of a phrase.
(187) NON-GAPPING (II): \{either_{0-1} < \rightarrow [\langle >_{0-?} \text{or}_{\text{nongap}} < \rangle]\}

A third alternant for (186) is (188). The alternation becomes a bit more awkward than with (187), but on the plus side we avoid a conjunct being a daughter of a phrase. But we also lose the generalization that conjunctions are always penultimate daughters, and instead have to stipulate the position of the phrase containing the conjunction.

(188) NON-GAPPING (III): \{either_{0-1} < \langle >_{0-?} [\text{or}_{\text{nongap}}] < \rangle\}

Some further possibilities are (189a–b). These do without Coordination as a constituent type, at the cost of allowing a phrase to have a conjunct as a daughter, and, in the case of (189b), allowing a phrase not to have a word as a daughter. In both of (189a–b) the gapping or will be s-dependent of the root of the first conjunct (the one following either). In (189b) either gets the same s-head as or. In (189a) the s-head of either is the s-head of the s-head of or.

(189) a. GAPPING (II): \[\text{[either}_{0-1} < \rightarrow [\langle >_{0-?} \text{or}_{\text{gaps}} [\langle >]]]\]

b. GAPPING (III): \[[\text{either}]_{0-1} < \rightarrow [\langle >_{0-?} \text{or}_{\text{gaps}} [\langle >]]\]

These might alternate with (190a–b), which require allowing a conjunct to have a word as a daughter, with (190c–d), which are variants on NON-GAPPING (I), (185), with (190e), which is a variants on NON-GAPPING (II), (187), and with (190f–g), which are variants on NON-GAPPING (III), (188).

(190) a. NON-GAPPING (IV): \[\text{[either}_{0-1} < \rightarrow [\langle >_{0-?} \text{or}_{\text{nongap}} [\langle >]]]\]

b. NON-GAPPING (V): \[[\text{either}]_{0-1} < \rightarrow [\langle >_{0-?} \text{or}_{\text{nongap}} [\langle >]]\]
As I said above, it is hard to see any particular merits in these structures.

3.5.4 TDG gapping

We move on now to the TDG analysis of gapping. On the whole I find it more successful than any of the alternatives, which supports the conclusion that the TDG analysis of coordination in general, and the definition of s-dependency and the maximally simple constituent structures that go along with it, is the best analysis available.

Before we begin, it should be noted that not all speakers who are tolerably happy with gapping will accept a correlative in the position shown in (191a). To my knowledge, no one accepts the correlative in (191b), so we have two lects: Lect A, which allows a correlative in the (191a) position, and Lect B, which won’t allow a correlative. Both lects are understandable: Lect A makes gapping look more like symmetric coordination, while for Lect B the grammar of gapping is substantially simpler.

(191) a. I realize that (%either) she wants to make his acquaintance or he hers.
   b. I realize that Amy wants to make Bill’s acquaintance, (*either) Celia Dave’s or Eleanor Fred’s.
The main significance of the difference between Lects A and B is that the acceptability of correlatives in gapping constructions in Lect A indicates that they involve coordination with an initial gapless conjunct. In Lect B, by contrast, it is possible that each conjunct is gapped — that is, that what in Lect A is an initial gapless conjunct is in Lect B not a conjunct at all. For Lect B, the conjuncts in (191b) would be Celia Dave's and Eleanor Fred's, but not Amy wants to make Bill's acquaintance.

Since it is more mnemonic to do so, I'll call Lect A gapping 'inclusive' gapping, because the coordination includes the first conjunct, and I'll call Lect B gapping 'exclusive' gapping, because the coordination excludes what, in inclusive gapping, is the first conjunct. It must be remembered that while inclusive and exclusive gapping are distinct constructions, they won't coexist within a single idiolect.

Let's start by seeing what structures unaltered TDG gives to gapping. If we take there to be an initial gapless conjunct, as there must be in Lect A, then we get (192), with the partial s-dependencies shown by stemma. This won't work, for Carol, Daphne, Eleanor and Fred are s-dependents of that but not dependents of that.

(192) [that [[[Ann] will [try [to [draw [Beth]]]]], [Carol] [Daphne] or [Eleanor] [Fred]]].

For Lect B, if there are no gapless conjuncts, then we get (193a). An alternative structure, (193b) is ruled out because it would lead to the discontinuity shown in (193c).

(193) a. [that [[Ann] will [try [to [draw [Beth]]]], [Carol] [Daphne] or [Eleanor] [Fred]]].

b. [that [[Ann] will [try [to [draw [Beth], [Carol] [Daphne] or [Eleanor] [Fred]]]]].

c. that Ann will want to meet Beth very badly, Carol Daphne or Eleanor Fred.

(193a) makes Carol, Daphne, Eleanor and Fred s-dependents of will. Each should therefore be a dependent of will. Carol and Eleanor are subjects of will, but Daphne and Fred aren't dependents of will. Finding a way to make Daphne and Fred dependents of will won't lead us to a solution, because that would not account for inclusive gapping.

Instead, we'll pursue a solution in which the remnants in gapped conjuncts are s-dependents only of the conjunction. When trying this with MWG, this was done by alter-
ing constituent structure; this time, with TDG, constituent structure will remain the same but the definition of s-dependency will be modified specifically for gapping. (More accurately, and more flatteringly for the TDG analysis, the definition of s-dependency is in fact not modified at all; what is altered is the definition of Supervener, which features in the definition of s-dependency.) For both inclusive and exclusive gapping the Substitute and Conjunct-mate relationships illustrated in (194) obtain.

(194) that Ann will try to draw Beth, Carol Daphne or Eleanor Fred.

We shall assume that there are subtypes of Conjunction, among which are Gapping-conjunction and Symmetric-conjunction. And or an or can be a gapping-conjunction, but other conjunctions, such as buts, perhaps can’t.\(^9\)

We’ll take Lect A (inclusive gapping) first. In (194), the coordinatees of the gapping-conjunction are will, Carol, Daphne, Eleanor and Fred. If this were symmetric coordination, then by the earlier definition of s-dependency, repeated here as (195), each coordinatee would be s-dependent of that, but in fact of the coordinatees only will is dependent of that, so a way must be found to prevent Carol, Daphne, Eleanor and Fred from being s-dependents of that. In addition, inclusive gapping clearly allows wards with gapping, as in that probably either Sophy loves Edgar or Edgar Sophy, where probably is ward of either. The rule for wards (repeated here as (196)) applies to symmetric coordination in such a way as to make the ward an s-dependent of each coordinatee. But probably is dependent of will but not of the other coordinatees, so it can’t be s-dependent of the other coordinatees.

\(^9\) Hudson (1982) also says but is bad in gapping. But Sag et al. (1985) give examples, some from Neijt (1979), that probably show that it is only for pragmatic reasons that but is usually odd in gapping.
(195) X is s-dependent of Y if either
X is instance of Word, and Y is instance of Word, and X is niece of Y,
or Z is supervener of X, and Z is s-dependent of Y
or Z is supervener of Y and X is ward of Z

(196) \( \forall x \forall y \, [x \text{ is ward of } y] \iff \)
\[ [x \text{ is } s\text{-dependent of } y, \text{ and }]
\[ \exists z \, [y \text{ is supervener of } z, \text{ and }]
\[ [\forall v \forall w \, [y \text{ is supervener of } v, \text{ and }]
\[ y \text{ is supervener of } w] \rightarrow 
\[ [x \text{ is } s\text{-dependent of } v, \text{ and }]
\[ [\text{either } v = w 
\text{ or } v \text{ is substitute of } w]]]]

The problem is to find a way to get these rules to apply only to the first coordinatee of
the gapping-conjunction. The solution is to amend the definition of Supervener from the
earlier version (197) to (198a–b).

(197) X is supervener of Y iff
either [Y is coordinatee of X and X is not complement of Z]
or [Y is coordinatee of Z and Z is complement of X]

(198) a. \( \forall x \forall y \forall z \, [x \text{ is complement of } y \text{ and } z \text{ is coordinatee of } x] \rightarrow 
[\text{y is supervener of } x]^{81}

b. \( \forall x \forall y \, [x \text{ is supervener of } y \text{ and}] \iff 
\exists z \, [y \text{ is coordinatee of } z, \text{ and }]
[\text{either } z \text{ is instance of inclusive Gapping-conjunction, }
\text{and } y \text{ is the first coordinatee of } z]
or \text{z is instance of Symmetric-conjunction], and
[\text{either } z \text{ is complement of } x 
or \text{z = x}]]

The effect of (198a–b) is that the remnants in gapped conjuncts have no supervener.
Therefore (196) doesn’t apply to the remnants at all, and (195) applies to remnants only

---

81 Alternatively, in the light of (205) below,
\( \forall x \forall y \forall z \, [y \text{ is proxy of } x \text{ and } z \text{ is coordinatee of } x] \rightarrow [y \text{ is supervener of } x] $
in that they are coordinatees and nieces of the conjunction, and therefore are s-dependents only of the conjunction.

Rule (199) will apply to the remnant coordinatees of gapping-conjunctions; it requires each to have a conjunct-mate. In addition we need a rule stating which words the remnant coordinatees can be substitutes of: this rule is given in §3.6.

\[(199) \forall x \forall y \forall z \ [x \text{ is coordinatee of } y, \text{ and } \]
\[y \text{ is instance of Gapping-conjunction, and } \]
\[z \text{ is not supervener of } x] \rightarrow \]
\[\exists w \ x \text{ is conjunct-mate of } w, \text{ and } \]
\[z \text{ is not supervener of } w] \]

The nature of Lect A is that it accommodates gapping with the minimum of differentiation between gapping and symmetric coordination. It adds to (194) the structure shown in (200).

Now for Lect B. We deal first with its one definite difference from Lect A — the unacceptability of the gapping conjunction occurring with a correlative. There are various ways this constraint could be effected. First, if gapping-conjunctions were always superveners of their coordinatees, that would mean they could never be complements of a correlative. But we have to discard this option, since we've just defined Supervener in a way such that coordinatees that are remnants in a gapped conjunct have no supervener.

Second, at minimum the constraint against the correlative could be the only respect in which Lect B differs from Lect A. This could be formulated in various ways, with there being nothing more to say about Lect B. It could be that EITHERS select for. an OR that is a symmetric-conjunction, and BOTHs for an AND that is a symmetric-conjunction.

[121]
Or, gapping-conjunctions could be forbidden from being complements; that would of course mean they couldn’t be complement of a correlative. Or, any conjunction that is a complement of a pronoun could be required to be a symmetric-conjunction.

However, it is rather more explanatory to assume that Lect B differs from Lect A not in some arbitrary restriction on correlatives, but rather in that it handles gapping in a somewhat simpler way than Lect A.\(^2\) I will assume that in Lect B we have the structure shown in (201), in addition to that in (194). Here, the explanation for the impossibility of the correlative is that a gapping-conjunction is always an advener. The reason for this restriction is given later on in this section. If all gapping-conjunctions are adveners, then they couldn’t be complements of a correlative, since complements of correlatives aren’t adveners. This possibility depends on the assumption that all conjuncts in gapping in Lect B are gapped, since only on such an assumption can we be sure that the conjunction will have an s-head that it is advener of: that is, we must assume the structure She would do anything for him [or he for her] with She would do anything for him outside the coordination and with or as s-dependent and advener of would.

![Diagram](image)

(201)

\[\text{[that } [\text{Ann} \text{ will } \text{try to } \text{draw } \text{Beth}]], [\text{[Carol]} \text{ [Daphne] or } \text{[Eleanor]} \text{ [Fred]}]].\]

The coordinatees here are Carol, Daphne, Eleanor and Fred. And is advener and s-dependent of will. The coordinatees are not s-dependents of will, and Daphne and Fred are not dependents of will. As with inclusive gapping, the definitions of Ward and S-dependent can be kept intact by redefining Supervener. For exclusive gapping, Supervener can be defined by (202), the Lect B counterpart to Lect A’s (198b).

---

\(^2\) [I now reject the analysis given in the remainder of this paragraph, in favour of the alternative analysis offered in Chapter 4, whereby (i) for every conjunction C the proxy of C is C or is supervener of a coordinate of C, (ii) every advener is a supervener, (iii) in exclusive gapping, the conjunction, which is not a supervener, is adjunct of its s-head.]
The structure of (201) allows for a simplification, for exclusive gapping, of (199) to (203).

(203) \( \forall x \forall y \ [x \text{ is coordinatee of } y, \ \text{and} \ \ y \text{ is instance of Gapping-conjunction}] \rightarrow \ [\exists w \ x \text{ is conjunct-mate of } w] \)

A downside to allowing all the coordinatees of some conjunctions to have no supervener is that we must revise (155), repeated here as (204a), to (204b).

(204) a. If \( X \) is advener of \( Y \) then \( X \) is s-dependent of \( Y \) and \( X \) is supervener of \( Z \).

b. If \( X \) is advener of \( Y \) then \( X \) is s-dependent of \( Y \) and either \( X \) is supervener of \( Z \) or \( X \) is instance of Conjunction.

Perhaps we might formalize the intuition that the correlative acts as the proxy of the conjunction, by means of (205).

(205) \( \forall x \ [x \text{ is instance of Conjunction}] \leftrightarrow \ [x \text{ is proxy of } x \text{ or } \exists y, x \text{ is complement of } y \text{ and } y \text{ is proxy of } x] \)

This affords a simplification of (204b), (198b) and (202) to (206a–c) respectively.

(206) a. \( \forall x \forall y \ [x \text{ is advener of } y] \rightarrow \ [\exists z, z \text{ is instance of Conjunction and } x \text{ is proxy of } z] \)

83 Under the preferred analysis offered in the last footnote, this downside does not hold, so (155)/(204a) stands, and (204b) is rejected.

84 Under the preferred analysis given in Chapter 4, this rule does not apply. In its stead is a rule that every advener is a supervener.
b. $\forall x \forall y [x$ is supervener of $y] \leftrightarrow \exists z [y$ is coordinatee of $z$, and $x$ is proxy of $z$, and
   [either $z$ is instance of inclusive Gapping-conjunction, and $y$ is the first coordinatee of $z]$
   or $z$ is instance of Symmetric-conjunction]]$

c. $\forall x \forall y [x$ is supervener of $y] \leftrightarrow \exists z [y$ is coordinatee of $z$, and $z$ is instance of Symmetric-conjunction, and $x$ is proxy of $z]$

I have concluded the discussion of Lects A and B for the present, though we return to them in §3.6. It is notable that the grammatical rules they require are few, and to square the analysis of gapping with the analysis of symmetric coordination, all it takes is a little tweaking of the definition of Supervener.

### 3.5.5 Why complex coordination is symmetric rather than gapping

Under the analysis of coordination developed so far, (207) is structurally ambiguous (though not semantically, as far as I can tell). If it is symmetric coordination, it has the bracketing given in (208a). If it is gapping, then in Lect A it has the structure in (208b), and in Lect B it has the structure in (208c).

(207) Give Sophy roses, Daphne laurel or Edgar tulips.
(208) a. [Give [[Sophy] [roses], [Daphne] [laurel] or [Edgar] [tulips]]].
   b. [[Give [Sophy] [roses]] [Daphne] [laurel] or [Edgar] [tulips]].
   c. [Give [Sophy] [roses] [[Daphne] [laurel] or [Edgar] [tulips]]].

In the light of this structural ambiguity we might wonder whether complex coordination can be considered gapping, as Sag et al. (1985) argue. In that case the analysis of symmetric coordination could be simplified a little, in that conjunct-mates would be restricted to gapping. There are two reasons for rejecting this step. The first and most compelling reason is given by Hudson (1989): a correlative is possible in complex coordinat-
ion, and it occurs in the position indicated in (209a). This could only be given the bracketing shown in (209b).

(209) a. Give either Sophy roses, Daphne laurel or Edgar tulips.
   b. [Give [either [[Sophy] [roses], [Daphne] [laurel] or [Edgar] [tulips]]].

The second piece of evidence comes from the requirement that in gapping the root of the first conjunct must be a predicate — i.e. it must be a word that has a subject. In inclusive gapping it is the first coordinatee that is a predicate; this is stated by (210a). In exclusive gapping it is the s-head of the gapping conjunction that is a predicate, (210). This rules out *large pictures of Sophy and small of Edgar.

(210) a. Inclusive gapping (Lect A):
   If X is instance of Gapping-conjunction and X is supervener of Y then Y is instance of ‘Predicate’.
   b. Exclusive gapping (Lect B):
   If X is instance of Gapping-conjunction then X is advener of Y and Y is instance of ‘Predicate’.

3.5.6 Supplementary coordination, and a generalization about conjunctions

Besides Symmetric-conjunction and Gapping-conjunction, we recognize an additional variety, Supplementary-conjunction, examples of which are given in (211a—d), and perhaps (211e) too.

---

85 Hudson (1989) gives an essentially similar formulation of this constraint, but because of examples like *Always eat slowly, never noisily, and sometimes moderately* he uses a disjunction whereby the root of the first conjunct must either have a subject or be a verb. The need for this disjunction is obviated by referring to a category like ‘Predicate’, which is reasonably well-motivated, as evidence presented in §6.3 shows.

86 There are other constraints on what the gap can contain which I haven’t dealt with. For example, Quirk et al. (1985: 975) give the following:

(i) Roberts won in 1979 and Dingwall in 1980.
(iii) * I looked at Agnes and she me.

This, they say, is because “for a construction to count as an acceptable gapping, it is necessary that each conjoin consist of a sequence of units at the same level of structure: *viz at the level of clause structure*. If I understand that correctly, it wrongly excludes *She wants to be friends with him and he with her.*
I don’t know how to analyse *too, either or not*, but now that we’ve seen how gapping works it’s fairly clear how supplementary-conjunctions work (at least if we ignore the NOTS, TOOS, etc.).

As with gapping, with supplementary coordination we may find an interlectal split according to whether the correlative is acceptable in *Either Sophy arrived late, or Edgar* — we may find both exclusive supplementaries and inclusive supplementaries. Rules for inclusive supplementaries, with which the correlative is acceptable, will differ from rules for exclusive supplementaries, with which the correlative is not acceptable, in much the same way as rules for inclusive and exclusive gapping. For inclusive supplementary coordination we’d have the bracketings in (212a–b), while for exclusive supplementary we have the bracketing in (213).

(212) a.  

\[
[\text{Either } [\text{[Sophy] had [arrived [late]]}, \text{ or } \text{[Edgar]]}].
\]

b.  

\[
[[\text{Sophy} \text{ had [arrived [late]]}}, \text{ or } \text{[Edgar]]}].
\]

(213)  

\[
[\text{Sophy} \text{ had [arrived [late]], or [Edgar]]}.]
\]

For the remainder of this section I’ll assume we have only exclusive supplementaries, but in §3.6.1 I give rules for both inclusive and exclusive versions.

As with exclusive gapping, the coordinatees of supplementary-conjunctions aren’t s-dependents of the word the conjunction is s-dependent of. For example, in (211a) *Edgar* can’t be s-dependent of *arrived*, for if it were it would have to precede *arrived*. Again like exclusive gapping, the conjunction can’t have a ward (*The grownups arrived late, small and children* [= “and small children”]). These facts are as problematic for the MWG analysis as we saw gapping to be, in §3.5.3. But in the TDG analysis we already have an explanation: the supplementary-conjunction is not supervener of the coordinatees. This follows from the rules for defining Supervener, the last versions of which were (206b–c).
The unacceptability of a correlative, and the failure of supplementary-conjunction to have a ward or to share its s-head with its coordinatees are all consequences of the conjunction not being a supervener. I said in §3.5.4 that an exclusive gapping-conjunction is always an advener, thus explaining why it can’t be complement of a correlative. This constraint will in fact follow from a general rule, (214), which applies to all conjunctions.

(214) \( \forall x \; [x \text{ is instance of Conjunction}] \rightarrow \)

\[ \exists y \; x \text{ is advener of } y, \text{ or} \]
\[ x \text{ is supervener of } y, \text{ or} \]
\[ x \text{ is complement of } y \]

If a conjunction is complement of a correlative then by (198a) the correlative must be supervener of a coordinatee of the conjunction. Coordinatees only have a supervener if the conjunction is symmetric, or if the coordinatee is the first coordinatee of an inclusive conjunction (gapping or supplementary). Coordinatees of exclusive gapping and supplementary conjunctions and noninitial coordinatees of inclusive gapping and supplementary conjunctions can’t have a supervener. Thus the effect of (214) is that unless the conjunction is a symmetric or inclusive gapping or supplementary, the conjunction must be an advener. That means the conjunction can’t be complement of a correlative, and it can’t occur in matrix coordination, where the conjunction (or correlative) has no s-head. In addition, these conjunctions can’t have a ward, since the rule defining wards, (166)/(196), states that only superveners can have wards. Thus, in summary, we have explained the correlation between the possibility of a conjunction (i) lacking an s-head, (ii) being complement of a correlative, (iii) having a ward. And we’ve also seen that these also correlate with whether s-dependencies can hold between coordinatees and words outside the coordination.

3.5.7 Agreement and other quirks of morphology

In §3.5.2 we saw that Hudson’s objection to taking the remnants in gapped conjuncts to be dependents of words in (what he takes to be) the initial ungapped conjunct is that this won’t account for verb agreement. The facts we need to account for are laid out in (215).
In each sentence the verb has two subjects, but not all the subjects are also s-dependents of the verb. In (215a–b) the verb has two subjects that are also its s-dependents, while in (215c–d) the verb has only one subject that is also its dependent, and the subject that is its s-dependent is singular. So, as a first stab at the rule for agreement we can offer (216).

(216) A singular verb has only one subject that is also its s-dependent, and the word that is its subject and s-dependent is a singular.

This must be amended to take account of Who did she deny is eligible?, where is agrees with who but where who is not s-dependent of is. Our revised version is (217).

(217) For every singular verb there is exactly one word that is both its subject and either its s-dependent or its extractee, and this word is a singular.

The essential point here is that the agreement rules refer not to subjects alone, but also to subjects of W that are also either s-dependents or extractees of W.87

A similar, though much more complex and problematic, account can be given for the form that instances of the lexemes Him, Her, Them, Us and Me take when they are subjects – that is, when do they have the form he/she/they/we/I? There is an enormous amount of interlectal variation here, so I can only sketch some of the factors that would be involved in a proper description. First, we have a contrast between (218a), where the pronoun is niece and therefore s-dependent and therefore dependent of are, (218b), where the pronoun is not niece of are but is s-dependent of are, and (218c), where the

---

87 I therefore feel entitled to ignore the agreement in Sophy or Edgar was here, and many other quirks in the grammar of agreement.
pronoun is neither niece nor s-dependent of *are*. In each sentence the pronoun is subject of *are*.

(218) a. They are here.
   b. Sophy and {he/him} are here.
   c. Sophy and Edgar are seeking thrills, and {he/him} enlightenment.

In other constructions where it seems plausible to think a pronoun is subject but not s-dependent of a finite verb we get mixed results. In (219a) we find the expected variability, but not, I think, in (219b), though in (219c), where the construction is more predicative, *she* is perhaps marginally acceptable.

(219) a. He arrived later than {her/she}.
   b. He arrived late, and {her/*?she} too.
   c. He is here, and {her/?she} too.

We should also have to consider cases where pronouns are both subjects and s-dependents, but not of a finite. In (220a) the pronoun is subject and s-dependent of a non-finite verb, while in (220b—c) the pronoun is subject and s-dependent of a noun.

(220) a. {%Her/%She} being a catholic, she punctured all the condoms with a toothpick.
   b. {%Her/%She} a catholic, she punctured all the condoms with a toothpick.
   c. She collects condoms, and {her/%she} a catholic to boot.

On the whole my impression of my own idiolect with respect to these matters is that if the pronoun is subject and niece of a finite then it must have its subjective form, and if it is not subject then it can’t have its subjective form, while if it is subject but not niece then anything goes, with the subjective form being consistently more formal and literary than the default form.

---

88 If my analysis of these constructions is correct then we would not expect to find anyone rejecting *him* in (i) but accepting it in (ii). Such is the idiosyncrasy of people’s judgements on these matters, I’m confident that one could find speakers to falsify that prediction.

(i) She chose vanilla and {he/*him} chocolate.
(ii) She chose vanilla and Sophy and him chocolate.
3.6 Substitutes

In §3.6.1 I give the rules that state for each variety of conjunction where the words its coordinatees are substitutes of are to be found. It is chiefly these rules that determine how many conjuncts there can be in each kind of coordination. In §§3.6.2–3 I give the rules that define dependency sharing and the properties of substitutes, as they pertain to wards and coordinatees.

3.6.1 What is substitute of what

With certain exceptions, for every coordinatee there is always a word that it is substitute of. The first exception is with Than, if thans really are conjunctions: in Be richer than Croesus, Croesus might be a coordinatee of the than, but there is no word it is substitute of, though there is in She is richer than Croesus. The second exception is with inclusive gapping (or supplementary) where the first, gapless, coordinatee is not substitute of any word. The third exception is with symmetric coordination, wherein a coordinatee needn’t be substitute of any word, so long as (i) it is conjunct-mate of another coordinatee that is substitute of some word, and (ii) it is an adjunct rather than a complement of the word its supervener is advener of.\(^8^9\) An example where this happens is (221); to (Thalia) is not a substitute, but it is a conjunct-mate of thistles and of on (Thursdays), both of which are substitutes, of freesias and on (Fridays) respectively.

(221) Give (either) thistles to Thalia on Thursdays or freesias on Fridays.

---

\(^{89}\) The basis for this descriptive statement is somewhat shaky. In order to test what precisely the nature and extent of this exception is, I conducted a small experiment in which I simply presented five native speakers (DH, CK, JW, JH, plus me) with all forty-eight possible sentences formed from Give (marigolds) (to Marina) (on Monday), (tulips) (to Tilly) (on Tuesday) and (wisteria) (to Walter) (on Wednesday) by omitting from each conjunct either one or none of the bracketed phrases. Even allowing for the tendency of this experimental method to give unreliable results, the results were nonetheless strikingly chaotic. There were only three sentences that all five accepted, and all these were ones amenable to analysis in which every conjunct-mate has a substitute. There were in fact more than three sentences amenable to such an analysis, and in these cases four out of five accepted them; DH had clearly scanted out as his acceptance of such sentences under conditions more conducive to valid judgements is on record in numerous books and articles published by him over the last twenty years. For 23 out of 48 sentences four or five out of the five respondents agreed in their judgements, but I’ve not detected any pattern in these agreements. Looking at individual sets of responses, DH had clearly scanted out, CK and JW seemed pretty erratic, and the remaining two, JH and me, were fairly clearly not erratic and agreed on 46 out of 48 sentences (which is an extremely surprisingly high rate of agreement for experiments of this sort). These apparently reliable responses of JH and me are consistent with the rule that I’ve suggested.

[130]
This third exception, incidentally, is not allowed for by Pickering & Barry's (1993) Dependency Categorial Grammar account of coordination, wherein, essentially, each fragment within a conjunct must be paralleled by a fragment within another conjunct, with the nature of the parallel relationship being that parallel fragments must have the same kind of 'classical' dependency link to words outside the coordination. ('Classical' in this context means 'excluding the extra dependencies that Word Grammar adds to traditional DG'; in the light of the present study we could understand that as meaning 'excluding non-skeletal dependencies'.) The problem for Pickering & Barry's account is that if W, an adjunct of a word external to the coordination, has a conjunct-mate that is substitute of a fellow coordinatee then W needn't be substitute of a fellow coordinatee. They appear to overlook this problem, though they do note without further comment that (222) (from Milward (1991)) is a “problematic example” (I've replaced their bracket notation with mine).

(222) You can call me {< [directly] > or < [after 5pm] [through my secretary] >}.

In fairness to Pickering & Barry, their account covers the full breadth of the data very impressively, and one would imagine that they could find a way to accommodate this minor exception. My analysis accommodates it simply by stipulation (see below). Furthermore, if (222) genuinely has the bracketing shown then it is not allowed for by the exception I've described, unless either after or through is substitute of directly. However, I suspect that (222) in fact has one of the structures indicated in (223a–b).

(222) a. You can call me [[directly] or [after [5pm] [through [my [secretary]]]]].
   b. You can call me [[directly] or [[after [5pm]] through [my [secretary]]]].

With the first exception ignored due to the dubiety surrounding it, the requirement on coordinatees is formalized in (224). Lines 2–4 of this rule deal with the second exception: they cause the rule to not apply to initial coordinatees in inclusive gapping or supplementary coordination. Lines 6–8 deal with the third exception.

\[90\] In some ways their coverage is too thorough. For example, their account of coordination includes right node raising, but, as I argue in §5.3, there are rather compelling reasons for thinking that right node raising really has nothing to do with coordination (except in as much that in actual patterns of usage, RNR tends to cooccur with coordination).
\[(224) \quad \forall a \exists b \quad [a \text{ is coordinatee of } b, \text{ and}\\ \quad \text{not } \exists c \quad c \text{ is supervener of } a, \text{ and}\\ \quad \text{not } \forall d \quad [d \text{ is coordinatee of } b] \rightarrow \\
\quad \quad [c \text{ is supervener of } d]]] \\
\rightarrow \\
\exists e \quad [\text{either } [a \text{ is substitute of } e]\\ \text{or } \exists f \exists g \quad [f \text{ is supervener of } a, \text{ and}\\ \quad g \text{ is conjunct-mate of } a, \text{ and}\\ \quad g \text{ is substitute of } e]]]]
\]

This rule explains why, say, *[Go and [Sophy]] and *[She and [arrived]] are ungrammatical: there is no word Sophie or arrived is substitute of. It also makes [And [she arrived]] ungrammatical, since arrived is not a substitute. The rule also has the effect of requiring all conjunct-mates in gapping to be substitutes, so Marina bought marigolds, and on Tuesday Tilly (tulips) is ungrammatical because there is no word of which on (Tuesday) is substitute.

(224) does not rule out give Sophie, Edgar, roses and tulips, for each coordinatee is a substitute (Sophy and Edgar are substitutes of each other and roses and tulips are substitutes of each other). We therefore need the further rule, (225), which specifies what relationships must hold between a pair of fellow coordinatees, \(v\) and \(w\). Line 6 of the rule has the effect of exempting first coordinatee in inclusive gapping or supplementary.

\[(225) \quad \forall v \exists w \exists x \quad [w \text{ is coordinatee of } x \text{ and } v \text{ is coordinatee of } x] \rightarrow \\
\quad [v \text{ is substitute of } w, \text{ or}\\ \quad v \text{ is conjunct-mate of } w, \text{ or}\\ \exists y \quad [[v \text{ is substitute of } y \text{ and } y \text{ is conjunct-mate of } w] \text{ or}\\ \quad [v \text{ is conjunct-mate of } y \text{ and } y \text{ is substitute of } w] \text{ or}\\ \exists z \quad [z \text{ is supervener of } v \text{ and } z \text{ is not supervener of } w]]
\]

Next we come to rules specifying where to find words a coordinatee is substitute of. (226) is the rule for symmetric coordination.
\[(226)\] \[\forall w \forall x \text{ [} w \text{ is coordinate of } x, \text{ and}\]
\[x \text{ is instance of Symmetric-conjunction, and}\]
\[\exists y w \text{ is substitute of } y] \rightarrow\]
\[\exists z z \text{ is coordinate of } x \text{ and } w \text{ is substitute of } z]\]

If the preceding coordinatees of a conjunction are its adjuncts, we would expect them to be optional by default. But in symmetric coordination the conjunction must have a preceding coordinatee: \emph{*Find and Sophy, *Large and sausages}. We do not, however, need a rule making a preceding adjunct coordinatee obligatory. Rather, the presence of a preceding coordinatee follows from the combination of (224) and (226). (224) ensures that either \(W\) is a substitute or \(W\) is a member of a set of mutual conjunct mates at least one member of which is a substitute, and (226) requires that if a coordinatee is a substitute then it is substitute of a fellow coordinatee. Therefore at least one of a symmetric-conjunction’s following, complement coordinatees must be substitute of a preceding coordinatee.

When (226) applies to an example where a supervener is itself a coordinatee, as in (227a—d), we get the substitute relationships shown above the line. But we also want the substitute relationships shown below the line to obtain. To achieve this end, we state rule (228), which passes on substitute relationships via supervener relationships:
If X is supervener of Y and X is substitute of Z then Y is substitute of Z.

The counterpart to (226) for inclusive supplementary coordination is (229a) and for exclusive supplementary coordination (229b). The noninitial coordinatee in inclusive supplementary coordination is substitute only of the initial coordinatee. The coordinatee in exclusive supplementary coordination is substitute only of a subordinate of the word the conjunction is advener of.

---

91 In §3.7.3 it is allowed that the coordinatee of an exclusive supplementary can have a supervener so long as the coordinatee isn't a substitute. Taking into account the discussion still to come in §3.7.1 and §3.7.3 (qv.), the rule for inclusive and exclusive supplementaries need to be revised a little, though I am not disposed to present the revision here.
(229) a. $\forall v \ [v$ is instance of [inclusive] Supplementary-conjunction] $\rightarrow$

$[\exists w \ \forall x \ \exists y \ w$ is supervener of $x$, and

$x$ is coordinate of $v$, and

$y$ is subordinate of $x$, and

$\forall z \ [z$ is coordinate of $v$, and

$w$ is not supervener of $z] \rightarrow$

$[z$ is substitute of $y]]$

b. $\forall v \ [v$ is instance of [exclusive] Supplementary-conjunction] $\rightarrow$

$[\exists x \ \exists y \ v$ is advener of $x$, and

$y$ is subordinate of $x$, and

$y$ is not subordinate of $v$, and

$\forall z \ [z$ is coordinate of $v] \rightarrow$

$[z$ is substitute of $y]]$

With inclusive supplementaries the initial coordinatee is obligatory because the other coordinatee must be substitute of a subordinate of the initial coordinatee. This rules out $*_{[either \ or \ [Sophy]]}$. No adjunct coordinatees apart from the initial coordinatee are possible with inclusive supplementaries, and no adjunct coordinatees at all are possible with exclusive supplementaries. This is because if there were such an adjunct coordinatee then some adjunct coordinatee and some complement coordinatee would be substitutes of each other, but that would conflict with (229a)'s requirement that the noninitial coordinatee be substitute only of a subordinate of the initial coordinatee, and with (229b)'s requirement that the coordinatee be substitute only of a subordinate of the word the conjunction is advener of. This rules out $*_{Sophy \ arrived, \ Thalia \ and \ Euphrosyne}$. (229a—b) also have the effect of preventing a coordinatee of a supplementary-conjunction from having a conjunct-mates, because each coordinatee (or, in the case of inclusive supplementaries, each noninitial coordinatee) must be substitute of same word (which is disallowed for conjunct-mates).

For inclusive and exclusive gapping the rules are (230a—b) respectively. In inclusive gapping, noninitial coordinatees are substitutes of a subordinate of the initial coordinatee. In exclusive gapping, coordinatees are substitutes of a subordinate of the word the conjunction is advener of.
As with inclusive supplementaries, the initial coordinatee in inclusive gapping is obligatory because the other coordinatees must be substitute of a subordinate of the initial coordinatee. Adjunct coordinatees apart from the initial coordinatee are optional with inclusive gapping and all adjunct coordinatees are optional with exclusive gapping; cf. Ann would do anything for Beth, (Carl for Dave) and Eleanor for Fred. The reason why adjunct coordinatees are not obligatorily present is that though some complement coordinatee must be a substitute, it needn’t be a substitute of another coordinatee, in contrast with symmetric coordination. The reason why adjunct coordinatees are not obligatory absent is that, in contrast to supplementary coordination, a coordinatee can be a substitute of a fellow coordinatee.

Nothing I’ve said so far accounts for the apparent fact that if a coordination contains a BOTH (or, in the lect Sag et al. (1985) report themselves to have, an EITHER) then it contains no more than two conjuncts. This is not actually a restriction to two coordinatees, for there are four coordinatees in \([\text{give BOTH \{Sophy\} \{tulips\} and \{Edgar\} \{Roses\}]\]. The explanation for this apparent restriction is that BOTHs require the meaning of the coordination to involve two expansions. This correctly predicts that collective interpretations are out with both: !Both John and Mary weigh 20 stone together. If there is a maximum of two expansions then any set of mutual substitutes must have a membership of two, but there may be any number of coordinatees as long as the expansions they’re associated with total two: it follows that if there are more than two coordinatees then the coordinatees must form two sets of mutual conjunct-mates.
3.6.2 Formalization and application of the notion ‘dependency type’

To repeat what was said in §3.3 when the Substitute relationship was introduced, the principal feature of substitutes is that if X is a substitute of Y then the dependencies holding between Y and Z, where Z is either s-head of Y or a subordinate of the s-head of Y but not a subordinate of Y, also hold between X and Z. This is difficult to formalize in first order predicate logic, which is what WG rules use. Nonetheless I shall try to manage it. Roughly speaking, we need to find a way to quantify over dependency types. First, for each basic dependency type we need rules like (231a—c).

\[(231)\]
\[\begin{align*}
 a. & \quad X \text{ is subject of } Y \text{ iff } Z \text{ is ‘subject-set’ of } Y \text{ and } X \text{ is member of } Z. \\
 b. & \quad X \text{ is object of } Y \text{ iff } Z \text{ is ‘object-set’ of } Y \text{ and } X \text{ is member of } Z. \\
 c. & \quad X \text{ is inject of } Y \text{ iff } Z \text{ is ‘inject-set’ of } Y \text{ and } X \text{ is member of } Z.
\end{align*}\]

Alternatively, we can discard Subject, Object, Inject, etc. and only ever make use of Subject-set, Object-set, Inject-set, etc. References in rules to “subject of W” can be replaced by “member of the subject set of W”. This is probably the right way to go, formally, but it’s a bit cumbersome for everyday prose, so I’ll continue to use ‘Subject’, ‘Object’, etc. as abbreviatory terms. The important thing to note here is that (231a—c) and other rules of their ilk don’t add to the number of rules required by the grammar.

Next, we create a class containing a word’s object-set, subject-set, and so on. Following a suggestion by Dick Hudson, I’ll call this class ‘Dependent-set’ ("set

---

92 My practise is to write rules in fairly normal English, but with an implicit guarantee they could be rewritten easily into standard predicate logic notation. (I refer anyone who objects to this rule format as allowing one to write all sorts of rules one would never wish to write, to McCawley’s (1973) critique of objections of that ilk.) There is no such guarantee with WG rules in Hudson’s work, and to the extent that they can be translated into predicate logic form they are not always first order, but nor are they consistent, so I feel entitled to throw Hudson’s rule format overboard. On the whole the reader will either have to verify my criticisms for themself or take them on trust, since I think it would be a waste to devote to a detailed critique of Hudson’s rule format the considerable space it would require, especially since only a little theoretical emphasis is laid (in Hudson 1990) on the format of rules. Not every rule written in standard predicate logic notation can be translated into Hudson’s (1990) notation, which might seem to be a strength of the latter, since it limits the expressive power of the rules, but in response to this I again cite McCawley, (1976): “Makkai’s [1975] aversion to first-order arbitrariness is coupled with a willingness to accept second-order arbitrariness, much in the same way that Immanuel Kant (Kaufmann 1973: 239—41) lived the minutest details of his life according to universal maxims but was quite arbitrary in his choice of maxims to live by (for example, he adopted the maxim that one must not smoke more than one pipe of tobacco per day)” (see also McCawley 1990).

93 It is manageable in first order logic because Word Grammar makes no formal logical distinction between sets and individuals; sets are individuals.
squared”). The motivation for the name is that whereas the members of, say, a subject-set are words, the members of a dependent-set will be sets of words.

\[(232)\]

a. If \(X\) is subject-set of \(Y\) then \(Z\) is dependent-set of \(Y\) and \(X\) is member of \(Z\).

b. If \(X\) is object-set of \(Y\) then \(Z\) is dependent-set of \(Y\) and \(X\) is member of \(Z\).

c. If \(X\) is inject-set of \(Y\) then \(Z\) is dependent-set of \(Y\) and \(X\) is member of \(Z\).

This will now give us a way to state \((233)\), the rule for dependency-type shaving for substitutes that aren’t superveners.\[94\]

\[(233)\]
\[\forall a \forall b \forall c \,[a \text{ is substitute of } b, \text{ and } a \text{ is not supervener of } c] \rightarrow \]
\[\forall d \forall e \forall f \,[\exists g \,[g \text{ is } s\text{-head of } b, \text{ and } d \text{ is unsuperordinate of } g, \text{ and } d \text{ is not unsuperordinate of } b, \text{ and } e \text{ is dependent-set of } d, \text{ and } f \text{ is member of } e, \text{ and } b \text{ is member of } f]] \rightarrow \]
\[[a \text{ is member of } f]]\]

The grammatical constraints that are normally thought of as preventing a word from having more than one subject or complement (or, more generally, non-adjunct) of the same type are in fact requirements that if a word has, say, two subjects then one subject must be substitute of the other. We want to state this rule as generally as possible, without having to state it separately for subjects, for objects, for injects, et cetera et cetera. Our rule requires us to define some additional sets of sets of words, analogous to Dependent-set: Adjunct-set and Complement-set. We thus have rules like \((234a-b)\). A

\[94\] [This rule is unnecessary under my now preferred analysis whereby coordinatees in of the conjunct-ion heading coordination \(X\) embedded within coordination \(Y\) are coordinatees of the conjunct-ion heading \(Y\).]

\[95\] A reminder: \(X\) is unsuperordinate of \(Y\) iff \(X\) is \(Y\) or \(X\) is subordinate of \(Y\). \(X\) is subordinate of \(Y\) iff \(X\) is \(s\)-dependent of \(Y\) or \(X\) is \(s\)-dependent of a subordinate of \(Y\).
similar set of rules would define Adjunct-set, but I can't give them since noone yet
knows what varieties of adjunct there are.

(234) a. If X is object-set of Y then Z is complement-set of Y and X is member of 

Z.

b. if X is inject-set of Y then Z is complement-set of Y and X is member of 

Z.

We can simplify matters a bit by scrapping (232b–c) and replacing them by (235).

(235) C is complement-set of X and D is dependent-set of X and Y is member 
of C iff Y is member of D.

And now we can state the generalization.

(236) \( \forall d \forall w \forall a \forall x \forall c \) [d is dependent-set of w, and a is adjunct-set of w, and 
x is member of d, and x is not member of a, and 
c is cardinality of x, and c is greater than 1] ⇔

[\( \forall y \) [y is member of x] →

[\( \exists z \) [z is member of x], and

[y is substitute of z, or z is substitute of y]]]

If a word has multiple dependents of the same variety, this rule doesn't require a sub­
stitute relationship to hold between each pair, since it is possible to say *She swims more 
than Sophy more than Edgar* (meaning “She swims more than Sophy more than Edgar
swims more than Sophy”), in which *swims* has three subjects, *she, Sophy* and *Edgar,*
and *Sophy* and *Edgar* are both substitute of *she* but neither *Sophy nor Edgar* is substitute
of the other.

The workings of (233) require that in (237), which for present purposes I shall sim­
ply assume has the dependency structure shown, both *glittering* and *(much-)admired* are
adjuncts of *gems.* However, it correctly does not require that *gems* be object and inject
of *glittering* as well as of *(much-)admired.* If W has conjoined s-dependsents, X and Y,
then the same dependencies that go from W to X (where W is the head argument of the

[139]
dependency and X is the dependent argument) must also go from W to Y, but there is no requirement that all the dependencies that go from X to W also go from Y to W.

(adjunct of subject of substitute of substitute of adjunct of subject of object of inject)

(237) glittering and much-admired gems

The same remarks apply to (238), which is okay because is and can are both clausals of where, even though where is subject of is but extractee of stay.

(of k of g of e of)

(238) I wonder where is a good place to visit and we can stay.

(233) applies vacuously to shared s-dependents of conjoined words, and to conjoined roots of a macrosyntagm. For example, in young boys and girls, boys and girls are substitutes of each other, but there is no word, Z, such that boys/girls is a dependent of Z and Z is not a subordinate of boys/girls.

3.6.3 Wards and dependency sharing

It is substitute relationships that are responsible, by rule (233), for coordinatees having the same kind of dependency, but this doesn’t, and shouldn’t, cover wards. For wards, we need a rule saying roughly that some variety of dependency by virtue of which the ward is s-dependent of one coordinatee is the same as a variety of dependency by which

---

96 Dick Hudson queries this judgement, offering in support the oddity of (i).

(i) ? I wonder who they invited and didn’t come.

To me, examples like these are at worst highly infelicitous. In colloquial speech, I don’t think (ii) would be found exceptionable.

(ii) I wonder who they asked to come but didn’t.

(iii) I wonder what what she was reading and then caused her to forget to get off the tube at the right station.

(iv) the book that caused such a fuss and several bishops tried to ban

However, even if I am right that these are all grammatical, I still fail to explain (a) why these examples aren’t wholly felicitous, and (b) why glittering and much-admired gems is. An explanation for (b) is afforded by the analysis in §6.4, according to which the adjuncts of gems would in fact be non-finite auxiliaries (whose complements are instances of the lexemes GLITTER and ADMIRE). Gems would be both subject and object of each of these auxiliaries. We could then propose that for complete felicity, not only must the same dependencies go from the s-head to the mutual substitutes, but also the same dependencies must go from the mutual substitutes to their s-head. Perhaps this effect can be attributed (with a certain degree of wandwaving) to processing.
the ward is s-dependent of the other coordinatees. Take (239a). She is subject of seemed and object of seemed, but only subject of laughed. This is permissible, because she is subject of both words. But (239b) is ungrammatical because there is no variety of dependency holding between the ward and each coordinatee.

(239) a. She laughed and seemed happy.
   b. *Chocolate either he loves or delights her.

The rule for this is very difficult to formulate. For each variety of dependency it requires rules like (240a–c), which define sets of sets of words (a set ‘squared’), and rules like (241a–c), which define a set of sets of sets of words (a set ‘cubed’). The terminology is selected for perspicuity rather than elegance.

(240) a. If X is subject-set of Y then X is member of Subject-set²
   b. If X is object-set of Y then X is member of Object-set²
   c. If X is inject-set of Y then X is member of Inject-set²

(241) a. Subject-set² is member of Dependent-set³
   b. Object-set² is member of Dependent-set³
   c. Inject-set² is member of Dependent-set³

We can then formulate the rule for wards as (242). I don’t think it could be made any simpler, and I don’t think it’s all that complicated.97

---

97 As far as I’m aware, no previous attempt has been made in WG to formalize rules of dependent and head sharing in coordinate structures. Hitherto, all statements have been informal.
3.7 Asyndetic coordination

What I hope I've achieved in this chapter is to show (i) that coordination, including various of its more recalcitrant forms, such as gapping, can be accommodated within a model of grammar in which continuity is achieved by a minimal constituent structure of unlabelled bracketing, and in which s-dependency, which is relevant to a number of kinds of grammatical rule, notably rules of word order, is defined partly on the basis of constituent structure but also partly on the basis of certain configurations of dependencies that obtain in coordinate structures, and (ii) that this analysis works better than some of its likely opponents, notably Hudson's analysis of coordination or the modified ('MWG') version of it. Hudson has from time to time claimed that one of the virtues of dependency is that it leaves constituency free to handle coordination, but I think it is preferable, and more in the spirit of the main analytical thrust of WG, that the range of devices employed by the grammar should be as uniform as possible, with — on my view — a maximally minimal constituent structure to which the grammar makes next to no reference at all, and with interword relationships doing the lion's share of the work.

Certain of what are commonly taken to be varieties of coordination have not so far been discussed. Right-node raising, and apparent examples of word-internal coordination are treated in §5.3, where it is argued that rather than being varieties of coordination, they result from an unorthodox correspondence between syntax and morphology. This leaves asyndetic, as in (243a—d), as the only major type of coordination that remains undiscussed. This could be handled easily by the MWG analysis, with the structure shown in (244), which differs from standard coordination in nothing but the omission of the conjunction.
(243) a. She ate chocolate, pears, peas, marshmallows.
b. She laughed, waved, smiled.
c. She loves, cherishes, adores him.
d. Give Edgar tulips, Sophy roses, Thalia dandelions.

(244) She ate \{<chocolate>, <pears>, <peas>, <marshmallows>\}.

It is not quite so obvious how TDG might go about handling it, so I'll finish the chapter by showing how TDG handles it. It turns out that TDG handles it pretty straightforwardly and without much room for doubt about alternative analyses. It should be noted, though, that the various rules presented in earlier sections don't take asyndetic coordination into account; the following is not integrated with what has been said above.

Asyndetic coordination is like symmetric coordination without the conjunction. You don't get asyndetic gapping, (245a), or supplementary coordination, (245b), but you do get asyndetic complex, (245c), and simple coordination involving wards, (245d). Not surprisingly, you don't get correlatives either, because these need a conjunction as their complement; cf. (245e-f).

(245) a. If Ann wished to meet Beth, *(and) Carl, Dave, then they should have planned accordingly.
b. He wished to meet Sophy forthwith, *(and) Euphrosyne.
c. He'd give Sophy roses, Thalia thistles, Euphrosyne tulips, day in day out.
d. She photographs newly-hatched sparrows, starlings, thrushes, finches, swallows in her ornithological researches.
e. * She met both Sophy, Edgar yesterday.
f. * She'll meet either Sophy, Edgar, Thalia, Euphrosyne tomorrow.

Inspection of any example of asyndetic coordination, e.g. (246), suggests that, as in symmetric coordination, the root of each conjunct is either substitute, or, in examples like (245c), conjunct-mate of each other conjunct root.

(246) If I, he, she, you each found ourself sitting opposite one another we'd yack till the ends of time.
(245d) can receive an interpretation in which *newly-hatched* modifies each of sparrows, starlings, thrushes, finches and swallows. The only way this can happen is if *newly-hatched* is a ward. If there were a conjunction present, *newly-hatched* would be ward of the conjunction, but since there is no conjunction present, it must be ward of something else. Since wards can occur in asyndetic matrix coordination, e.g. *She laughed, danced, rejoiced*, the ward must be ward of one of the conjunct roots. Since only superveners can have wards, some of the conjunct roots must be superveners of other conjunct roots. That is how the ward can get to be dependent of each conjunct root: if X is ward of Y and Y is supervener of Z, then X is s-dependent of Z. It seems simplest to assume that each conjunct root is supervener of each of its substitutes, which gives us then the structure shown in (247).

![Diagram](image)

Young sparrows, starlings, thrushes, finches, newly hatched from their shell were filmed.

We could cope with (248a), which, according to Hudson (1989) is acceptable for many speakers (though it must be rather limited, for (248b) seems much worse), by analysing it as involving asyndetic coordination. (248a–b) are MWG-style, and (248c) is TDG. The variation between speakers in their acceptability judgements of (248a) is explicable as variation as to which grammatical contexts asyndeton is tolerated in.
(248) a. % John {<has {<finished his thesis>,
    <taken it to the binders}>},
    and <is impatient for the viva>}. 

b. *? John {<should {<have {<shaking hands>},
    <kissing babies}>},
    <gone onto the rostrum}>,
    <vow never to make the same mistake again>},
    and <will always be remembered for the disaster that ensued>}. 

However, treating these as partially asyndetic makes some fairly bad predictions. If asyndetic coordination is always interpreted as conjunctive rather than disjunctive, then (249a) ought to mean the same as (249b); but it doesn’t. (249c) ought to be acceptable, since and supposedly has only two coordinatees, have and be. (249d) ought to be ungrammatical even for speakers who accept (248a), since either it has no complement, or if its complement is or then if the coordinatees of or are have and be there’s discontinuity, and if the coordinatees are finished, taken and be, then they’re not properly substitutes of each other.

(249) a. John will either have finished his thesis, taken it to the binders or be impatient for the viva. 

b. John will either have finished his thesis and taken it to the binders or be impatient for the viva.

c. John will both have finished his thesis, taken it to the binders and be impatient for the viva.

d. * John will have either finished his thesis, taken it to the binders or be impatient for the viva.
I suspect that the true story is not that these involve asyndeton but rather that they are regular symmetric coordination, but in certain lects the rules specifying what coordinatees must be substitutes of are rather less stringent than in Standard English. So, in (247d) the coordinatees of or would be finished, taken and be, but be is substitute not of the other coordinatees but, instead, of have, which its supervener, either, is advener of.
4

Conjunctions

and s-dependency loops

4.1 S-dependency loops and coordinacy

I'll begin by recapitulating some of the main lessons of Chapters 2–3. The grammar generates a tangled web of syntactically significant pairwise associations between words in the sentence. Most of these associations are dependencies, a dependency being an interword association that can sometimes be paralleled by an s-dependency. Some of the dependencies in the sentence are paralleled by an s-dependency. It is via s-dependencies that the relative linear ordering of heads and their dependents are stated; words X and Y are ordered with respect to each other if and only if they're linked by an s-dependency.

The s-dependency structure of a sentence corresponds to a skeletal constituent structure, whereby most, but sometimes not all, s-dependents of W are nieces of W. It is this skeletal constituent structure that effects global word order requirements of continuity.

The need for skeletal constituency over and above s-dependency is demonstrated by the existence of s-dependency loops. Because there are s-dependency loops, continuity cannot be effected by a prohibition against s-dependencies tangling (which includes looping).

So far, we've seen s-dependency loops only in coordination. In this chapter I'll be arguing first that coordination is not the only place where we find s-dependency loops, and, second, that even if 'coordination' is understood to mean 'structure involving a non-subordinating conjunction', coordination would be very far from being the only place where we find s-dependency loops, were it not that the class of (non-subordinating) conjunctions is much larger (I'll be contending) than is commonly thought.
The types of dependency introduced in Chapter 3 to handle coordination, i.e. Supervener, Advener, Ward, Coordinatee, Substitute and Conjunct-mate, do two different kinds of job. On the one hand, they provide the apparatus that allows s-dependency loops. On the other hand, they provide the apparatus that allows 'coordinacy' — i.e. when two words are in the same type of dependency relationship with a third. It is coordinacy that is normally taken to be the defining property of coordination: two parallel structures that could be expanded into two sentences are conflated into one.

The six dependency types listed above are not equally responsible for s-dependency loops and for coordinacy. Where we find s-dependency loops we will find superveners, and only where we find superveners will we find adveners or wards. Where we find coordinacy we will find substitutes, and vice versa, and only where we find substitutes will we find conjunct-mates. As for coordinatees, these occur only wherever there are conjunctions. I'll reprise this in §4.3, but in the meantime, (250a—e) summarize.

(250) a. If X has two s-heads then X has a supervener or is ward of a supervener.
   b. If X is an advener then X is a supervener.
   c. If X has a ward then X is a supervener.
   d. If X is a conjunct-mate, then X or a conjunct-mate of X is a substitute.
   e. X has a coordinatee iff X is a conjunction.

In symmetric coordination we find both s-dependency loops and coordinacy; the coordinatees both are substitutes and have a supervener. But s-dependency looping and coordinacy do not always cooccur. In exclusive gapping and supplementary coordination there is coordinacy but no s-dependency looping; therefore the coordinatees are substitutes but have no supervener. And, in contrast, §4.2 concludes that in pied piping there is s-dependency looping but no coordinacy: pied piping therefore involves superveners but not conjunctions or substitutes.

---

98 It is purely for ease of exposition that I'm calling Supervener, Substitute and Conjunct-mate varieties of dependency. In fact, if the defining feature of a variety of dependency is that it is sometimes paralleled by an s-dependency, then Supervener, Substitute and Conjunct-mate are not varieties of Dependent, but only of a more general kind of interword relationship that we might call 'Associate'.

99 This statement contradicts the TDG analysis in §3.5.4 and §3.5.6, which allows exclusive gapping and supplementary conjunctions to be adveners. However, there is no impediment to analysing exclusive conjunctions as adjuncts of W (and s-dependents of W, and therefore not extractees of W) rather than adveners of W. This step then preserves the generalization that only superveners can be adveners.
In §§4.3—6 we see further examples of conjunctions that involve either s-dependency looping or coordinacy but not both. I’ll also be arguing that there are conjunctions belonging to lexemes other than AND, OR and BUT — i.e. other than those lexemes commonly recognized as comprising the conjunctions. And once we’ve established that, we’ll see that there are further kinds of conjunction, with properties no AND, OR or BUT ever possesses. The dependency structure of these conjunctions tends to be a combination of its valency with some or all of the six coordinational dependencies.

The picture that emerges is that coordination is not all that special. A conjunction’s valency combines with the dependencies it gets through its conjunctionhood. Instances of a wide range of lexemes can be conjunctions. S-dependency loops are found not only with conjunctions. Thus, the less discrete from the rest of syntax we find coordination to be, and the more we find structures partially resembling coordination, the more we should favour the TDG analysis of coordination over, say, Hudson’s (1984, 1988, 1990) approach, where coordination is utterly different from the other devices of syntax which by comparison are relatively homogeneous. Not only is TDG preferable to its rivals because it handles the facts about coordination better, but also it is preferable because an analysis in which there is no great gulf between the devices used to analyse coordination and the devices employed for all other constructions can cope better with constructions that resemble coordination.

4.2 Pied piping

§4.2.1 exposes a fatal flaw in the analysis of pied piping in Rosta (1994), and sketches the beginnings of a remedy. §4.2.2 argues that Hudson’s (1990) analysis of pied piping is no remedy. §4.2.3 presents a solution to the problems of pied piping, and shows that, like coordination, this construction involves s-dependency loops. §4.2.4 shows how the analysis covers main clause pied piping.

4.2.1 A problem with the Rosta (1994) analysis of pied piping

In §2.6.1 I propose for pied piping the partial associacy structure shown in (251a—b).

\[
\begin{array}{c}
\text{libraries in nooks of which she dozed}
\end{array}
\]
The analysis of pied piping in Rosta (1994) is essentially in agreement with this, but differs in that it treats proxies as dependents and potential s-dependents, giving the analysis shown in (252).

There are at least three objections to this analysis. First, it is an exception to the otherwise exceptionless rule that complements of a preposition P are either s-dependent of P or are promoted or extracted. In (252) which is complement of of but is neither s-dependent of of nor extracted or promoted. Second, no exceptions to the rule that proxy associacies are not dependencies are provided by any of the other cases where I have discerned a role for proxies — the examples in §2.6.2-4, and either as proxy of or in either him or her (see §3.5.4).

The first two objections are only weak evidence against (252), but they do suggest that it is in some respects wanting. However, the third objection constitutes very strong evidence against (252). When applied to (253), the analysis applied to (252) leads to tangling s-dependencies.

In (253) the s-dependency from gift to to crosses the s-dependency from flowers to which. If we use Hudson’s (1994b) notation for the same structure, as in (254), there is apparently even more tangling. (Due to the limitations of my word processing software I must render his arcs rectilinearly.)
Here we see the gift-to s-dependency tangling not only with the which-thought but also with the flowers-which and which-her.

The s-dependency structure assumed by Rosta (1994) therefore cannot be correct. What should the s-dependency structure be, then? Suppose thought were s-dependent of her rather than of which, as in (255a–b) (the same s-dependency structures shown with the two notations). On its own, this change doesn’t help: there is no longer a tangle of which-thought with gift-to, but flowers-which still tangles with gift-to and now tangles with her-thought.

Similarly, if we adapt (253)/(254) by making her rather than which s-dependent of flowers, tangling remains, as (256a–b) show. We lose the tangling of gift-to with flowers-which and which-her, but gift-to continues to tangle with which-thought.

However, if we combine the revisions leading to (255) and (256) this will give us the tangle-free structure in (257a–b).
The s-dependency structure of (257a–b) requires the presence of additional dependencies that have not yet been independently motivated. These are shown in (258) labelled ‘?’.

What are these mystery dependencies? We could start by turning to Hudson’s (1990) analysis of pied piping, which would give the dependency/associacy structure shown in (259).

But this analysis in which her is adjunct of flowers, and thought is complement of her is open to a range of criticisms, which I detail in the following section.

4.2.2 Proxies versus ‘projections’

The relation ‘proxy’ is intended to supersede the relation ‘projection’, introduced in Hudson (1990) to handle constructions like pied-piping and so clever a person. The two relations are very similar, but there are two important differences. First, the projection of W takes over W’s role as a dependent. So in (259), her (instead of which) is adjunct of flowers. Second, the projection of W takes over W’s role as a head. So in

---

100 Proxies, and also, to a greater extent, projections, work in a similar way to Generalized Phrase Structure Grammar’s foot features (Gazdar et al. (1985)), which percolate up the tree from non-head nodes.

101 In Rosta (1994) I said that proxies differed from projections in that proxies of W are dependents of W while projections of W aren’t. I now think that proxies are like projections in this respect.
(259), thought is complement of her (instead of of which). I'll give arguments against each of the two properties in turn.

We'll deal first with the way that, in the projection analysis, in wonder/person with whom she went, it is not whom but with, the projection of whom, that is the complement of wonder or the adjunct of person. One problem with this is that the syntax is completely out of step with the semantics. It is the meaning of whom, not with, that has the semantic role of wonderee and that (if it is nonrestrictive) is coreferential with person; Hudson's analysis would require a possibly unmanageably extreme mismatch between syntactic and semantic structure. Another problem, which Hudson notes, is that in his analysis, in talk about how often she went or talk about how big it is, the complement of about is often and big, which are the projection of how; however, the complement of a preposition is usually required to be a noun, rather than an adverb or adjective, and Hudson would have to exceptionally allow adverbs and adjectives to be complement of a preposition, so long as they are projections of a how. In contrast, in the analysis I offer, it would be how, a noun, that is unexceptionally the complement of about.

We turn now to arguments against the way the projection of W takes over the valency of W, so that, for example, in wonder on what she sat, on, being the projection of what, has a clausal, sat. First, there is the difficulty of formulating an appropriate rule (Hudson (1990) does not attempt to). One might try (260). But this overgenerates. Assuming that projections are defined in the same way as proxies (as described in §2.6.1), every head of a relative wh-pronoun W is a projection of W and every head of every projection of W is a projection of W. Therefore in (261a) every italicized word is a projection of which. Each of the words in boldface is both an extractee of some word and a projection of which. This wrongly predicts the clausal dependencies illustrated by coindexing the extractee and its clausal. If such clausal dependencies obtained then in principle weird word orders such as that in (261b) would be possible unless ruled out by sundry other conditions.

(260) If X is extractee of Y, and X is projection of Z then Y is clausal of X.

(261) a. Who, did, she say, he was, hard, to, persuade, to, put, up in the house of someone who, they say, keeps, dogs a, fair proportion of which bite?

[153]
Any more adequate rule than (260) would very probably require novel types of dependency to make it work. (Note too that we haven’t even dealt with a rule for ‘pied piped’ subjects, as in someone the guardian of whom munches asphodels.)

We can start to investigate further problems with the projection of W inheriting W’s valency by comparing the proxy-based and projection-based analyses in (262a—b): in (262a), that is a complement of so and extraposes across so’s proxy,\footnote{The fixed order of alleged extraposees in (264—265) casts doubt on the analysis of them as extraposees. If they were extraposees, they ought to be able to occur in any order. The fixedness of their ordering suggests the bracketing in (i) with a dependency structure to match. (i) \textit{She was [so [much [more [sad [that he died]] [than me]] [that she wept]]}}

Technically, this would invalidate this paragraph’s argument against projections, but the essence of the argument hold good, even if the data chosen to illustrate it proved to be inappropriate.} while in (262b) that is a complement of so’s projection, and there is no extraposition.

\begin{itemize}
  \item \textbf{a.} The book was so poignant that she wept.
  \begin{itemize}
    \item \textit{of } \textit{X}
    \item \textit{of } \textit{E}
    \item \textit{of } \textit{e}
    \item \textit{of } \textit{prox}
  \end{itemize}
  \item \textbf{b.} The book was so poignant that she wept.
  \begin{itemize}
    \item \textit{of } \textit{X}
    \item \textit{of } \textit{c}
    \item \textit{of } \textit{prox}
  \end{itemize}
\end{itemize}

The proxy analysis requires that the complement of a so be the extraposee of the word the so is an adjunct of, while the projection analysis requires that it is not a so but a projection of a so that subcategorizes for a complement that is proxy of a finite. If we complicate the construction slightly, to give (263a—b), we see that the proxy analysis is the better one.

\begin{itemize}
  \item \textbf{a.} She was so sad that he died that she wept.
  \begin{itemize}
    \item \textit{of } \textit{X}
    \item \textit{of } \textit{c}
    \item \textit{of } \textit{prox}
    \item \textit{of } \textit{e}
  \end{itemize}
\end{itemize}
Under the proxy analysis, the general rule that complements precede other dependents will explain why the order of the two *that*-clauses can’t be reversed. The projection analysis would have to stipulate the order – that’s the first problem. In order to make this stipulation, the two alleged complements of *sad* have to be of distinct types. They also have to be of distinct types in order to preserve the apparently exceptionless rule that a word cannot have more than one complement of the same type (coordination apart), and in order to ensure that the meaning of each complement is attached as an argument to the right predicate (the meaning of *so* and of *sad*). This is the second problem, for it would require the ad hoc introduction of new kinds of complement, whereas all other evidence suggests that there are (in English) only a handful of types of complement, not all of which have properties compatible with the complements in (263a–b). This second problem is even more evident in (264a–b), which I also adduce to illustrate a third problem with the projection analysis. (264a) is the proxy analysis and (264b) the projection analysis.
Here, *more* is also behaving like *so* in terms of how the competing analyses would handle them. Note how in (264b) *sad* has three complements: *that (he died)*, *than* and *that (she wept)*. We can see the third problem\(^\text{103}\) with the projection analysis by asking how and why *than* and *that (she wept)* end up as extrapoloses or complements of *sad*. In the proxy analysis all we need is a rule that makes each complement and extrapolose of \(W\) an extrapolose of \(W\)'s proxy, where \(W\) is *so* or *more*. In the projection analysis it is the projection of each of these words that subcategorizes for the complement: since *so* has three projections (*much*, *more* and *sad*), this should mean that *that (she wept)* is a complement of each of these three projections. These unintended extra complement relations are shown in square brackets in (264b). It should be noted that constructions like (264a—b) can get much more complicated (and hard to process). In (265), for example, there are four words behaving like *so*: *too*, *more*, *differently* and *easy*. (The diagram is intended just to guide the reader in parsing this monstrosity.)

\[
\text{(265) \quad Hers was too much more differently easy a book to read from yours than his to bear.}
\]

Under the projection analysis, *a* has five complements (*book*, *to (read)*, *from*, *than* and *to (bear)*), *easy* has four complements (*to (read)*, *from*, *than* and *to (bear)*), *differently* has three complements (*from*, *than* and *to (bear)*), *more* has two (*than* and *to (bear)*), and *too* and *much* each have one (*to (bear)*). Under the proxy analysis, the only complement relations there are correspond to the lines in the diagram.

### 4.2.3 Pied piping with s-dependency loops

We have established that pied piping involves the associacy structure shown in (266a), and the constituency structure shown by stemma in (266b). We also require the s-dependencies shown in (266c): they will preserve the generalizations that dependents of \(W\) are s-dependents of \(W\) unless they are 'moved' (extraposed, extracted, promoted, etc.). Adding together all of these gives us the overall pattern in (266d).

---

\(^{103}\) To be fair, it is more a weird and unforeseen consequence than a problem.
The only remaining step is to complete the associacy structure (266d) in such a way as to yield the requisite s-dependency structure. The single addition we need is provided by rule (267).

(267) If X is a wh-pronoun, and Y is clausal of X, and Z is proxy of X, and [Z is subject of Y or Z is extractee of Y] then Z is supervener of X.

The Supervener associacy has the special property of licensing s-dependency loops — it allows s-dependencies to tangle in specified ways. The permissible tanglings arise from rules stating (268a–b).

(268) a. If X is supervener of Y and Z is ward of X then Z is s-dependent of X and Z is s-dependent of Y (but Z exceptionally isn’t niece of Y).

b. If X is supervener of Y and X is advener of Z then X is s-dependent of Z and Y is s-dependent of Z (but Y exceptionally isn’t niece of Z).

Any supervener can be an advener and have a ward.
We end up with the associacy and s-dependency structure illustrated in (269a—b).\footnote{A construction which shows how compelling this analysis is is provided by a challenging example from Pollard & Sag (1994: 69), and given here in (i). Note that it can be complicated to (ii).} In the diagram labels of dependents of W that are also s-dependents of W are italicized and labels of (s-)dependents of W that are also nieces of W are capitalized. The single lines in the stemma represent the part of s-dependency structure that corre-

\begin{itemize}
  \item[(i)] Whoever’s dogs are in my garden is in big trouble.
  \item[(ii)] Whoever’s children’s dogs are in my garden is in big trouble.
\end{itemize}

The problem is that we can tell from the form of the auxiliary (are/is) that ’s (dogs) is subject of are but whoever is subject of are. In free relatives like Whoever is in my garden (is in big trouble), Hudson’s (1990) analysis of relative clauses works fine: whoever is the root of the relative clause, the structure being [Whoever [is in my garden]] is in big trouble. But in (i—ii) whoever is subordinate of ’s (dogs). There therefore cannot be a projective subject dependency from is to whoever. Hudson’s (1990) ‘projection’ analysis of pied piping, does not deal with this construction, but if that analysis were applied to this construction, ’s (dogs) would be projection of whoever and subject of is, and hence we would have a plural pronoun, ’s (dogs), agreeing with is, which is surely in incorrect result. But under my analysis the problems evaporate. The structure is shown in (iii), with s-dependencies that are branch-dependencies shown capitalized, italicized and in boldface, s-dependencies that aren’t branch-dependencies shown just capitalized and italicized, dependencies that aren’t s-dependencies shown just capitalized, and associacies that aren’t dependencies shown in ordinary lowercase.

\begin{itemize}
  \item[(iii)] Whoever ’s children ’s dogs are ... is ...
\end{itemize}

As the diagram shows, ’s (dogs) is proxy of whoever and therefore can be supervener of whoever. Because ’s (dogs) is a supervener, it can be advener and have a ward. Because it is advener of is and supervener of whoever, whoever can be subject and non-branch-dependent s-dependent of is, in accordance with the agreement form of is. The subject of are is the plural pronoun ’s (dogs), again in accordance with the agreement form. Since are is ward of ’s (dogs) and ’s (dogs) is supervener of whoever, are can be complement and non-branch-dependent s-dependent of whoever.

On the matter of ’s (dogs) being proxy of whoever, it has already been shown that if a relative pronoun (or a proxy of a relative pronoun) is inject of X then X is proxy of the relative pronoun. Pied piping in free relatives is possible only with -ever pronouns, and only with this particular route to proxyhood, as (iv—v) show, though, as (v) shows, in -ever clauses of ‘logical independence’ (with an “irrespective of” meaning), other routes to proxyhood are possible.

\begin{itemize}
  \item[(iv)] * The manners of whoever are reprehensible is not to be allowed entrance to the club.
  \item[(v)] Whoever’s manners are reprehensible is not to be allowed entrance to the club.
  \item[(vi)] In however many pockets you stash the odd five pound note, you can never be sure of having enough dosh on you.
\end{itemize}

A similar constraint to that on pied piping in free relatives is found with interrogative subjects, as (vi—vii) show.

\begin{itemize}
  \item[(vi)] (I wonder) whose mother(’s personal guru) arrived late
  \item[(vii)] * (I wonder) (the personal guru of) the mother of whom arrived late
\end{itemize}

An apparent defect of my analysis is that it appears not to explain why the relative and interrogative pronoun in free relatives and subject interrogatives must be at the extreme left of the clause. A solution that remedies this defect, yet leaves the analysis of pied piping unaltered, is provided in §5.4.1.
sponds to constituent structure, and the double lines represent s-dependencies that do not correspond to constituent structure.

(269) a. wonder with whom to dine

(269a—b) have strong parallels with a coordination structure like invite friends and family of the bride, as can be seen by comparing (269a—b) with (270).

I conclude this section with diagrams of further varieties of pied piping that haven’t been considered so far, but which are handled unproblematically by the analysis. On whose as the morphological/phonological realization of a sequence of instances of the lexemes WHICH (or WHO) and IT-THE, see §6.1.3.
(271) a. book whose title she liked best.

b. % books the which she read

c. % She discussed sunsets, pictures of the which she exhibited endlessly.

d. She forgets titles the initial letter of whose initial word is Z.
4.2.4 Main clause pied piping

To complete the analysis of pied piping we should consider main clause pied piping, as in (272), which has the indicated dependency and s-dependency structure.\(^\text{105}\)

\(^{105}\) Shortage of space prevents any exposition or justification of the analysis of 'subject-auxiliary' inversion.
There are no grounds for taking \textit{should} to be clausal of \textit{whom} or \textit{whom} to be dependent of anything but \textit{with}. The fact that the syntactic root of the sentence is \textit{should} but the ‘semantic root’ of the sentence is \textit{whom} (in the sense that the sentence is a question about the identity of the referent of \textit{whom}) does not motivate a dependency between the two words, since wh-questions don’t require extraction: cf. \textit{(And) she should dine with whom?} (which, N.B., needn’t be an echo question). If (272) worked like subordinate clause pied piping, then the root of the sentence would, counterintuitively, be \textit{with}, but since \textit{should} is not clausal of \textit{whom}, rule (267) fails to apply and consequently \textit{with} is not supervener of \textit{whom} and therefore there are no s-dependency loops. As for how the difference between subordinate and main clause interrogative pronouns should be captured, see §5.4.1.

It might seem that (273a) is a counterexample to my denial that main clause pied piping works the same way as subordinate clause pied piping. If (273a) were taken to be a parentheticalized counterpart of (273b) then we might indeed expect (273a) to have the structure shown in (273c); \textit{with} would be extractee both of \textit{should} and of \textit{wonder}. That is, \textit{wonder} is not the root of the sentence, for \textit{should} is not subordinated to \textit{wonder}, so if \textit{whom} is complement of \textit{wonder} then the way for \textit{wonder} to not be the root is for it to be subordinated to \textit{with}, as in subordinate clause pied piping. Put yet another way, if \textit{whom} is complement of \textit{wonder}, it’s hard to see any alternative to (273c).

\begin{itemize}
  \item[(273) a.] With whom, I wonder, should I get in touch?
  \item[(273) b.] I wonder with whom I should get in touch.
  \item[(273) c.] [With [whom], [I wonder], [should [I] [get in touch]]]?
\end{itemize}

However, I conclude from (274a), in which there is by all appearances no interrogative, that \textit{wonder} can have a complement that is a finite instead of an interrogative, so long as (i) such a finite complement is also its extractee, and (ii) such a \textit{wonder} is an unsupervenial of a parenthetical, and possibly (iii) the parenthetical is parenthetical of an ‘inquisitive’ (question-asking) finite.\textsuperscript{106} (On parentheticals, see §2.3.3.) In this case,

\begin{itemize}
  \item[(i)] He wonders that she should be so forward.
  \item[(ii)] She’ll be there already, I shouldn’t wonder.
\end{itemize}

\textsuperscript{106} I am talking about the \textit{W}\textit{ONDER} whose sense is, roughly, “ask about”, not the \textit{W}\textit{ONDER} exemplified in (i), whose sense is, roughly, “be surprised at”. I assume that it is the latter \textit{W}\textit{ONDER} that appears as the parenthetical in (ii).
(274a) can be taken to have the structure in (274b), and (273a) can be taken to have essentially the same structure, as in (274c).

(274) a. Is she, I wonder, a credible witness?
    b. [Is [she], [I wonder], [a credible witness]]?
    c. [[With [whom]], [I wonder], should [I] [get in touch]]?

4.3 A taxonomy of conjunctions
In Chapter 3 we met three kinds of conjunction: symmetric, gapping and supplementary. I’ll here be introducing two more: ‘unary conjunctions’, and ‘subjunctions’. The relationship between them is as follows.

- Every subjunction is a supplementary.
- Every supplementary is a unary.
- Every unary is a gapping.
- Every gapping is a conjunction.

As for symmetries, it should prove possible to treat them as just conjunctions that aren’t gapping-conjunctions. These will all be introduced and explained in due course, but here are some preliminary examples to orient the reader. The conjunction is italicized, the coordination is in curly brackets, the conjuncts are in pointy brackets, and the independent phrases within conjuncts are in square brackets.

(275) a. **Conjunction (symmetric, not gapping).**
   \{< [I chose chocolate] >, < [Ellen chose strawberry] > and < [Lucien chose vanilla] > \}.

    b. **Gapping (not unary).**
   inclusive: \{< [I chose chocolate] >, < [Ellen] [strawberry] > and < [Lucien] [vanilla] > \}.
   exclusive: I chose chocolate, \{< [Ellen] [strawberry] > and < [Lucien] [vanilla] > \}.

    c. **Unary (not supplementary).**
   Sophy could at least do your portrait, if not as easily \{as < [Edgar] [mine] > \}.  

[163]
d. **Supplementary (not subjunction).**

Assam will do me fine, {or < [darjeeling] > }.

e. **Subjunction.**

a {not < [very] } long article

The properties a word acquires through being a conjunction are that it has coordinatees. By default a conjunction can have any number of coordinatees. And also by default, a coordinatee can have a conjunct-mate and a supervener.

Gapping conjunctions differ from default (symmetric) conjunctions in terms of which of their coordinatees have a supervener: something like rule (276) applies.

(276) If W is coordinatee of a gapping-conjunction then W cannot both have a supervener and be a substitute.

In addition, for exclusive gapping, none of the coordinatees have a supervener, while for inclusive gapping only the initial coordinatee has a supervener. Alternatively, one could get the same results by requiring each coordinatee of exclusive gapping to be a substitute, and all but the initial coordinatee of inclusive gapping to be a substitute.

Unaries differ from gappings in that they allow only complement coordinatees: that is, they can have only one conjunct,\(^\text{107}\) but there can be multiple coordinatees so long as they’re conjunct-mates of each other.

Supplementaries differ from unaries in that their coordinatees can’t be conjunct-mates, which has the result that they can have only one coordinatee.

Finally, subjunctions differ from supplementaries in that their coordinatee has a supervener.\(^\text{108}\) Given (276), it follows that since a subjunction’s coordinatee has a supervener, the coordinatee can’t be a substitute. Subjunctions therefore give rise to s-dependency loops but not coordinacy.

\(^{107}\) This is why I’ve called them 'unary'.

\(^{108}\) I’ve called them 'subjunctions' because they’re conjunctions and comprise a large number of what Quirk et al. (1985) call ‘subjuncts’. The term is not an allusion to subjunction in Dependency Phonology (Durand 1990) which allows (contrary to the tenets of most versions of Dependency Grammar) the head of a terminal node to be a non-terminal node.
(277) summarizes.

<table>
<thead>
<tr>
<th>CLASS</th>
<th>PROPERTY DISTINGUISHING IT FROM SUPERORDINATE</th>
</tr>
</thead>
<tbody>
<tr>
<td>conjunction</td>
<td>Coordinatees can't both be substitute and have supervener, &amp; {no/only the first} coordinatee has supervener; every supervenerless coordinatee has a conjunct-mate</td>
</tr>
<tr>
<td>gapping</td>
<td>Each coordinatee is conjunct-mate of every other coordinatee.</td>
</tr>
<tr>
<td>unary</td>
<td>Coordinatee can't be conjunct-mate.</td>
</tr>
<tr>
<td>supplementary</td>
<td>Coordinatee has supervener.</td>
</tr>
</tbody>
</table>

I have not formalized the rules or devices necessary for accounting for which of these conjunction classes can intersect with which lexemes, but the basic facts are as follows. Only **And**, **Or** and **But** have instances that are gappings but not unaries. And only **And**, **Or** and **But** have instances that are gappings and instances that are conjunctions but not gappings. It therefore looks as though if a lexeme can have instances that are gappings but not unaries then it can also have instances that are conjunctions but not gappings. It also appears that if a lexeme can have instances that are gappings but not unaries then it can also have instances that are unaries, and if a lexeme can have instances that are unaries but not supplementarys then it can also have instances that are supplementarys. That means that **And**, **Or** and **But** can have instances that are unaries or supplementarys.  But they don’t have instances that are subjunctions: for a lexeme to have instances that are subjunctions there must be an explicit rule permitting this.

§4.4 presents further symmetric conjunctions. §4.5 presents some unary conjunctions. §4.6—8 introduce subjunctions.

---

109 Supplementary **but**s generally require a supplementary **not** as their coordinatee: *She arrived late but *(not) him.* (See §4.6 on **not**.) If supplementary **but**s are ever exempt from this requirement it is when they're in negated *-clefts*, as in *It was*(n’t) him that arrived late, but her (cf. *She didn’t tell Sophy they’d be late but Edgar*).
4.4 More symmetric conjunctions

In Chapter 3 I wrote as if the only symmetric conjunctions there are are *and*, *or* and *but* (perhaps along with *neither... nor* as some negative version of *either... or*). To this list can unproblematically be added *for*, which is unusual only in that its coordinatees must be finite, in that it can't have a ward (*She worried *for* was late), and in that, for some speakers, it cannot be an s-dependent (i.e. it cannot conjoin subordinate clauses — cf. Quirk et al. (1985: §13.18)). Its failure to have a ward or to be an advener would simultaneously be explained if *for* is not supervener of its coordinatees. *But* too has restrictions on the syntactic function of its coordinatees; e.g. *but* can't coordinate subjects or objects, except in the *not... but* construction. The limitation to two coordinatees ought hopefully to be explicable on the semantic grounds that semantically it has two arguments (and, perhaps, syntactically two valents, e.g. a subject and a clausal), just as with *but*, and, for slightly different reasons, *both* (see §3.6.1).

Other symmetric conjunctions are *so*, *yet* and *then*. Like *but* and *for* they're limited to two coordinatees, presumably for the same reasons, and are restricted in what the function of their coordinatees can be. But they differ from other conjunctions in that they can occur following another conjunction, as in (278a—b). They can also be sentence-initial, as in (279a—b).

(278)  a. I'm not too keen on binding, *(and)* so am unfamiliar with the data.
       b. I'm keen on binding, *(and/but)* yet am unfamiliar with the data.
       c. She rose to her feet, *(and/but)* then couldn't think of anything to say.

(279)  a. I'm not too keen on binding. *So* I'm unfamiliar with the data.
       b. I'm keen on binding. *Yet* I'm unfamiliar with the data.
       c. She rose to her feet. *Then* she couldn't think of anything to say.

In my idiolect *nor* is restricted to these additional contexts, (280a—b), except where it's a supplementary or a complement of *neither*, and except, somewhat marginally, in some cases where the ward of *nor* is subject of *nor*'s coordinatees, e.g. (281a—c).

(280)  a. I realize that she didn't come, *(but/and)* nor did he.
       b. I realize that she didn't come. Nor did he.
(281) a. I am not Prince Hamlet, nor was meant to be.
b. She can't come, nor would wish to.
c. She tried not to speak nor to betray any sign of sensibility.

At present I have no account to offer of the syntax of nor and of yet and so when they're not acting unambiguously as conjunctions.

A fair number of other words appear to function on occasion as conjunctions (cf. Quirk et al. (1985: §13.20, §13.103)). I shall here discuss only though, and address only the question of what its dependency structure is when it is a conjunction. Thoughs that aren't conjunctions are exemplified in (282–284), and thoughs that are conjunctions are exemplified in (285).

(282) a. She can still work though tired.
b. Though tired she can still work.

(283) a. She can still work though she is tired.
b. Though she is tired she can still work.

(284) a. Tired though she is, she can still work.
b. She can still work, tired though she is.

(285) a. She is tired though can still work.
b. * Though can still work, she is tired.

The dependency structure of the thoughs in (282a–b) is shown in (286a). Though is adjunct of can and in general the subject of an adjunct of W is W, so we take can to be subject of though. Tired is xcomp of though because it gets its subject via though. The subject of tired is she, and the subject of the xcomp of W is the object of W, so therefore she is object of though. The dependency structure of the thoughs in (283a–b), shown in (286b) is essentially the same, except that the though's object is also its clausal (which means the clausal can't be xcomp of the though, because the subject of the xcomp of W is the object of W, and no word is its own subject). The dependency structure of the thoughs in (284a–b) is shown in (286c).
Though tired she can still work.

b. Though she is tired she can still work.

c. Tired though she is, she can still work.

WHILES are like THOUGHTS in having the diathetic patterns of both (286a–b), as are AFTERS and BEFOREs, except that the xcomp of these must be an ing-participle. WITHOUTS have just the pattern in (286a), again with an ing-participle xcomp. The pattern of (286b) is shared by numerous ‘subordinating conjunctions’ such as BECAUSEs and WHEREASes. The pattern of (286c) is shared, I think, only by ‘complementizer’ THATs and ASes, as in fool that I am, tired as I am, which has essentially the same distribution as tired though she is (which is why I conclude (on the basis of insufficient deliberation) that fool that I am and tired as I am are not relative clauses but instead involve though-movement).

When we come to the dependency structure of the conjunction THOUGH in (282g) the simplest analysis is simply to add the Supervener and Coordinatee dependencies onto the dependency structure of (286b).
4.5 Unary conjunctions

Supervenience, or even, as can be seen in §§4.6—8, conjunctionhood, is often hard to detect. For example, pied piping is a case of a construction that superficially doesn’t look in the least like coordination, but turns out to work only if we assume it shares certain properties — the presence of Supervener — with coordination. This claimed similarity of pied piping is contingent not only on assumptions internal to WG, but also on assumptions internal to TDG. This cloud of dubiety is rather unwelcome, since I wish to establish not only, at maximum, the validity of the totality of my proposals (e.g. that there are manifold varieties of conjunction) but also, at minimum, that coordination uses the same kind of mechanism — namely, dependency — as the rest of syntax. But thankfully there are constructions that much more clearly and less controversially appear to be quasi-coordinational. A particularly clear illustration of this is unary conjunctions, exemplified in (288a–f) ((288e–f adapted from Hudson (1989)).

(288) a. [A]pparent counterexamples require that the feature inventory be adjusted rather than [the principle] [abandoned]. (Ladefoged & Maddieson 1996: 332)

b. He’d rather give Thalia roses than [Euphrosyne] [tulips].

c. He’d as lief/soon give Thalia roses as [Euphrosyne] [tulips].

d. He longs to have eaten as many apples as [she did] [pears].[^10]

e. The plaster cast got dirtier and dirtier and at last he shed it like [a snake] [its skin].

f. The villain ended up with the woman instead of [the hero] [with her].

(288a–f) look very much as though the italicized words are mutual conjunct-mates. If they’re conjunct-mates, we’d expect that they’re coordinatees respectively of than, of the AS that is complement of the other AS, of like, and of the OF that is complement of the INSTEAD. In this case we’d therefore also expect that than, as, like and of are, at least in these constructions, conjunctions. In the case of than, this is in fact what non-theoretical grammar has traditionally held, albeit not for the same reasons. If we do take them to be conjunctions, and it seems entirely desirable and feasible to do so, then this way the similarity between comparatives and coordination — e.g. the way they both

[^10] It’s not clear what did is substitute of here, but my account certainly claims that there is something that it is substitute of.
involve substitutes (see §3.6.1) — is explained: both comparatives and coordinations involve conjunctions.

These unary conjunctions in (288a—f) are most similar to exclusive gapping, the difference being that no preceding (adjunct) coordinatees are permitted. A plausible dependency structure is shown in (289).

(289)

```
of of c
of co
of of c
of co
of cm of
of cm
of ss
of ss
[than [Euphrosyne] [tulips]]
[as [Euphrosyne] [tulips]]
[as [she] did] [pears]]
[like [a snake] [its [skin]]]
[of [the hero] [with [her]]]
```

In some ways this dependency structure is the opposite of that for pied piping. For pied piping we require Supervener, Ward and Advener dependencies, but not Conjunct-mate, Substitute, and Coordinatee. For unaries, we need Conjunct-mate, Substitute and Coordinatee, but not Supervener, Ward or Advener. The reason for thinking that the unaries are not superveners is that their coordinatees do not behave like s-dependents of anything other than the conjunction. Consider (288e): a (snake) is subject of shed, but if it were also s-dependent of shed then we would expect it to precede shed. The conjunctions are also not adveners.\(^1\) In (288c—d) the second _as_ is complement of the first _as_, and presumably extraposee of give in (288c) and of many in (288d). _Like_ in (288e) is adjunct of _shed_. And _of_ in (288f) is complement of _instead_, which is adjunct of _ended_.

As for _than_ in (288a—b), it can be argued that almost every _than_ is complement of a _more_. In §6.2.2 it is proposed that comparative adjectives are morphological— phonological indicators of a sequence of a _more_ plus an adjective. In §6.1.1 it is suggested that a sequence of a _more_ plus an _else_ is pronounced _other_. Similarly, it could be that _rather_ has changed diachronically less than we might have thought, and it remains a way of pronouncing a _more_ plus a _rathe_ (the actual change being that _rathes_ never occur independently from _mores_. That leaves a few not-entirely-standard cases where a _than_ is complement of something other than a _more_, e.g. the italicized words in (290a—b).

\(^1\) Or at least, even if they are advener of W they are also complement of W, and therefore unlike common-or-garden adveners.
(290) a. She is more different \textit{than} \{he is/him\} than \{I am/me\}.

b. Scarcely/hardly had he entered the room \textit{than} the music started up.\textsuperscript{112}

Whereas the symmetric conjunction \textit{THOUGH} in (287) differs from non-conjunction \textit{THOUGH}s not at all semantically, and, syntactically, only in that Supervener and Coordinatee dependencies are added to its usual valency, some of these unaries, such as the instances of \textit{LIKE} and \textit{OF} in (288d–e) have properties that are not merely the combination of the properties of \textit{LIKES} and \textit{OFFS} in general and the properties of unaries in general. The conjunction \textit{LIKES} are atypical as \textit{LIKES} because it is usually the meaning of a \textit{LIKE}'s subject and of its complement that are alike (i.e. that are arguments of the predicate “be like”); but in the meaning of (288d) it is two sheddings (i.e. expansions of the meaning of the word the \textit{LIKE} is adjunct of) that are said to be alike, not snakes or skins. Similar observations apply to \textit{instead of} in (288e).

4.6 Subjunctions and more supplementaries

So far, we've seen examples of words that are (i) symmetric conjunctions (and therefore also superveners), (ii) symmetric conjunctions, namely \textit{for}, that, exceptionally, are not superveners, (iii) gapping and unary conjunctions (that are therefore not superveners), and (iv) superveners but not conjunctions (as in pied piping). Examples of (i–iii) are fairly easy to spot, since the syntactic characteristics of symmetric, gapping and unary conjunctions, notably conjunct-mates and multiple coordinatees (which are essentially coordinacy properties), are fairly unmistakable, and once a word is identified as a conjunction it is easy to test whether the substitute relationships and additional looping dependencies obtain as they should. Supplementary conjunctions are a little harder to spot, since they allow only one coordinatee and therefore do not so clearly evidence their syntactic character. Hardest to spot are superveners without coordinacy: our earlier example of this was the proxy of a wh-pronoun in pied piping, and it took considerable effort in §4.2 to demonstrate the necessity of that analysis. In this section and the remaining sections of this chapter, I'll speculate very very tentatively about possible further examples of (non-subjunction) supplementary conjunctions and of subjunctions. These are mutually exclusive, for as we saw in §3.5.6 and §4.3, the coordinatees of supplementary conjunctions are substitutes but have no supervener, whereas the coord-

\textsuperscript{112} This locution looks like it originates by analogy with \textit{no sooner} \ldots \textit{than}, which does involve a \textit{MORE}. (See O.E.D. entry for \textit{than}.)
inatees of subjunctions have a supervener but aren't substitutes. But I deal with them both together in this section because we find a partial overlap between the respective memberships of the sets of candidates for the two constructions. §4.6.1 looks at candidates from negation, cursorily, and then in §4.6.2 at candidates from among subjunct adverbials, in a little more depth. §4.6.3—6 survey a diverse range of other candidate subjunctions.

4.6.1 Not and other than as supplementaries and subjunctions

Not. The strong syntactic parallel between (291—292) rather suggests that the Not is a supplementary conjunction, just like the And or Or in (292a—b).

(291) a. It was Sophy that emerged, not I/me.
    b. Sophy emerged flustered, not I/me.
(292) a. It could have been Sophy that emerged, and/or I/me.
    b. Sophy could have emerged flustered, and/or I/me.

The parallel between (291—292) is not only syntactic but also semantic: logicosemantically, negation and conjunctions are 'connectives' (or 'operators').

It is also principally for semantic reasons, at this stage of superficial inspection of the data, that we could consider taking some or all Not in other syntactic contexts to be superveners, e.g. in (293), where to is complement of not and not is supervener of to. The semantic reason is that the meaning of not is in no way a modifier (either restrictive or nonrestrictive) or argument of the meaning of to (laugh). The meaning of not is a connective whose argument is the meaning of to (laugh). Whatever rules build semantic structures off the other connectives, e.g. and and or, so that the meaning of the whole is not merely the sum of the meaning of the parts, could, one hopes, also operate on negation.

(293) She tried not to laugh.
**Other than.** The semantic reasons for giving *other than* in (294) the structure shown are the same as with *not* (though the negation is of the scalar rather than contrary variety), but the syntactic grounds are somewhat stronger. We’ve already decided in §4.4 that *than* is a conjunction, so we’d expect *with* to be coordinatee and complement of *than*. But *with* should also be xcomp of *found* (as it is in *these dependencies are found with conjunctions*).

![Diagram](image)

(294) These dependencies are found other than with conjunctions.

I have here taken *with* to be coordinatee of *than*, because this is what it should be if *than* is a conjunction, which it indeed in other contexts seems to be. But in this case *with* ought to be substitute of something, which it appears not to be. Put another way, this *than* lacks the coordinacy properties we’d expect a conjunction to have. One way out of this, which I reject, is to set up some inclusive category, ‘Conjunctor’, say, every instance of which is either a conjunction or a supervener. We could then state the generalization that every *than* is a conjunctor, which would allow some *thans* to be conjunctions and some to be superveners. Another way out of the problem, which I adopt, is to categorize the *than* in (294) as a subjunction. We distinguish between (non-subjunction) supplementaries that were in §4.3 claimed to be adjuncts, e.g. those in (291–292), and those that are subjunctions and not adjuncts but adveners, i.e. those like the *than* in (294).

### 4.6.2 Subjunct adverbials

A hefty cantle of my candidates for supplementaryhood in general and subjunctionhood in particular are all classified by Quirk et al. (1985) as ‘subjuncts’ — a kind of adverbial. In particular, Quirk et al.’s lists of ‘downtoners’ and ‘focusing subjuncts’, given in full in (295), include some promising items that might possibly be supplementaries as well as or instead of adverbs and adverbials.
Among these words we find, I venture, both subjunctions and (non-subjunction) supplementaries. The supplementaries, we see in the quasi-appositive construction in (296a—d).

(296) a. She eats but little meat, only/merely/simply/just **squid** garnished with lime juice.
   b. She studies several subjects, chiefly/especially/mainly/mostly/notably/particularly/primarily/principally/in particular **botany**.
   c. She studies botany, **specifically** botanical **osmology**.
   d. She read all his works, **even the** juvenilia.

Since the italicized subjuncts must be adjacent to the phrase they ‘focus’ (though as well as preceding the focused phrase, as in (296a—d), they can also follow it), the subjunct is either s-head or s-dependent of the root of the focused phrase’s root (‘FPR’), which is indicated in (296a—d) in boldface. If the subjunct were s-dependent, then presumably
the FPR would be appositive of the object of *eats/studies/read*, and subjuncts of this sort would function as adjuncts of words functioning as appositives. The problem with that analysis, though, is that the semantic relationship between the supposedly apposed words is not one of coreference, as in standard apposition, but rather is specified by the subjunct. Since the subjunct mediates semantically between the two quasi-apposed words, this can best be captured by supposing that it somehow mediates syntactically between them too. In this case, the FPR is s-dependent of the subjunct. In support of this conclusion, note that whereas appositives are normally nouns, the FPR needn’t be a noun; cf. (297–298).

(297) a. She was delighted, even jubilant.
   b. She was delighted, jubilant even.

(298) a. She writes with several implements, chiefly with crayons.
   b. She writes with several implements, with crayons, chiefly.

If the subjunct is s-head of the FPR, a possible structure is (299a–b), with the subjunct serving as a kind of quasi-appositive linker.

(299) a. She studies several subjects, chiefly botany.
   b. She was delighted, jubilant even.

This analysis is highly plausible for certain of Quirk et al.’s ‘appositive conjuncts’, such as *namely, for example, for instance* and *e.g.*, but for the subjuncts in question, however, it is suspect, because, as (297–298) show, the word class of the supposed complement of the subjunct is limited not by the requirements of the subjunct but by the function of the word that the subjunct is supposedly adjunct of. This can instead best be captured by taking the subjuncts to be ordinary non-subjunction supplementaries, with the dependency structure of (300a–b).

(300) a. She studies several subjects, chiefly botany.
We turn now to the possible subjunctions among the subjuncts of (295). If a subjunct is a subjunction, then its dependency structure is going to look like (301a), like that in (293-294). The alternative analyses, wherein the subjunct is not a subjunction, are, as far as I can see, (301b—c).

I'll present a range of arguments in favour of the (301a) analysis and against the (301b—c). I'll begin with a couple of arguments specifically against (301c), which is the analysis implied by calling these subjuncts 'adverbials'. First, if the subjunct must be s-dependent of the verb, or root of the clause, then discontinuity will result in such examples as (302a—b).

Second, the (301c) analysis would require exceptions to such almost exceptionless rules as those requiring unpromoted injects of ditransitives and unraised objects of finite aux-
iliaries to be tail of their s-heads (no s-dependent of W can intervene between W and its tail); cf. (303–304).

(303) a. She gave even Edgar flowers.
    b. * She gave today Edgar flowers.

(304) a. Will even Edgar relent?
    b. * Will today Edgar relent?

If tail is indeed defined as a prohibition against any s-dependent of W intervening between W and its tail, then (303–304) is evidence not only against (301c) but also against (301a), in which case (301b) would of course emerge as the best analysis. However, if tail is instead defined as a requirement that every word intervening between W and its tail be superordinate or subordinate of the tail then (303–303) do not constitute evidence against (301a), since even, which intervenes between give/will and Edgar, the tail of give/will, is superordinate of Edgar.

An initial though not very compelling reason for preferring (301a) to (301b–c) is that, like other complements, the complement coordinatee of a conjunction follows its s-head, which gives us the order in (301a), while if the subjuncts are adjuncts, as in (301b–c), we would expect them to follow their s-head, contrary to the actual word order. These default tendencies are at least suggestive, even though they must have exceptions, for I would analyse too in the same way as even, even though the word order is reversed: that would mean that the complement of certain conjunctions, such as too, precedes its conjunction s-head. As for how strong the tendency for adjuncts to follow their s-heads is, this depends what we end up calling subjunctions and what we end up calling adjuncts; it is possible that in time it may prove that apparent adjuncts that precede their heads but are not extracted or otherwise preposed are in fact subjunctions, e.g. degree modifiers and attributive adjectives (see §4.6.5).

A second reason for favouring (301a) is that adjuncts are on the fussy side with respect to what they're happily adjuncts of. Adjectives, for example, are generally adjuncts only of nouns, though when functioning predicatively they can also be absolute adjuncts of verbs or other predicative words (as in Shy, she hesitated or [Sophy hesitat-

---

113 The only exception I'm aware of is that parentheticals can irrupt between the aux and its unraised object, as in (i).

(i) Will, I wonder, all the shops still be open?
Adverbs are, according to traditional grammar at least, modifiers only of adjectives, verbs or adverbs; at any rate, they're not adjuncts of nouns functioning as subjects or objects. Subjuncts, however, are far less fussy about the class of the words they enter into dependency relationships with, which suggests they're not adjuncts of the words they focus or downtone. In this respect they're very much like conjunctions, for words of pretty much any class can be coordinatees. Thus, by distributional criteria of this sort, subjuncts come out looking more like conjunctions than adjuncts. It is noted in Hudson, Rosta & Gisborne (1994), reporting an observation by Larry Trask (p.c.) that some of these subjuncts, namely just, c.ily, even, almost, have an additional special property that differentiates them from common-or-garden adverbs: they can't be modified by degree modifiers — *very just/only/even 'almost. This fact is compatible with their being conjunctions: cf. Sophy (*very) and/or Edgar.

A third reason is semantic. If X is adjunct of Y, then typically, X and Y are in a conjunctive semantic relationship. Both virgins in boy virgins and boys in virgin boys refer to entities that are both virgins and boys; paper in blue paper refers to something that is paper and that is blue. Less often, when the adjunct is appositive, X and Y are in an ‘equative’ semantic relationship; they are coreferential or ‘cosensual’. In contrast, if X is a conjunction, and Y a coordinatee, then by analogy with those words about whose status as conjunctions we can be fairly confident, we would expect other conjunctions to have meanings that involve comparison or some element of meaning that is relevant to the logical portion of semantic structure; they might well be sensitive to scope, for example, just like and and or are. If we cursorily examine the meaning of subjuncts, we find that they have rather more in common with conjunctions than with adjuncts.

To take a few examples, only functions semantically as a quantifier. Only mammals lactate can be paraphrased as “every lactator is a mammal”, though since only triggers negative polarity items (Only she had seen anyone), the grammar of only ought to incorporate negation somewhere — perhaps “Nothing except” is an appropriate gloss for only (see McCawley (1993: 83)). Quantifiers and connectives are both scope-sensitive, and each reduces to the other (McCawley 1993: §8.5). Approximators such as almost are truth-conditionally negators; if something is almost ruined, it is not ruined but in

---

114 They don’t behave like negators, in that they don’t trigger negative polarity items: *She’s almost ever seen anyone at all. This is because the meaning of approximators involves double negation: they express a small amount of negativity. There are a fair number of analogous examples of lexicalized double negation, e.g. allow (“not prevent; not cause X to not do Y”), near (a small amount of distance, of “not-at-ness” — cf. *near to ever having won anything vs far from ever having won anything). (Cf. Cruse 1986.)
a state that is close to ruination. *Sort of* functions sort of as a sort of fuzzy negator — a ‘subnegator’. *Even* means, very roughly, “and moreover” (see Kay (1990), McCawley (1993: 320) for a proper analysis). *Also, too* and *as well* mean, again very roughly, “and, additionally”. To summarize, whereas the semantic relationship between X and Y, when X is adjunct of Y, is typically conjunctive or equative, the relationship between a conjunction and its coordinatee we might call ‘supplantive’, because the coordinatee’s extrasentential meaning is partly or wholly supplanted by the conjunction.

Fourth, the semantic relationships subjuncts are involved in support a dependency structure like (301a), for only in (301a) is the subjunct linked by dependencies to both *she* and *would*. Consider *Even someone she likes won’t suffice*. *Even* doesn’t just focus *someone*; it focuses *someone* with respect to *won’t suffice*, and not with respect to, say, *likes*. The meaning is something like “the set of those that won’t suffice also, surprisingly, contains X”, where X = someone she likes. A semantic structure that relates the meanings of words V and W is that much the easier to build if there is a dependency between V and W that the semantic relationship can be built off.

Fifth, the primary symptom of superveners is that, when C is candidate for being supervener of W, the internal structure of the C+W phrase suggests C is s-head of W, but on distributional grounds W appears to be s-dependent of a word outside C+W. This is how things look in pied piping. Now, some of the subjuncts that appear to involve more than one word, such as *sort of* and *all but*, will seem much less eccentric if we assume that the *of* and *but* are normally functioning prepositions. The sort of structure I’m envisaging is shown in (305). The dependency structure is left somewhat sketchy, but the important relationships are that *sort* is advener of *has* and supervener of *agreed*, which will make *agreed* an s-dependent but not niece of *has*.

(305) \[
\text{[(She) has [sort [of [agreed]]]]}.
\]

Some of these remarks about *all but* are recapitulated in the following section.
4.6.3 All but and more/less than

What is the structure of the italicized phrase in (306a—b)?

(306) a. The fee more than doubles.
   b. The ring all but betrays him in the end. [T.A. Shippey, The Road to Middle Earth]

There are both syntactic and semantic reasons for taking doubles to be complement of than, and than to be complement of more in (306a), and betrays to be complement of but and but to be complement (or at least modifier) of all in (306b).

The syntactic reason is that MORES subcategorize for an optional THAN complement, and THANS subcategorize for an obligatory complement. The semantic reason is that (according to WG) the ‘comparand’ argument of the meaning of a MORE is equated with the meaning of its THAN complement, while the meaning of a THAN is equated with the meaning of its complement. As for all but betrays, we can simply note the similarity to everyone except/but Sophy.

It might be claimed that than and but are prepositions, and that since the complement of a preposition is usually a noun, not a finite verb, doubles/betrays can’t be complement of than/but. My response to this objection in Rosta (1994) was to point out that not every complement of a preposition is a noun: the complement of a preposition is a preposition in (307a) and (307b) (from Defoe, quoted in Jespersen 1909–49 cited by Aarts 1992), and it’s an adjective in (307c) (from Kayne 1984, also via Aarts 1992).

(307) a. She came out from under the table.
   b. I would have given half I had in the world for him back again.
   c. She was distressed at the thought of him alone.

However, all that these data show is either that the complement of a preposition can be a preposition, as in (307a), or a ‘predicator’ (i.e. a preposition or adjective that has a subject), as in (307b—c), or that some prepositions and some predicators can count as
nouns. At any rate, they don’t demonstrate that prepositions can have finite complements. Nonetheless, (308–309) show rather clearly that a than can have a finite dependent.

(308) It pleases rather than upsets her.116
(309) a. She walks more than she talks.
   b. She saw more than he saw.
   c. It’s more likely she will be here than she won’t be here.

So, all in all, the most explanatory conclusion is that than and but are not prepositions, and that they can have a finite complement.

However, this minimal dependency structure for more than doubles/all but betray is not adequate, for the distribution of the phrase is determined by the complement of than/but. For example, the complement of than/but can be a finite, as in (306a), but in (310) the complement cannot be a finite; rather, it must be a perfect.

(310) She has more than taken/*took it to heart.

How can we explain this? I’ll start by rejecting the possibility that more is a proxy of taken/took, and that auxiliary has/es subcategorize not for a perfect but for a proxy of a perfect instead (with perfects being able to be their own proxies, like finites, as we saw in §2.6.2). The structure would be as in (311).

![Diagram](181)

115 See Jaworska (1986) on prepositions behaving nominally.
116 I in fact suspect that this does not involve than having a finite complement. There are two reasons for this. First, if than did have a finite complement then an impossible loop of s-dependencies would result: it (like all unextracted subjects of finites) should be s-dependent of upsets, upsets should be s-dependent of than, and than should be subordinate of pleases. There’s no way to get a projective structure out of that. Second, (i) seems rather poor, so constitutes a clear example of than not allowing a finite complement.
(i) *? She should rather than will.

In reality, the syntactic structure is probably like that in (ii), in which than has a nonfinite complement. See Chapter 5 for explanation.

(ii) IT DO PLEASE RATHER THAN UPSET HER
But this possibility cannot be correct. Suppose we analysed (312) in the same way: *more* would be the complement of *that* and the proxy of double(s); and the badness of *double*, which is nonfinite, would be because we’d be saying that a THAT subcategorizes not for a finite but for a proxy of a finite.

(312) She knew that the price more than doubles/*double.

But as we have seen earlier, a THAT itself can be a proxy of a finite, so were we to hold that a THAT subcategorizes for a proxy of a finite, rather than for a finite, we should fail to predict the ungrammaticality of *That that that that she went surprises him*.

We must conclude that in (312) it is *doubles* that is the complement of *that*, and, generalizing from this, that in these *more than X* constructions it is *X* that is the dependent of a word superordinate to the words in the phrase; (311), for instance, has the partial dependency structure shown in (313).

\[
\begin{array}{c}
\text{of}
\end{array}
\begin{array}{c}
of \\
of
\end{array}
\begin{array}{c}x \\
\text{has more than halved}
\end{array}
\]

(313)

In support of this conclusion, consider (314a) (pointed out to me by Dick Hudson) and (314b).

(314) a. More than one person has/*have complained.
   b. Less than two people have/*has complained.

In (314a) *more than one person* must refer to a plurality of people, yet the verb shows singular agreement. Similarly, *less than two people* may refer to a single person, yet the verb shows plural agreement. If it is *one/two* that is subject of *has/have* then the number agreement is explicable (i.e. *one* is syntactically singular, and *two* is syntactically plural).\(^{117}\)

\(^{117}\) The same test shows that in *All but Sophy are arriving* it is *all* rather than *Sophy* that is subject of *are*. I conclude that in *all but X*, sometimes *X* is dependent only of *but*, and sometimes it is dependent of *but* and of some word outside the phrase. If the *but* in *All but Sophy were* is a conjunction rather than a preposition, then it must be a variety of (non-subjunction) supplementary, such that *Sophy* is coordinatee of *but* and substitute of *all*. 

[182]
Note that the dependency structure in (313) is insufficient to provide continuity: there is no way for *more* to get an s-head. There are two ways of augmenting the dependency structure so that it provides continuity. On the one hand, there is (315a—b), where *more* and *all* are adjuncts of the finite, and on the other hand there is (316a—b), where *more* and *all* are superveners of the finite and adveners of *that*, and *than* and *but* are subjunctions. (In the diagrams constituent structure is shown by bracketing and by stemma, and where a dependency is paralleled by an s-dependency the letter indicating the dependency type is capitalized.)

![Diagram](image)

(315) a. [that [[it] [more [than]] halved]]

(316) a. [that [[it] more [than [halved]]]]

In both analyses, we can presume that the *than* and the *but* are subjunctions. This accounts for the relative lack of restrictions on what the word-class of the complement of the *than* and *but* can be. It also preserves §4.5’s claim that every *than* is a conjunction, and also the generalization that most, if not all, *buts* are conjunctions. And it also reflects the observation in §4.6.2 that words functioning supplantively rather than conjunctively or equatively are conjunctions; *more than happy* means not a state of happiness but a state exceeding happiness, and *all but betrayed* entails “not betrayed”.

[183]
Overall, I would argue that the subjunction analysis in (316a–b) is superior to the adjunct analysis in (315a–b), though the evidence is not compelling. First, we should prefer the subjunction analysis on the grounds of simplicity, since, unlike the adjunct analysis, it preserves the generalization that a complement of W is either (i) promoted to subject, (ii) moved (extracted or extraposed), or (iii) s-dependent of W. (This may not technically exhaust the possibilities, but even so it remains the case that a complement of W is either s-dependent of W or raised to become s-dependent of a superordinate of W.) In (315a–b), *halved/betrays* is complement of *than/but*, but none of (i–iii) apply. But in (316a–b), *halved/betrays* is s-dependent of *than/but*.

Second, evidence from echo-questions appears to favour the subjunction analysis. Consider (317a–b).

(317)  

<table>
<thead>
<tr>
<th>ECHO</th>
<th>NON-ECHO</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>% *</td>
</tr>
<tr>
<td>b.</td>
<td>% *</td>
</tr>
</tbody>
</table>

On a non-echo reading, (317a–b) are ungrammatical for all speakers. Both the adjunct and the subjunction analyses can explain this. According to the adjunct analysis, (317a–b) would give rise to the tangling shown in (318a–b). This is because *more/all* must be s-dependent of *how/what* because it has no other head. According to the subjunction analysis, (317a–b) would give rise to the tangling shown in (319a–b). This is because *how/what* is coordinatee of *than/but* and is therefore required to be s-dependent of *than/but*.

(318)  

| a. How was she more than? |
| b. What was she all but? |

(319)  

| a. How was she more than? |
| b. What was she all but? |

As echo-questions, (317a–b) are rejected by some speakers, who don’t like extracted echo-*wh* at all, but are accepted by other speakers (e.g. me). Given that all sorts of
transgressive things are permissible on echo readings, e.g. (320a—c), it is sorely tempting to conclude that these are not generated by the grammar at all. However, in this case we would have to account for the inversion in these examples of extracted echo-

wh.

(320)  a. She WHATTED him?
   b. WHAT did you say you love peanut butter and?
   c. WHO has she discovered a man who was in love with?

From data like (320b—c) I conclude that echo-wh can be extracted, triggering inversion in the normal way,\textsuperscript{118} but that the extraction of echo-wh is not subject to normal constraints on extraction. It may, for example, be extracted in a single swoop rather than in the normal stepping stone way.

Now, if (317a—b) are grammatical if the wh-word is of an echo variety then how can this be reconciled with the adjunct or the subjunction analysis, both of which appear to disallow (317a—b), as (317—319) showed? Clearly, some constraint must be relaxed. The only way the adjunct analysis, (315a—b), can tolerate (317a—b) would be if the global requirement of continuity were relaxed for echo questions. That seems excessive, even if it were possible. In contrast, to get the subjunction analysis to allow (317a—b), we need only to allow that a coordinatee of W needn’t be an s-dependent of W so long as it’s an echo-wh. Alternatively, we could leave unrelaxed the requirement that coordinatees be s-dependents, and instead allow extracted echo-wh to have two s-heads, so long its aunt is superordinate of its other s-head. That would predict that any dependent of W that is normally unextractable because it is required to be s-dependent of W would be able to extract if it is echo-wh; examples are \textit{WHO did she give flowers?} and \textit{WHAT did you break the handle of their?}, which strike me as possible, albeit not necessarily probable. Thus, the s-dependency structure I propose is that shown in (321), with the s-dependency from \textit{but} to \textit{what} a bit more tentative than the others.

(321)  WHAT had she been all but?

\textsuperscript{118} I assume that inversion results from extraction to subject, which prevents the object of the auxiliary from raising to subject; e.g. in \textit{She will put it there}, she is object and subject of \textit{will}, while in \textit{Where will she put it}, she is object of \textit{will} and \textit{where} is subject and extractee of \textit{will}.
4.6.4 Preposition subjunctions

PPish DPs: prepositional ‘limiters’ of cardinals. The italicized prepositions (‘limiters’) in (322a–b), I take to be subjunctions, with the associacy/dependency/s-dependency/constituency structure shown.

The arguments for treating these prepositions as subjunctions are essentially the same as those mustered in §§4.6.2–3. The meaning is supplantive: e.g. if under thirty went then it is not the case that thirty went. Taking thirty to be s-dependent and complement of the preposition preserves the generalization that all complements of W are either s-dependents of W or extractees or ‘promotees’ of W. In these examples thirty must be s-dependent of the preposition-subjunction: this is explained if thirty is its coordinatee.

Alternative analyses are possible. For example, prepositions could be allowed to also be nouns, so long as their complement be a cardinal; being a noun, over would be allowed to be a subject. But this would give the wrong semantics — “over [thirty people]” instead of “[over thirty] people”. And the analysis would not easily apply to up to thirty, unless up to be treated as a single word.

More than thirty people went on the reading “[more than thirty] people”, “people numbering more than thirty”, gets the structure in (323) — the same as (322b).
But what is the structure on the reading “more [than thirty people]”, “something that is more than thirty people”? The basic dependency and s-dependency structure is as in (324a), judging by the agreement in (324b).

\[
\begin{array}{cccc}
S & OF & C & OF \\
& & & \\
(324) a. & [(More [than [thirty [people]]]) went]. & b. & More than thirty people has been troubling her.
\end{array}
\]

Since thirty is here not subject of went/has, more can't be supervener of thirty. If thirty is coordinatee of than, and has no supervener, then it must be a substitute. But at present it is not clear what it is substitute of. I leave the resolution of this matter for future work on the grammar of More and Than.

**French and Italian partitives.** Another kind of prepositional subjunction is perhaps found with partitives such as French de and Italian di.\(^{119}\) In English, partitive of is rather limited in distribution, and it has a kind of antipassivizing, demotional function: cf. (325a—b), which strongly suggests that the is s-dependent of of and of is s-dependent of ate.

\[
(325) a.\text{ She ate of the meat.} \\
\quad b. \ast \text{ Of the meat was eaten.}
\]

But in French and Italian, 'partitive NPs' can be subjects.

\[
(326) a. \text{ De la viande est ici. [lit. "Of the meat is here" = "Some meat is here"]} \\
\quad b. \text{ Vengono dei miei amici. [lit. "They-come of the my friends" = "Some friends of mine are coming." ]}
\]

Let us suppose, if only for the sake of consistency with the rest of the grammar, that (i) the DE/DI (\(dei = di + i - \) see §5.2.3) is a preposition, (ii) the article is its complement

\[^{119}\text{This proposal stems from a suggestion by Dick Hudson.}\]
and s-dependent, and (iii) subjects must be nouns. What, then, is the subject of est/vengono in (326a–b)? One possibility is that the de/di is a noun as well as a preposition, in which case it could be subject of the verbs, and the partitive construction ceases to be relevant to the present discussion. Another possibility is that the de/di is a preposition and a subjunction, in which case (326a–b) would have the structure shown in (327).

(327) a. De la viande est ici.
   b. Vengono di i miei amici.

These type of dogs. In these types of dog the number agreement shows that it is types that is the complement of these; of is a complement of types, and dog is a complement of of. This construction is unremarkable. But in These type of dogs, which, if non-standard, is a not unusual construction, the number agreement shows that it is dogs that is the complement of these. There are also good reasons for taking dogs as the complement of of: first, these type of dogs means “these types of dog”; second, of invariably has a complement; and third, the argument of the meaning of type is equated with the meaning of of, whose meaning is always equated with the meaning of its complement. The complement of a pronoun is always its s-dependent, so dogs is s-dependent of these. And if it is not extracted or otherwise raised, the complement of a preposition is always its s-dependent, so dogs is s-dependent of of. We therefore have an s-dependency loop, which can be effected by making of a subjunction as well as a preposition, and by making type supervener of dogs, the coordinatee of of. (328a) shows the more standard construction, where of is not a subjunction, and (328b) shows the structure when of is a subjunction.
'Binominal' NPs. Should we wish to avail ourselves of it, the analysis in (328b) is also available for (English) partitive NPs, as in (329a), and for what Aarts (1994) has called 'binominal NPs' (see also Austin (1980)), as in (329b).

(329)  

a. An ounce of cannabis with a street value of ten million pounds was seized today by customs officials.  
b. In Sgt Steiner (James Coburn), the almost mythically indestructible leader of a battle-weary Wild Bunch of a platoon, he has another projection of his own embattled position within the film industry [Kathleen Murphy, Blood of a poet: The cinema according to Sam Peckinpah]

Subjunction analyses of partitives and binominals are given in (330a—b). In both sentences it is the of that is the subjunction. (330c) is, roughly, a WG version of Aarts's (1994) Government-Binding analysis. The advantage of (330b) over (330c) is that it would allow us to state that only an an or a plural common noun can be coordinatee of this 'evaluative' subjunction of, thus ruling out *that ginormous gorgon of that counterexample and *those ginormous gorgons of those counterexamples.)
b. She came up with that ginormous gorgon of a counterexample.

c. She came up with that ginormous gorgon of a counterexample.

It has been observed that partitive and binominal NPs can’t be extraposed out of (Akmajian & Lehrer 1976, Selkirk 1977, Aarts 1994); cf. (331a-b).

(331) a. * Dozens were announced of conferences on extraposition.
    b. * That ginormous gorgon was raised of a counterexample (to the analysis).

The analyses shown in (330a—c) would be one way of accounting for (331a—b). (330c) would straightforwardly give rise to the nonprojective s-dependency structure in (332).

(332) That ginormous gorgon was raised of a counterexample.

As for the (330a—b) subjunction analysis, these does not make (331a—b) discontinuous because the s-dependency from of to dozens/gorgon is nonprojective.

4.6.5 Adjuncts?

In §§4.6.1—3 I’ve been arguing that a sizeable number of what many would reckon to be adjuncts are in fact subjunctions, a kind of conjunction. A more conservative, and less satisfactory, WG analysis would replace the supervener associacies posited in §§4.6.1—3 by adjunct s-dependencies (as can be verified from Rosta (1994), which offers just such an analysis of much the same data). The subjunction analyses I’ve been advocating are shown schematically in (333a—b), and their respective adjunct-based alternatives are represented in (334a—b). ‘W’ represents any word, and ‘S’ represents a subjunction. As usual, single lines in the stemma represent niece s-dependencies and double lines represent non-niece s-dependencies. (333a)/(334a) would correspond to, say, invite under thirty, and (333b)/(334b) to invite up to thirty.
Adjuncts that don’t follow the (333)/(334) schema appear to fit one of essentially only four other schemas, (335a–d). (335a) is the pattern for appositive adjuncts. (335b) is the pattern for bare relatives and parentheticals. (335c) is the most common pattern, which applies for the majority of adjuncts; generally, if X is adjunct of Y, then X is predicated of Y. (335d) is the pattern for absolutes and depictives, and perhaps other things such as adverbs: semantically, W1 appears to be an argument of W2, but W1 is not subject of W2.
Since some adjunct constructions have been reanalysed as involving subjunctions, could it be that some or all of the putative adjuncts that follow some or all of the patterns in (335a—d) are in fact subjunctions? Both the pros and the cons of such a step are too numerous and too ramified to do justice to them here, so I shall confine myself here to noting some especially salient and general pros and cons.

Possibly the defining feature of adjuncts as opposed at least to valents (i.e. 'arguments' — subjects, objects, etc.) is that a word can have many adjuncts of the same type without them being asyndetically or syndetically coordinated. For example, (336a) is possible, and is not necessarily synonymous with (336b).

(336) a. She smoked during meals during termtime during her thirties.
   b. She smoked during meals, during termtime and during her thirties.

Semantically, (336a) (when understood as not synonymous with (336b)) has an interpretation reflected in the bracketing in (337a) and the glosses in (337b—c).

(337) a. [[[She smoked] during meals] during termtime] during her thirties].
   b. It was during her thirties that she smoked during meals during termtime.
   c. During her thirties, it was during termtime that she smoked during meals.

Suppose that during were in fact a subjunction. In this case we could give (336a) the structure shown in (338).

(338)
This analysis has several merits. First, note how the constituency closely reflects the semantic structure. While I've given hardly a hint about how the semantic structure actually gets built off the syntax, it should at least be clear that the closer the correspondence between syntax and semantics is, the simpler the rules that effect it need to be. Very roughly speaking, if (i) X is valent or coordinatee of Y, (ii) Z is advener of Y and supervener of X, and (iii) A is an argument of the meaning of Y that is usually expressed by X, then A is a synthesis of the meaning of X and the meaning of Z.

Second, if the analysis in (338) can be extended to all adjuncts, then we can state the generalization that if W has more than one dependent of the same type, then one of the dependents must have each of the others as its substitute. In general, that would mean they must be coordinated. Adjuncts would be exempt from this requirement, because the adjunct of W wouldn't be a dependent of W; rather, it would be a supervener of W. It might even turn out that the Adjunct relationship can be dispensed with altogether.\(^{120}\)

Third, we also explain why in general adjuncts of W follow complements of W. If the bracketing is \([W [C] [A]]\), with A as adjunct and s-dependent of W, then the ordering of A with respect to complement C must be stipulated. But if A is supervener of W and W is s-dependent of A then the bracketing is \([W [C]] A\), which immediately explains why the order is \(W - C - A\) rather than \(W - A - C\). When we do find \(W - A - C\), this can be dealt with by permitting the subjunction A to have a ward, as in (339).

\[
(339) \quad [[[\text{She}] \text{ knew}] \text{ for sure} [[[\text{he}] \text{ was} \text{ [there]]}]].
\]

This suggestion that some or all adjuncts of patterns in (335) are subjunctions is not altogether easy to square with such adjuncts' extractability. The predicted structure of

\(^{120}\) If all adjuncts are in fact superveners, then what I have been analysing as exclusive gapping, unary and supplementary conjunctions, which are adjuncts and not superveners, are consequently impossible. Only the inclusive versions of these conjunctions would exist. The only significant consequence of this, so far as I can see, is that we lose the earlier explanation for why, in some lects, non-symmetric conjunctions can't be complement of a (correlative) pronoun. However, this is not too hard to rectify. For example, 'exclusive' but not 'inclusive' lects might have a rule requiring a pronoun to be supervener of each coordinatee of its conjunction complement. That would serve to restrict correlatives to symmetric coordination.
an extracted adjunct is shown in (340). (Note that the putative subjunctions in §§4.6.1—3 don’t extract, even when the subjunction is adjunct/supervener of a verb.)

\[
\begin{array}{c}
\text{AV} \\
\text{OF} \\
\text{C}
\end{array}
\]

(340) When do you think she will arrive?

The most glaring problem with these proposals is that they make completely wrong predictions about the word order of adjuncts of adjuncts. On the standard WG assumption that adjuncts are dependents, we correctly predict the grammaticality of (341a—b) and the ungrammaticality of (342a—b). In contrast, under my proposals, as (343—44) show, (341a—b) are incorrectly predicted to be ungrammatical, while (342a—b) are incorrectly predicted to be grammatical.

(341) a. [Run [[very] quickly]].
   b. [[Needing a cigarette [badly]], she reached for her bag].

(342) a. * Very run quickly.
   b. * Needing a cigarette, she reached for her bag badly.

(343) a. Run very quickly.
   b. Needing a cigarette badly, she reached for her bag.

(344) a. * [Very [[run] quickly]].
   b. * [[Needing a cigarette, [she reached for her bag]] badly].

It is possible to stipulate our way out of this problem. For example, a rule to the effect of (345a—b) would rule out (344a—b) and instead impose the structure in (346a—b). But clearly, if we need to stipulate (345), then we are paying for the removal of arbitrariness in one area with arbitrariness in a new area.
(345) a. If X is coordinatee of Y and X is precessor of Y and Y is coordinatee of Z then X is ward of Z and X is precessor of Z.

b. If X is coordinatee of Y and X is successor of Y and Y is coordinatee of Z then X is ward of Z and X is successor of Z.

\[
\begin{array}{c}
\text{W} \\
\text{OF} \\
\text{OF} \\
\text{OF} \\
\text{CO}
\end{array}
\]

(346) a. Run very quickly.

\[
\begin{array}{c}
\text{CO} \\
\text{OF} \\
\text{OF} \\
\text{OF} \\
\text{W}
\end{array}
\]

b. Needing a cigarette badly, she reached for her bag.

Even though the need for (345) rather takes the shine off the adjuncts-as-subjunctions proposal, we should arguably retain it until some way is found either to render (345) less arbitrary or to explain the ordering and recursion properties of adjuncts. Within WG, the only rival explanation for the ordering of adjuncts is found in Hudson (1984), which suggests that the more specific the level at which a dependent is selected, or licensed, the nearer its head it will be. For example, according to Hudson, the ability of read in read books now to have an object is granted by virtue of its being an instance of the lexeme READ, while its ability to have an adjunct is granted by virtue of its being an instance of the more general category Verb. As it stands, this suggestion (not repeated in subsequent works on WG) is inadequate on several counts. For example, it is hard to formalize, and were it formalized, new devices would be required. More importantly, the correlation of ordering with specificity of selection/licensing is only a tendency. Much complementation is not lexeme-specific (pace Hudson (1984, 1990)) and a few adjuncts are lexeme-specific (e.g. need badly, blind drunk).

However, in Hudson’s suggestion we can find the basis of a possible alternative explanatory account of adjuncts. In essence, Hudson’s suggestion can be understood as making terminal nodes project non-terminals, such that complements link to terminals, and adjuncts to non-terminals. The resulting two-dimensional structure is then constrained not to tangle, and in consequence the ordering of adjuncts is correctly predicted, as (347a−b) show. (Projected nodes are shown by a double line.)

\[
\begin{array}{c}
\text{read books now}
\end{array}
\]
If projection of non-terminal nodes is recursive, then we also have a way to capture the recursion of adjuncts. This is shown in (348).

These non-terminal nodes are highly reminiscent of the distinction in Dependency Phonology between adjunction of nodes and subjunction of nodes. In (347—348), the double lines correspond to subjunction and single lines to subjunction. The problem with this solution in the context of WG, though, is that there are no entities that the non-terminal nodes can be understood as, and there is no other motivation for thinking that non-terminal nodes exist.
This chapter argues that two or more 'coenunciate' words can be enounced simultaneously. The ramifications of this are considerable, affording some very illuminating and far reaching analyses of a wide range of areas of English grammar.

5.1 Planes, structurings and strata

It will be useful for the discussion to come in this chapter for us to begin with a consideration of the kind of thing that gets called a 'level', a 'structure' or a 'stratum' in grammatical theory. I shall present, largely by means of analogy, a not necessarily theory-neutral discussion and definition of three notions: 'plane', 'structuring' and 'stratum'. Briefly, by 'structuring' I mean some kind of overall organization of elements. For example, twenty bricks laid end to end in a row would be one possible structuring of the bricks, and if they were piled one on top of another into a column, or stacked into a pyramid, these would be other structurings.

Planes. A structuring of one plane differs from a structuring of another plane in that they are structurings of different kinds of thing. Consider a workplace such as a company headquarters. To 'analyse' this workplace we must consider both the organization of the building, and the organization of the personnel. Certain structurings, for example, relationships like 'next to' and 'opposite', apply to the rooms in the building. And certain structurings, such as 'boss of' and 'secretary of', apply to the personnel. This analysis of this workplace involves two 'planes', the building plane, and the personnel plane. The two planes are associated with each other: for example, rooms and personnel

---

121 I define these terms only for the purposes of the present discussion. They are relevant when discussing the overall architecture of a grammatical theory, and such architectural matters are, for the time being at least, mostly settled for WG, and so do not require frequent discussion.
are connected by relationships like ‘office of’. We wouldn’t consider them planes if they weren’t in some kind of correspondence with each other.

**Strata.** Suppose several different structurings apply to the same plane — that is, they are alternative structurings of the same thing. On the one hand, these structurings can be mutually compatible, as, say, if the company personnel were sorted by age or by rank or structured in terms of their official role in the company or in terms of their social role in the company (e.g. X is Y’s shoulder to cry on, X is the office lothario). On the other hand, these structurings can be incompatible, as with the various ways of arranging the twenty bricks. If two structurings apply to the same plane, and they are mutually compatible, then they apply at the same stratum — they’re ‘homostratal’. If two structurings apply to the same plane, and they are mutually incompatible, then they apply at different strata — the structurings are ‘heterostratal’. If the plane comprises only one stratum, it’s monostratal, while if it comprises more than one stratum, it’s polystratal. Another example of homostratal structurings is physical and political maps: both sorts of map apply to the same thing (the surface of the earth), so are on the same plane,¹²² and each chunks the surface of the earth up differently, so they’re different structurings. But both structurings can be true at the same time, so they’re homostratal. For some further examples of heterostratal structurings, suppose Sophy has a gross income of £4000 and owes £2000 in tax, while Edgar has a gross income of £3000 and owes no tax. At the stratum of gross income, Sophy has a higher income than Edgar, but at the stratum of net income, Sophy has a lower income (£2000) than Edgar. The statement “Sophy’s income is higher than Edgar’s” is true at the gross stratum but false at the net stratum. Or suppose Sophy and Edgar each work part-time in the same two companies. In one company, Sophy is Edgar’s boss, and in the other company Edgar is Sophy’s boss. Here we have one plane, since the boss-of structurings apply to the same people, but two strata, one for each company, since, boss-of not being a symmetric relationship, it cannot be that Sophy is Edgar’s boss and Edgar is Sophy’s; to answer the question “Is Sophy Edgar’s boss?” we first have to ascertain which stratum — which company — is being talked about.

¹²² If the British Isles disappeared from the physical map, the political map would necessarily be altered. That shows the structurings apply in the same plane. In contrast, in our company headquarters example, if the head of the payroll section left the company, the building’s room structure would be unaffected — the room formerly occupied by the departing employee would not disappear.
To summarize non-analogically, if two structurings in no way apply to the same kind of entity, then the structurings are on different planes. And if two structurings are contradictory, they are defined over different strata.

Grammar must involve at least two planes: the plane of sound, and the plane of meaning. I take the view that morphological structurings and phonological structurings are homostratal on the same plane. In the next few sections I shall be considering whether syntactic structurings operate on the same plane as phonology and morphology, and will end up concluding that syntax constitutes its own plane — in other words, that the ‘matter’ or ‘substance’ of words (qua syntactic units), the clay of which they are formed, is not phonological. I’ll also be asking whether s-dependency structure and dependency structure are homostratal structurings on the syntactic plane, or whether s-dependency structure operates on the phonological plane, and will conclude that despite the attractions of the latter, the former is ultimately more justified. At least on the semantic and syntactic planes, there is no polystratality (the phonological/morphological plane is pretty much unexplored in WG). WG, unlike Transformational Grammar or Relational Grammar, strives to be monostratal.

Parsimony dictates that, other things being equal, it is preferable that grammar involve the minimum of two planes and no more than that. So, well-motivated though it is, my postulation of a syntactic plane distinct from the phonological-morphological plane is going against the dictates of parsimony. In the discussion that is to come, my basic position is that desirable though a single plane for syntax/morphology/phonology would be, there are compelling data and alluring analyses that draw us to put syntax on a different plane. But, with my hands very much waving, I speculate that there is no specifically semantic plane distinct from the syntactic plane: specifically, although words would still have senses and referents, the grammar delivers nothing but a structure of speech sounds (or gestural signs, or whatever) at one end, and a structure of syntactic words at the other end. What would have been on a semantic plane will instead be either on the syntactic plane or outside grammar altogether. On this view, there are still just two planes — form and meaning — but syntax, the structure composed of grammatical relations, is located not on the form plane but on the meaning plane. There are definite empirical consequences of this position: for example, lexical decomposition is possible only if it is syntactic, and if logical scope is partly grammatically-determined, then it must be represented syntactically. In the discussion that follows I offer no defence of these proposals, but by the discussion’s end one should have an impression of how and why they might be credible.

Strictly speaking, it may be that TG is not polystratal. A decent definition of monostratality is that for any pair of (possibly arbitrarily defined) strata, every proposition that is true at/of one stratum is true at/of the other, so that there is no point in distinguishing strata. RG, then, is polystratal, because, say, an NP can be 1 (and not chômeur) at one stratum but chômeur (and not 1) at another stratum. In a derivationally formulated TG, an NP might be in comp of PP (and nowhere else) at d-structure and in spec of CP (and nowhere else) at s-structure. But one could equally well say that there is only one stratum, and at it there is a chain occupying comp of PP and spec of CP.

---

123 Parsimony dictates that, other things being equal, it is preferable that grammar involve the minimum of two planes and no more than that. So, well-motivated though it is, my postulation of a syntactic plane distinct from the phonological-morphological plane is going against the dictates of parsimony. In the discussion that is to come, my basic position is that desirable though a single plane for syntax/morphology/phonology would be, there are compelling data and alluring analyses that draw us to put syntax on a different plane. But, with my hands very much waving, I speculate that there is no specifically semantic plane distinct from the syntactic plane: specifically, although words would still have senses and referents, the grammar delivers nothing but a structure of speech sounds (or gestural signs, or whatever) at one end, and a structure of syntactic words at the other end. What would have been on a semantic plane will instead be either on the syntactic plane or outside grammar altogether. On this view, there are still just two planes — form and meaning — but syntax, the structure composed of grammatical relations, is located not on the form plane but on the meaning plane. There are definite empirical consequences of this position: for example, lexical decomposition is possible only if it is syntactic, and if logical scope is partly grammatically-determined, then it must be represented syntactically. In the discussion that follows I offer no defence of these proposals, but by the discussion’s end one should have an impression of how and why they might be credible.

124 Strictly speaking, it may be that TG is not polystratal. A decent definition of monostratality is that for any pair of (possibly arbitrarily defined) strata, every proposition that is true at/of one stratum is true at/of the other, so that there is no point in distinguishing strata. RG, then, is polystratal, because, say, an NP can be 1 (and not chômeur) at one stratum but chômeur (and not 1) at another stratum. In a derivationally formulated TG, an NP might be in comp of PP (and nowhere else) at d-structure and in spec of CP (and nowhere else) at s-structure. But one could equally well say that there is only one stratum, and at it there is a chain occupying comp of PP and spec of CP.
5.2 What are words?

The answer is going to be that (i) they’re conceptualized as actions, (ii) they constitute nodes in dependency structure, (iii) they constitute nodes in s-dependency structure, and (iv) they in no way consist of sounds or any other kind of phonological or morphological entity. §5.2.1 establishes the term ‘vocable’ to denote ‘phonological words’. §5.2.2 discusses alternative possible relationships between nodes in dependency structure and the sounds of words. §5.2.3 introduces evidence to show which of those alternative relationships is correct. §5.2.4–6 elaborate the analysis of the relationship between words and their sounds. §5.2.7–8 briefly apply the analysis to clitics and articles.

5.2.1 Vocables

It is standardly agreed that there is such a thing as a minimal morphological unit — i.e. the morpheme — and a maximal morphological unit, such that in *Forswear underwear* there are four morphemes and two maximal morphological units. Maximal morphological units are standardly called ‘words’, but I shall call them ‘vocables’, because I’ll be arguing that they’re not words, and because the term’s etymology serves as a reminder that the criteria for recognizing a vocable are phonological as well as morphological — a vocable is a ‘phonological word’. The exact nature of these criteria needn’t concern us here; a rule of thumb can be that the graphological indicator of a vocable boundary is the blank space between graphological words. So *Milk’s* is one vocable, even though it is by now fairly widely recognized that syntactically it seems to represent two words.

I shall be assuming that a vocable is a constituent structure whose terminal nodes are phonological structures. The substance of a vocable, then, is phonological. Though it is not strictly relevant to the points that will be under discussion, it follows from this view that there will be no phonologically zero morphemes and hence no phonologically zero vocables. Hence *(two) sheep (are)* is a vocable containing only one morpheme, *sheep*.

---

125 The — pretty orthodox — idea is that the terminal nodes are morphemes. Obviously this approach can’t cope with nonconcatenative morphology, whether it involves intercalation, infixation or suprasegmentals. But since the internal structure of vocables (i.e. morphology, in effect) is not really an issue in this discussion I shan’t worry about this.

[200]
5.2.2 The nature of terminal nodes in syntactic structure

It is also standardly agreed that there is such a thing as syntactic structure, wherein grammatical relations such as ‘subject’ and ‘object’ are represented (either directly or derivably). Unsurprisingly, I agree with this.

What is the nature of terminal nodes in syntactic structure? My impression is that it is rather taken for granted that terminal nodes are morphological units, that they are typically vocables, and that when they aren’t vocables, as in the case of clitics (e.g. in *Milk’s*), this is a challenging though tractable analytical problem. It is, I think, generally assumed that a grammatical structure is formed by inserting elements from the lexicon. In a lexicalist theory (such as WG), the structure is thereby complete. In a theory with a computational component, derivation follows lexical insertion. In the theories of Transformational Grammar, lexical items are inserted not only with syntactic features but also with semantic and phonological features. So one and the same structure is interpretable as a syntactic, a semantic or a phonological structure, simply by ignoring irrelevant features. These features (or feature values) are not to be understood quite as they would be in WG: in WG a word’s ‘features’ would be the categories it belongs to. Thus, if the value for the Sex feature of Chris is female, then Chris belongs to the class of females. Now, suppose cat had the syntactic feature [noun], the semantic feature [feline], and the phonological feature [KAT]: this word belongs to the class of nouns, but not to the class of felines, so the semantic [feline] feature needs to be understood not as a WG feature value but as, say, an instruction to activate the concept ‘Cat’. As for the phonological [KAT] feature, from the perspective of production, this is to be understood as an instruction to manifest phonetically the phonological features. The function of a word, then, is as a package delivering pairings of semantic and phonological instructions. I’ll label this the ‘pantechnicon’ approach: the idea is

---

126 From a reading of secondary sources, such as Matthews (1993), I gather (perhaps mistakenly) that it was a view of structuralist linguistics (e.g. Harris (1951), which Matthews discusses), continuing into mainstream transformationalist theoretical linguistics, that it is morphemes (in general, rather than stems and inflections) that are terminal nodes in syntactic structure. That would give the syntactic structure of *Forswear underwear* four terminal nodes. Such a view is surely not current, so I shall ignore it (except briefly in §5.2.5), and the question of whether this report is historically accurate.

127 This, at least, is the view of Chomsky (1993, 1995), though a contrary position is stated in Jackendoff (1993), who claims to work within the same theoretical confession. In terms of the pantechnicon/hand-in-hand distinction outlined below, Chomsky’s position is pantechnicon par excellence, and Jackendoff’s (if I read him correctly) is a particularly clear exposition of a hand-in-hand relationship between distinct syntactic, semantic and phonological planes (though on the evidence of Jackendoff 1996, he makes no distinction between structurings and planes).

128 ‘[KAT]’ is not here intended as some kind of phonetic representation of a sonic event.
that syntactic structure carries within it all features necessary for both semantic and phonological structure.

Semantic structure in WG is effected somewhat differently. In WG terms, it does not make sense to say a word has semantic features: *cat* cannot have the feature [feline], since, as I have said, a word is not a feline. Rather, a word is linked by conceptual relations to its meaning: we have two different entities, an instance of the word *cat* and a concept of a cat, linked by a relationship such as Referent. This means that a syntactic structure, consisting of words, is paralleled by a semantic structure whose components are linked to components of the syntactic structure by conceptual relations. I’ll call this the ‘hand-in-hand’ approach: whereas in the pantechnicon approach everything is contained in a single structure, in the hand-in-hand approach we have separate but connected structures. (349a—b) illustrate diagrammatically the notional difference between the pantechnicon and hand-in-hand approaches respectively, for the sequences of syntactic and semantic entities in *Sophy loves chocolate*.

(349) a.  

![Diagram of syntactic and semantic structure](image)

In the terms of the discussion of planes in §5.1, the pantechnicon approach conflates heterogeneous kinds of item (e.g. syntactic, semantic, phonological) onto one plane, while the hand-in-hand approach filters them onto different planes. The pantechnicon approach makes the prediction that structurings on the one plane apply equally to these different kinds of item. As the more restrictive of the two, the pantechnicon approach is therefore in principle to be preferred as far as possible. Only if structuring V applies to one kind of item, X, (e.g. syntactic) while structuring W applies to a different kind of item, Y, (e.g. semantic), and V does not apply to Y and W does not apply to X, should separate planes be recognized. WG takes the view that semantics and syntax are indeed separate planes: structurings of semantic units and structurings of syntactic units do not necessarily match. Put another way, a syntactic unit is not a semantic unit; words are distinct from the concepts they have as their meanings.

[202]
A question I shall be focusing on is whether terminal nodes in syntactic structure are, on the one hand, vocables (or sometimes concatenations of vocables) — that is, entities with both syntactic and morphological-phonological properties, packaged together pantechniconwise — or, on the other hand, the relationship between syntactic terminal nodes and morphological-phonological properties is hand-in-hand, analogous to their relationship with their meanings, so that just as the referent of an instance of the word *cat* is a cat, so the ‘enunciation’ of the word *cat* is an instance of the morphological-phonological structure */kæt/*. I’ll articulate the difference between these alternatives in more depth later on. At this stage it is sufficient to note that either position is compatible with the basic tenets of WG. On the pantechnicon view, it would be possible for instances of, say, the lexemes *Heir*, *Ere* and *Air* to also be instances of a category ‘/e:/’ (in a nonrhotic lect). The instances of *Heir* and *Air* would also be instances of Noun, and instances of *Ere* would also be instances of Preposition. For the utterance *I can* it would be the action of articulating *I* that would be a pronoun and the subject of the action of articulating *can*, and whose referent the utterer would be. Metaphorically, if words could be mashed up or dissected, the flesh and blood substance we’d be dealing with would be phonological (i.e. conceptualizations of phonetic actions). On the hand-in-hand view, instances of *Heir*, *Ere* and *Air* would not be instances of */e:/*, but the enunciation of the instances of *Heir*, *Ere* and *Air* would be an instance of */e:/*. In this case, a mental representation of the action of uttering [kæt] is represented as an instance of categories of phonological actions but not as an instance of *Cat*, a category of lexical actions. The structure of lexical actions is paralleled by a structure of phonological actions; the two structures are linked by conceptual relations of representation. Phonological actions function as symptoms from which the occurrence of lexical actions can be inferred.

Let me present the dichotomy in another, complementary way. We could think of a word as a pairing of a sound with a meaning. A word would be like a predicate with two arguments, one for the sound and one for the meaning. This would imply two planes, one of sound and one of meaning, linked, hand-in-hand, by relationships that words would be. Structurings could not apply to words as such, but only to sounds or

--

129 Dick Hudson has used the term ‘form’ to mean ‘enunciation’ (though not in published work). I have tried using ‘form’ rather than ‘enunciation’, but have found that it leads to an unacceptable amount of ambiguity, given that ‘form’ is a common word often used in a relatively non-technical sense. Also, the term ‘enunciation’ makes it clear that we are talking about conceptualizations of the phonetic manifestation of word tokens rather than of some abstracted archetypal phonological structure, which the term ‘form’ is more likely to suggest.
to meanings; we would have to see syntax as ultimately involving structurings of sounds. Syntax and phonology would be on the same plane pantechniconwise. I shall be presenting evidence against this intuitively obvious view, and shall argue that a separate, biplanar, hand-in-hand relationship holds between phonological-morphological structure on the one hand, and dependency structure on the other, but for s-dependency structure both pantechnicon and hand-in-hand positions have their attractions. We'll take dependency structure first, looking at evidence from 'articled prepositions', and then s-dependency structure, looking at various conflicting evidence.

5.2.3 'Articled prepositions' and contracted auxiliaries

A number of European languages contain 'fused' vocables that occur in contexts where we would expect a sequence of a preposition and a determiner.\(^\text{130}\) Examples from German are zum = zu + dem, vom = von + dem, zur = zu + der, ins = in + das and im = in + dem. Examples from French are du = de+le, des = de+les, au = à + le, and aux = à + les. Examples from Italian are shown in the following table (from Speight (1962), with items in boldface from Dardano & Trifone (1983));\(^\text{131}\) the prepositions in the leftmost column fuse with the articles in the topmost row to give the resulting forms in the table.

<table>
<thead>
<tr>
<th></th>
<th>il</th>
<th>i</th>
<th>lo</th>
<th>gli</th>
<th>la</th>
<th>le</th>
<th>l'</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>al</td>
<td>ai</td>
<td>allo</td>
<td>agli</td>
<td>alla</td>
<td>alle</td>
<td>all'</td>
</tr>
<tr>
<td>con</td>
<td>col</td>
<td>coi</td>
<td>collo</td>
<td>cogli</td>
<td>colla</td>
<td>colle</td>
<td>coll'</td>
</tr>
<tr>
<td>da</td>
<td>dal</td>
<td>dai</td>
<td>dallo</td>
<td>dagli</td>
<td>dalla</td>
<td>dalle</td>
<td>dall'</td>
</tr>
<tr>
<td>di</td>
<td>del</td>
<td>dei</td>
<td>dello</td>
<td>degli</td>
<td>della</td>
<td>delle</td>
<td>dell'</td>
</tr>
<tr>
<td>in</td>
<td>nel</td>
<td>nei</td>
<td>nello</td>
<td>negli</td>
<td>nella</td>
<td>nelle</td>
<td>nell'</td>
</tr>
<tr>
<td>per</td>
<td>pel</td>
<td>pei</td>
<td>pello</td>
<td>pegli</td>
<td>pella</td>
<td>pelle</td>
<td>pell'</td>
</tr>
<tr>
<td>su</td>
<td>sul</td>
<td>sui</td>
<td>sullo</td>
<td>sugli</td>
<td>sulla</td>
<td>sulle</td>
<td>sull'</td>
</tr>
</tbody>
</table>

\(^{130}\) The Italian term for these is preposizioni articolate — hence my use of 'articled prepositions'. I don't intend the term to imply that these vocables are necessarily prepositions, or even, for that matter, that they are words.

\(^{131}\) I have no authority for pell’, but the possibility of pella entails the possibility of pell’, since the alternation is regular.
Some of the Italian items alternate with an unfused sequence of two vocables, e.g. *per il* as well as *pel*. Some of the fused vocables of Italian, and all the French ones wholly preclude an unfused sequence of two vocables:

\[(350)\]

a. de la parisienne ["of the (female) Parisian"]

b. *de le parisien ["of the (male) Parisian"]

c. du parisien ["of the (male) Parisian"]

In English, auxiliary verbs have forms that fuse in this way. Obvious examples are *you’re*, which is homophonous with *yore* not *ever*, *we’re*, which is homophonous with *wear* not *weir*, and *they’re*, which rhymes with *lair* rather than *layer*.\(^{132}\) In addition, these contractions have weak forms \(/j@\)/, \(/w@\)/ and \(/b@\)/. It is probably obvious that these contractions are single vocables, but in support of that conclusion consider the utterances in (351).\(^{133}\) These show that repetition involves repetition of complete words: (351a−b) show that it is not just the last syllable that is repeated. (351d−e) show that *you’re* is a single phonological word (i.e. a vocable).

\[(351)\]

a. a polite... polite... polite person

b. † a polite... lite... lite person

c. you are... are... are mistaken

d. † you’re... are... are mistaken

e. you’re... you’re... you’re mistaken

It is extremely hard to see how these articulated prepositions and these subject–auxiliary contractions can be analysed as single terminal nodes in syntactic structure, especially in those cases where the articulated-preposition bleeds the possibility of the unfused preposition–article sequence. Syntactically, the articulated prepositions constitute a preposition whose complement is a determiner, and the contracted auxiliaries constitute an auxiliary whose subject, in the cases cited, is a personal pronoun. This conclusion is supported by (352a−d). The Spanish example (352a), from McCawley (1989), suggests that *del* truly represents a DE + EL sequence, since the DE has other (nonad-

---

\(^{132}\) These pronunciations are for the strong forms in my idiolect. Wells (1990) confirms these pronunciations of *you’re* and *they’re* for RP, but gives only *we’re* = *weir*.

\(^{133}\) ‘†’ here indicates something incompatible with the normal rules of performance. It is intended as a performance counterpart of ‘*’.
jacent and unfused) complements, *la* and *las*. In (352b) we have a pronoun that is subject of two auxiliaries, with one of which it is fused. In (352c–d) the auxiliary has two subjects, with one of which it is fused.

(352) a. Puerto Rico ... había salido victorioso del mal tiempo, *la* posible turbulencia politica, y *las* dificultades logisticas. ["Puerto Rico has emerged victorious from the bad weather, the possible political turbulence and the logistic difficulties."]

b. We’re hungry and won’t wait any longer.

c. Either Edgar or Sophy’s got to do it.

d. He or she’d do it like a shot.

Since some of the fusions, especially the Italian ones, could be analysed as a vocable consisting of a sequence of two non-maximal morphological units, there might seem to be available an analysis wherein non-maximal morphological units can constitute syntactic terminals. Many of the fusions resist this analysis, though. There’s not much grounds for a morpheme boundary in, say, /jo:/ (you’re), /we:/ (we’re), or — most glaringly — in /o/ (au). As it happens, for these last three a case could nonetheless be made for discerning a combination of two non-maximal but non-concatenated morphological units. If we assume an element-based phonological theory such as Government Phonology, we could offer the following analysis:

• *au*

An instance of the morpheme A is the enunciation of the preposition à. An instance of the morpheme U is an allomorphic enunciation of le. The A and the U fuse phonologically in a single segment realized as [o].

• *you’re*

An instance of the morpheme IU is the enunciation of you. An instance of the morpheme A(r) is the enunciation of are. The U and the A fuse phonologically in a single segment realized as [o], giving [jo:(r)].

——

134 I have been unable to elicit analogous Italian examples from my Italian informant.
135 In my lect; /wi@/ in some others.

[206]
• *we’re*

An instance of the morpheme UI is the enunciation of *we*. An instance of the morpheme A(r) is the enunciation of *are*. The I and the A fuse phonologically in a single segment realized as [e], giving [we:(r)].

However, the assumption that the ability of a vocable to effectively occupy two syntactic terminals is contingent on it containing two non-maximal morphological units such that each non-maximal unit is a syntactic terminal makes very implausible predictions. Suppose French underwent a phonological change such that /o/ became /u/. The phonological form of *au* would contain but a single element, U. There would simply not be enough phonological material to constitute more than one morphological unit. The prediction, which is surely incorrect, would therefore be that such a phonological change would necessarily precipitate a radical alteration in French syntax.

We must conclude that syntactic terminals do not necessarily correspond one-to-one with morphological units, maximal or otherwise. One morphological unit can correspond to more than one syntactic terminal. We can capture this by taking the view that syntactic terminals are in a hand-in-hand relationship, rather than a pantechnicon relationship, with vocables (or concatenations of vocables). A word (i.e. a syntactic terminal) has an enunciation, and the word is not the enunciation, and they belong to different sorts of category; words belong to syntactic and lexeme classes, while vocables belong to classes of phonological structure and to morpheme classes. On reflection, this view turns out to make more sense than the initially more intuitive pantechnicon view: it does not seem strange that the phonological forms of a sequence of two words might blend together, and nor, since such a blend is merely a matter of phonology, is it strange to take the resulting blended form to have no effect on the syntactic structure, so long as the correspondence between the blended form and the underlying word sequence is learnable.

One of the advantages of the hand-in-hand enunciation-based approach is that it explains why the internal morphological structure of words is invisible to syntax: the units that compose syntactic structure have no internal structure. Other manifold advantages are the analyses it affords. §5.4 shows how ‘coenunciation’ — when two or more words share the same enunciation — can solve certain recalcitrant problems with interrogative clauses, with gerunds, and with demoted subjects of unactive verbs, and in §5.5 an even more recalcitrant problem with depictive adjuncts is solved using coenunciation.
The rest of the chapter thereafter seeks to discover coenunciation elsewhere in the grammar, most importantly in the syntax of verbs and auxiliaries.

I shall discuss in the next section the ontological status of words (within the mental model of language and the world in general). In the meantime I’ll simply diagram some illustrative examples of coenunciation — (353a–d). In these diagrams words are represented by the names, in capital letters, of the lexemes they belong to, and the enunciation of each word is represented immediately below it in normal orthography. Sequences of words with the same enunciation are underlined.

\[
\begin{align*}
\text{of c} & \\
of c \\
\text{NOM DE \ LS PARISIEN} & \\
nom du \ parisien \\
\text{of c} & \\
of c \\
\text{NOM DE \ LS PARISIENNE} & \\
nom de \ la \ parisienne \\
s \text{of} & \\
o \text{of} & \\
Y \text{OU BE RIGHT} & \\
\text{YOU BE RIGHT} & \\
o \text{of} & \\
go \text{ of} & \\
Y \text{OU BE RIGHT} & \\
\text{YOU BE RIGHT} & \\
\end{align*}
\]

(353) a.

b.

c.

d.

5.2.4 Words and their enunciations

A central, and rather insightful, claim of WG is that words are conceptualized as instances of actions. They therefore are naturally sequenced with respect to each other, just as any group of events would be. I assume that every word is associated, by the

---

136 In general, the name of the lexeme should be construed simply as a diagrammatical device for marking the presence of a word, rather than as a theoretically well-motivated or well-thought-out claim about which lexeme the word actually belongs to.

137 I only use this notation when it is relevant to do so. In diagrams elsewhere, words are represented notationally by the normal orthographic representation of their enunciation.

138 This claim pays off in particular in the analysis of deixis, of illocutionary adverbs, and of grammatically encoded illocutionary force (e.g. imperatives). In brief, the characterization of the meaning of such a word involves the word itself; it features in its own definition.

139 For this reason we don't take the view (e.g. of HPSG (Pollard & Sag 1987)) that syntax is defined only in relations of 'immediate dominance' while linear precedence relations are found only in the morphological or phonological components. But see §5.5.
Enunciation relationship, with a concept of a phonetic action\textsuperscript{140} that is, as we shall see below, composed of vocables and that represents the word — the lexical action. The concept of the phonetic action is the enunciation of the concept of the lexical action. The enunciation functions as evidence or a symptom of the word.

In semiotic terms, we are making a distinction between producing a signans and employing a sign. A word is an act of employing a sign. Its enunciation is the act of producing the signans. These might seem to be the same thing, but they’re not. While employing the sign entails producing the signans, producing the signans does not entail employing the sign. It is possible, for instance, to enounce \textit{hello, good morning, and welcome} in a phonetics class without thereby greeting anyone.

Current WG practise holds that every word in an utterance has an enunciation; no word is phonologically vacuous. This is a stipulation, but no more costly or stipulative than saying only some words have an enunciation. It is allowed that performance can leave some enunciations unrealized phonetically, as in utterance-initial ‘ellipses’ like \textit{Coming?} or \textit{Don’t think so}. WG’s position\textsuperscript{141} at present on these cases is that the grammar generates such sentences complete with the omitted \textit{Are you} and \textit{I}, including their phonology; all that happens is that the articulatory apparatus is late in switching on. At any rate, as far as the grammar is concerned, every word in a sentence is to be pronounced. The requirement that every word be associated with an enunciation must aid learning, for it would be difficult to learn about the behaviour and properties of phonetically vacuous words.

\textsuperscript{140} This model would in principle also allow for words to be associated directly with graphetic forms without phonology being involved. But in fact I believe that is not how orthography works. I believe that orthographic forms represent not words but their enunciations. That is, \textit{<cat>} is the graphology of the enunciation of instances of CAT, rather than the graphology of instances of CAT themselves; a spelling is the spelling of the sound of a word, not the spelling of the word. This helps to explain why so many native speakers’ spelling is confused by homophones like \textit{their/they’re/there} — they are all graphologies of a structure /\textipa{t}/, and learning and applying the criterion that distinguishes them involves conscious or semi-conscious consideration of the relationship between the /\textipa{t}/ and the word it is the enunciation or morphological-phonological form of. A prime example of this phenomenon involves ‘s. Given that the syntactic difference between \textit{years} and \textit{years’} is considerable, and something every speaker of English knows, how can we explain the near-ubiquitous graphological confusion between them? For example, the spelling \textit{20 years’ old} appeared in The Economist (2 April 1994, p100) after having passed through the word processors of several literate and fastidious journalists and subeditors. If \textit{<years>} were the spelling of \textit{years} and \textit{<years’>} the spelling of \textit{years’}, it would be hard to explain the mistake. But if \textit{<years>} and \textit{<years’>} are both spellings of an instance of the structure /\textipa{z}/, we have the basis for an explanation of the confusion.

\textsuperscript{141} By which I mean that this is the view of Dick Hudson and me. Neither of us has investigated the phenomenon in depth. This interpretation of the data was put forward by Hudson in a message to the Linguist email discussion group.
It is incontrovertibly the case that words' enunciations are sequenced. It is less obvious, however, that the sequence of words is the same as the sequence of their associated enunciations, though we take it to be true that the sequence is the same. We assume, though, that associations pattern as in (354a), without crossing association lines, rather than as in (354b) with crossing association lines. 'W' stands for a word, a lexical action and 'V' for a vocable, a phonological action.

(354) a. $\begin{array}{c}
W_1 \\
V_1 \\
W_2 \\
V_2 \\
W_3 \\
V_3 \\
W_4 \\
V_4 \\
\end{array}$

b. $\begin{array}{c}
W_1 \\
V_1 \\
W_2 \\
V_2 \\
W_3 \\
V_3 \\
W_4 \\
V_4 \\
\end{array}$

The association pattern in (354b) needs to be ruled out by the rule (355).

(355) The enunciation of a follower of any word $W$ does not precede the enunciation of $W$.\textsuperscript{142}

To this rule there are certain tightly circumscribed and highly stylistically marked exceptions for right node raising and related phenomena, as we are to see in §5.3.

As already mentioned, every word is associated with a non-void phonological representation. Rule (356) states this, and thereby prohibits the association pattern in (357).

(356) For every word, $W$, something is the enunciation of $W$.

(357) $\begin{array}{c}
W_1 \\
V_1 \\
W_2 \\
V_2 \\
W_3 \\
V_3 \\
W_4 \\
\end{array}$

However, the obverse of (357), in which not every vocable is associated with a word, as in (358), is possible.

\textsuperscript{142} It is worth noting that no other rule of grammar appears to require the notion 'follower' — i.e. 'word that follows $W$'.

[210]
An example of this phenomenon is the use in speech of so-called vocalized pauses like *uuh*; that is, *uuh* would correspond to V2 in (358). One might have imagined that such things were invisible to the processes that derive utterances from the grammar, so that in terms of how (358) conforms to the grammar V2 is not really present, but in fact this appears not to be so. The evidence for this comes from the articles *an* and *the*, which have the phonological form */@n/* and */ôi/* if they’re followed by a vowel, and */@/ and */ô@/* if they’re not. The alternation is peculiar to these words; there are no other words that lose a final */n/* when followed by a consonant or gain one when followed by a vowel, or that alternate */@/ and */i/* according to whether a consonant or a vowel follows.¹⁴³ Now, when the phonologies of an *an* or *the* are followed by *uuh*, they get pronounced as */@n/* and */ôi/* . The *uuh* must therefore be visible to the phonological actions associated with an *an* or *the*.

If associations from words to vocables can’t cross, it might still be possible for them to diverge or converge, as in (359a–b) respectively.

Examples that might appear to involve divergence are sequences of vocables that appear to collectively constitute the enunciation of a single word. One example is *by and large*. This apparent phrase looks like a coordination of a preposition and an adjective, and functions as a kind of adverb. How come the grammar doesn’t choke on such a monster? Perhaps the answer is that we have a single lexeme, *by-and-large*, the enunciation of the instances of which is a sequence of the stems associated with the lexemes *by*, *and* and *large*. Another large class of examples is names that aren’t proper nouns, such as *Pigs Might Fly* (the name of a horse) and *Ladies whose bright eyes* (the title of a

¹⁴³ See §5.2.8 for a brief consideration of the morpho-phonological mechanism underlying the variation.
The same goes, perhaps, for direct speech. Within a sentence, names like these, and perhaps also direct speech, function as a single word. But the names themselves consist of the vocables of the words of a potential utterance. On the whole, I think it is probably best to say that these examples involve a single word with a single enunciation, but with this single enunciation exceptionally consisting not of a single vocable but of a sequence of vocables. We therefore maintain the rule that every word has exactly one enunciation, and refine (356) to the more restrictive (360).

(360) Every word has exactly one enunciation.

We can diagram the enunciation of a by-and-large by (361). The single line represents the enunciation relationship, and the double lines indicate that the vocables are parts of a vocable sequence that is an enunciation.

(361) by-and-large

by and large

Is convergence of associations from words to vocables also prohibited? There is a default rule that says “for every enunciation of every word, there is only one word it is the enunciation of” (cf. “for every hand of every person, there is only one person it is the hand of”). This default rule is overridden in cases such as those presented in the previous section and in other cases to be presented in following sections. In these exceptional cases one (possibly singleton) vocable sequence is the enunciation of two words. These coenunciations involve the association pattern in (362a–b).

(362) a. YOU ARE

you’re

144 This shows, I think, that the morphological processes that create vocables that are enunciation of more than one word can operate independently of the coenunciation. Consider the italicized vocables in (i–ii). The vocable man’s and the vocable sequences pig’s ear and dog’s breakfast are each the enunciation of a single word (instances of the lexemes MAN’S, PIG’S-EAR and DOG’S-BREAKFAST). These examples show that the composition of the vocables man’s, pig’s and dog’s is not contingent on them being the enunciation of the possessive determiner and of the word preceding it.

(i) three green man’s shoes
(ii) She made a right {pig’s ear/dog’s breakfast} of it
Let me sketch a scenario which might make this idea of words with the same enunciation easier to grasp. Take the task of counting sheep; one can't count the second sheep without having counted the first, one can't count the fifth without having counted the fourth, and so on. One way to count sheep is one, two, three, four, five, six etc. But imagine that the step of counting the fourth sheep and the step of counting the fifth sheep could be simultaneously signalled by uttering the sound zim. Then sheep could be counted one, two, three, zim, six... This seems quite reasonable, and quite an efficient way to count sheep (counting one, two, three, five might mean the sheep counter had forgotten to count four, and would thus end up with a count one greater than the total number of sheep); zim is the enunciation of both the act of counting the fourth sheep and the act of counting the fifth sheep.

I am not sure how we should analyse the combination of coenunciation with divergence of associations from words to vocables, as in Pigs Might Fly's racing round the bend. It could be as in (363a), where Pigs Might Fly's is enunciation of both the PIGS-MIGHT-FLY and the be, or as in (363b), where Pigs Might Fly's is enunciation of the PIGS-MIGHT-FLY and Fly's is enunciation of the be.

(363) a. PIGS-MIGHT-FLY BE
    Pigs Might Fly's

b. PIGS-MIGHT-FLY BE
    Pigs Might Fly's

Allowing only (363a) would rule out bidirectional coenunciations, as in (364), where V2 shares an enunciation with V1 and with V3, but where V1 and V3 don't share an enunciation.

(364) W1 W2 W3
    V1 V2
It is worth mentioning at this juncture that the analysis being developed in this section has no bearing on the nature of idiomaticity. Idioms consist of several words, each with its customary enunciation (and without its customary meaning). For example, *do it over* ("rob it") has the enunciation structure in (365a) rather than the structure in (365b). See Hudson (1990: 149) for a sample WG treatment of idioms.

(365) a. \[ \text{DO IT OVER} \]
    \[ \quad \text{do it over} \]

b. \[ \text{DO-OVER IT} \]
    \[ \quad \text{do it over} \]

5.2.5 The relationship between phonology and syntax

At least for English, there are two extreme positions we can take on the relationship between phonology and syntax. On grounds of simplicity, either of these extreme positions is preferable to some compromise between them.

One of the extremes is taken by some structuralist theory. Syntax and phonology are linked pantechniconwise. Put another way, syntax is morphology, and therefore phonology. The basic unit of syntax is the morpheme. If it is correct to assume that morphology is constituency-based, either hierarchically or terms of slot/filler templates, then syntax is at least partly constituency-based, and any use of dependency would be an enrichment of the basic elements of the model. Words, if they exist, are merely a kind of phrase, a higher level grouping of morphemes. Word-and-paradigm morphology is not available to this model.

The other extreme is the one I am taking for WG. Syntax and phonology are linked hand-in-hand. This allows for word-and-paradigm morphology. Syntax is not phonology and therefore not morphology. The job of the phonological-morphological component is to specify the sound structure of the enunciation of words, and to create new lexemes. The basic unit of syntax is the word. Syntax cannot access parts of

---

145 In English, at least, though maybe also, but not necessarily, in morphologically very different languages, e.g. Altaic, or Athabaskan languages.

146 I envisage such lexeme creation as working in the following way. Morphemes are associated with meanings. When morphemes combine, so do their meanings. A lexeme is the 'lemma' of a morphological combination. The lemma of a combination of *un-, re-,* the stem of *read,* the stem of *able* and *-ity,* into *un-re-read-abil-ity* is the lexeme of which the word *unrereadability* is an instance. The meaning of the lexeme is inherited from the morphological combination that it is the lemma of.
words, since words have no parts; words don’t consist of morphemes or of phonemes. Since, other things being equal, dependency structures are simpler than constituency structures, e.g. in requiring fewer nodes, and since the structural devices used in morphology are not necessarily those used in syntax, we can assume, for the sake of simplicity, that syntax is dependency-based.

The former model is simpler, but inferior to the latter on both empirical grounds and conceptual grounds. We predict that knowledge of the non-productive morphological structure (i) is not necessary in order to know the word, (ii) can vary greatly between speakers, and (iii) can be acquired gradually over a lifetime. We predict that radical deformation of phonological structures, either diachronically or, say, in allegro production, needn’t necessarily affect the corresponding syntactic structure, if there remain adequate means for the learner or addressee to discover the syntactic structure. We predict that syntax is blind to morphology and phonology, instead of having to stipulate this. We allow for the possibility that the same bit of phonological structure corresponds to more than one bit of syntactic structure; it is this that much of this chapter is concerned with.

To the extent that Hudson’s (1984, 1990) version of WG is committed to a position on the relationship between syntax and phonology, it is intermediate between the pan-technicon and hand-in-hand approaches, but is closer to the former. He assumes that terminal nodes in syntax must be words, but for certain cases — clitics, compounds, gerunds — where there is a mismatch between syntax and phonology, such that one phonological word seems to correspond to more than one syntactic word, he allows that a word can contain other words. He is assuming that phonological criteria are relevant to wordhood, and that words are composed of a constituency structure of morphemes, and, ultimately, of sound. The analysis relies on the internal structure of the phonologically-defined word being agglutinative. These examples that motivate his conclusion can be reanalysed in terms of my simpler model; see §5.2.7.

The hand-in-hand analysis developed here may seem reminiscent of Sadock’s (1985, 1991) theory of Autolexical Syntax. But while both approaches are initially motivated by the need to accommodate mismatches between syntactic and phonological structures, the actual architecture of the two models is different. In the present analysis the syntactic plane and the morphological-phonological plane are distinct but hand-in-hand. In Autolexical Syntax, as far as I can tell, the syntactic plane and the morphological-phon-
ological plane are the same, but there are two separate structurings for syntax and for morphology, and these structurings can conflict with each other, which, as Sadock (1994) points out, makes Autolexical grammar polystratal. There is no polystratality in WG.\textsuperscript{147}

5.2.6 Types of coenunciation

I’ll assume, if only for descriptive convenience, that a word that the grammar requires to share its enunciation with another word is an instance of a category ‘Coenunciate’. At least descriptively, coenunciates are of various different types. One thing all coenunciates have in common is that they share their enunciation with an adjacent word. This is a result of the prohibition against associations between words and their enunciations crossing.

The first factor differentiating coenunciates is which word the coenunciate is coenounced with: most coenunciates must share their enunciation with a specified dependent (e.g. subject, object), while a minority are coenounced with an unsuperordinate of a specified dependent. (An unsuperordinate of W is W or a subordinate of W.) Examples of the latter type are contracted auxiliaries, which are coenounced with an unsuperordinate of their subject, and possessive ‘s, which, as explained in §6.1.4, is coenounced with an unsuperordinate of its inject.

A second factor is whether, for a given lexeme or other class of words, all or only some instances of that word class are coenunciates. For example, instances of some types of auxiliary optionally coenounce with the preceding word, while, as we shall see later, instances of certain other types of auxiliary obligatorily coenounce with their complement. A third and related factor is whether in a given context all instances of the word class are coenunciates. An auxiliary immediately following its subject may or may not be a coenunciate. In contrast, as discussed in §6.1.2, a sequence of deictic this plus day is coenounced as today if and only if they’re adjacent, but they’re not always adjacent (e.g. in this very day).

Additionally, rules for coenunciates must specify what the morpho-phonological form of their enunciation is.

\textsuperscript{147} It may be, however, that this can be claimed only because WG has had so little to say about morphology and phonology. It seems quite likely that the morphological-phonological plane is polystratal, with phonological structuring (e.g. syllabification) and morphological structuring conflicting.
5.2.7 Clitics and other mismatches between syntax and phonology

Definitions of exactly what counts as a clitic vary, but the major feature of clitics is that they involve one vocable that corresponds to more than one syntactic terminal. Debate about clitics often involves deliberation of the issue of whether some morphological unit is a clitic or an inflection. This issue is actually a conflation of two independent questions:

- First, does the vocable containing the putative clitic genuinely correspond to more than one syntactic terminal?
- Second, for each syntactic terminal that the vocable corresponds to, is it possible to identify within the vocable a morphological unit that corresponds to that syntactic terminal?

The reason why these questions are not usually seen as independent is that, as I observed earlier, it is generally assumed that it is morphological units that are syntactic terminals. Therefore, on that view the answer to the first question can be yes (a single vocable can correspond to more than one syntactic terminal) only if the answer to the second question is yes (it is possible to identify within the vocable a morphological unit that corresponds to each syntactic terminal) – unless, that is, phonologically vacuous morphological units are permitted. I've been arguing, however, that syntactic terminals are not morphological units; every syntactic terminal has an enunciation, which is either a vocable or a sequence of vocables. In this case, the answer to the second question is entirely irrelevant to syntax; it has no bearing on the answer to the first question. More generally, morphological structure does not constrain syntactic structure.

In consequence of the above, my analysis of standard clitic data is somewhat different from usual analyses. In *Sophy les aime*, for example, there are two vocables: *Sophy* and *les aime*, and there are at least three words. There is a proper noun, whose enunciation is *Sophy*, there is a verb, whose enunciation is *les aime*, and there is a pronoun, whose enunciation is the verb's enunciation – i.e. *les aime*. The morphological structure of the verb's and the pronoun's enunciation is composed of a bit of morphology (*les*) associated with the pronoun, and a bit of morphology (*aime*) associated with the verb. One consequence of this analysis is that the order of morphemes in *les aime* needn't reveal the order of the various words it is the enunciation of.

---

This approach to clitics is new to WG, even though the separation of word and enunciation has been lurking as an unexplored potential in WG, e.g. in Hudson (1990)'s Word-and-Paradigm treatment of morphology. What WG treatments of clitics there have been, most notably Volino (1990), and also, briefly, Hudson's (1990) discussion of English gerunds, in practice take a traditional view in which a word is not distinct from its enunciation. Finally, in Hudson's (1984) analysis of clitics, carried over without revision into Hudson (1990), one word is allowed to contain other words; the clitic and the host are words that are daughters of another word. Clearly, I've rejected this analysis, believing it results from a failure to separate words from their enunciations, or to recognize the implications of that separation. Once we do recognize this separation, it is hard to conceive of a way to make sense of the notion that one word can be daughter of another; put another way, it is hard to see what would distinguish, on the one hand, a word that is a mother, from, on the other hand, a phrase.

I turn now to other examples where phonological criteria for identifying 'words' conflict with syntactic criteria. An example from English is when nouns premodify nouns. Oversimplifying rather, stress normally falls on the last 'word' in the phrase (i.e. on the last vocable), e.g. young ADult, but in when one noun premodifies another the stress falls on the first, e.g. ADult rate. This is the stress pattern found with true compounds, e.g. GREENhouse, BUTterfly, which are single words, but ADult rate can't be a single word, since the premodifying word can itself be modified by an adjective, e.g. young ADult rate. This is not possible for true compounds; cf. !pale GREENhouse and !rancid BUTterfly, which contrasts with the homophonous but syntactically distinct rancid BUTterfly ("a fly associated with rancid butter"). A possible analysis for these pseudocompounds is that they are coenunciative: ADult rate is a single vocable, as evidenced by its stress, and is enunciation of both the instance of ADULT and the instance of RATE.

Some similar examples are discussed as problems in Spencer (1991: 41—43). In Czech, stress falls on the first syllable of the word (i.e. vocable), but "a monosyllabic preposition before an unmodified noun will usually attract stress to itself". Thus we get ten STUL, "that table", na ten STUL, "onto that table", and, with stress shift, NA stül, "onto a/the table". This can be analysed as involving coenunciation. If a NA immediately precedes its complement then it shares its enunciation with its complement. The enunciation is a single vocable composed of the morphological form proper to the pre-
position and the morphological form proper to the noun. Because it is a single vocable the stress falls on the first syllable.

Another of Spencer's examples is the well known problem of Latin -que. As (366a–b) show, -que causes the stress shift that we would expect if puerîque is a single vocable. (Stressed nuclei are italicized.)

(366) a. Puellæ et puerî canunt. [“the girls and the boys sing”]
     b. Puellæ puerîque canunt. [“the girls and the boys sing”]

The problem with taking -que to be a suffix is that it can attach to anything: if two clauses are conjoined, it attaches to the first word of the clause, irrespective of what that word is. Spencer's (367a–b) show this. Further, if W+que is the enunciation of a single word, it becomes difficult to analyse these examples as involving coordination.

(367) a. canuntque fēminaē [“and the women sing”]
     b. haecque canunt fēminaē [“and these women sing”]

Even if we ignored the matter of stress and simply took -que to be a separate word, we'd find that it always occurs between the first and second words of the final conjunct. Even though Latin is notorious for its tolerance of discontinuity, there remains the problem of accounting for this second position of -que. Coenunciation offers an extremely simple solution to all these problems. Suppose an et can optionally coenounce with the following word, W, the shared enunciation being a single vocable consisting of the morphology proper to W plus the morpheme -que. (368) shows this. Then the problems of syntax and phonology vanish. -Que shifts stress because it is part of the same vocable as its host. No word -que occurs as the second word within a conjunct; the second position is an illusion created by the morphology of the vocable that is enunciation of the ET and the HAEC.

(368) PUELLA ET Puer CANO ET HAEc CANO FEMINA
     puellæ et puerî canunt haecque canunt fēminaē
5.2.8 Articles revisited

An, The. A purely phonological explanation is in principle possible for the alternations /@ ~ @n/ and /ô@ ~ ôi/. Suppose the words have a single underlying phonological form with the segmental content /@n/ and /ô@i/, but with the /n/ and /i/ not associated to a skeletal timing unit. If a consonant-initial word, which therefore has a filled onset, follows, then the /n/ and /i/ would be unable to associate with a a skeletal position, but if a vowel-initial word follows, and if vowel-initial words begin with an empty onset, then the /n/ and /i/ can fill the empty onset position. The eventual form /ôi/ rather than /ô@i/ would result from the /i/ spreading into the preceding nucleus, by a process that by some means or other fails to apply to, say, the yacht or rather yellow. However, since these two alternations are specific to one lexeme each, and therefore show no indication of being instances of prevalent, across-the-board, phonological processes, it might be better to analyse them as phonologically-conditioned allomorphy. I assume that triggers of allomorphs of a morpheme are within the same vocable as the allomorph, so since the allomorphs of the articles are triggered by the immediately following phonological material, it follows that the morphology of articles is in the same vocable as the morphology of the following word. This can be achieved by having the article have the same enunciation as the following word, as shown in (369a–c). As (369c) shows, it looks as if voiced pauses can be part of the phonological structure of a word (and so can trigger the /ôi/ allomorph).

(369) a. EAT THE FISH FIRST
    |   |   |
    eat [ô@] fish first

 b. EAT THE ORANGE FIRST
    |   |   |
    eat [ôi] orange first

 c. EAT THE FISH FIRST
    |   |   |
    eat [ôi] uuh fish first

Le. The realization of Le as l' apparently bleeds fusions like au (= A + LE) and du (= DE + LE). As with English articles, the alternation le ~ l', /læ ~ l/, is conditioned by whether an empty onset follows (Kaye & Lowenstamm (1984), Durand (1990: 204), Harris (1994: 179)), and also as in English, the alternation is lexeme- or morpheme-specific (or at least it is specific to a limited class of lexemes or morphemes, com-
prising, e.g. me, te, se). Thus, just as in English, for the allomorphy of the morphology of a LE to be triggered by the following phonological context, the enunciation of the LE should be shared with the following word, as shown in (370a–b).

\[
\begin{align*}
\text{VOIR LE MATIN DEMAIN} \\
\text{voir le matin demain}
\end{align*}
\]

\[
\begin{align*}
\text{VOIR LE ARBRE DEMAIN} \\
\text{voir l'arbre demain}
\end{align*}
\]

Then, the rule for DE + LE and A + LE coenunciations simply gives the preposition the same enunciation as its LE complement; the rule is blind to the allomorphy. The word—enunciation associations are as shown in (371a–b).

\[
\begin{align*}
\text{NOM DE LE MATIN DEMAIN} \\
\text{nom du matin demain}
\end{align*}
\]

\[
\begin{align*}
\text{NOM DE LE ARBRE DEMAIN} \\
\text{nom de l'arbre demain}
\end{align*}
\]

The actual phonological realization of the enunciations is a matter of morphology, the workings of which are outside the scope of this discussion. However, in order to make this buck-passing palatable to the reader, I will sketch very very roughly how it might work. The relevant conclusion to draw is that there is at least one way this could work, even if there may be numerous ways and the one offered here is wrong.

From the data in (372) we can derive rules such as (373–374). These rules recognize two varieties of stem: prevocalic stems are followed by a vowel-initial morpheme and preconsonantal stems are followed by a consonant-initial morpheme.

\[
\begin{align*}
\text{(372) mon cousin ma cousine} \\
\text{mon ami mon amie}
\end{align*}
\]

\[
\begin{align*}
\text{le cousin la cousine} \\
\text{l'ami l'amie}
\end{align*}
\]
(373) a. stem of MON = mon
    b. preconsonantal stem of fem. MON = ma

(374) a. prevocalic stem of LE = l’
    b. preconsonantal stem of masc. LE = le
    c. preconsonantal stem of fem. LE = la

Next, assume a rule that says that by default, for every X, X a lexeme and subclass of articles, the stem associated with the intersection of X and the class of words that coenounce with a preposition is made up of the stem of X and the stem of the preposition. Assume too that the enunciation of a coenunciate article is a realization of the stem of the article. This default correctly gives us le la, de l’. It also wrongly gives us de le, so for this case we override the default rule with a rule like “preconsonantal stem of masc. LE coenounced with DE = du”. Let me reiterate that the purpose of this analysis is merely to indicate that there is a feasible account of how the realization of le as l’ bleeds the realization as du.

5.3 Right Node Raising and related phenomena

In this section I employ coenunciation to provide an analysis of right-node-raising and similar phenomena. The analysis requires us to accept certain exceptions to the rules presented in §5.2.4 that govern the correspondence between words and their enunciations.

5.3.1 A coenunciative analysis of ‘node raising’

Right-node-raising (RNR) is standardly taken to be a form of coordination. (Accordingly, Quirk et al. (1985) call it “interpolated coordination”.) Typical examples are (375a–c).

(375) a. She discovered __ and thought about __ the problem.
    b. I wanted to buy __, but failed to find __, the book.
    c. She made Sophy take pictures of __ and Bill be friends with __, anyone who called at the front door.
Probably the standard view of RNR is that it involves a kind of across-the-board (ATB) rightward extraction out of conjuncts. I would argue that this view is wrong: (i) RNR is not a variety of coordination or necessarily associated with coordination, and (ii) it doesn’t involve anything akin to extraction. I’ll address claim (i) first, the validity of which is easily demonstrable by data like (376a–d). We might argue that the ‘subordinating conjunction’ though in (376a) is behaving like a coordinating conjunction here (as indeed THOUNGS sometimes can — see §3.7) but in (376b–d) there is no word that that can be taken to be acting as a coordinating conjunction. Examples (376c–d) are taken from Postal (1994) who takes them from Hudson (1976); both Hudson and Postal use them to make the same point I’m making with this data.

(376) a. John left __, though tried to eat __, the pudding.
    b. John left __, having tried to eat __, the pudding.
    c. Of the people questioned, those who like __ outnumbered by two to one those who disliked __ the way the devaluation of the pound had been handled.
    d. It’s interesting to compare the people who like __ with the people who dislike the power of the big unions.

Further support for the conclusion that RNR is not a form of coordination in opposition to the mutually incompatible symmetric, gapping and supplementary varieties, but rather is something orthogonal to them and combinable with them comes from the fact that it can combine not only with simple symmetric coordination, as in (375a–c), but also with complex symmetric coordination, as in (377a–f), and with gapping, as in (378).

(377) RNR plus complex coordination.
    a. John told each of the fathers their eldest sons __ and each of the mothers their youngest daughters __ were coming. [adapted from Pickering & Barry (1993)]
    b. She told him she __ and he told her he __ was going to be there.
    c. Tell Sophy he __ and Edgar she __ is too shy to make direct advances.
    d. Give him red __ and her blue __ books.
    e. Give him Sophy’s __ and her Edgar’s __ address.
f. Give him the correct postcode for Sophy's _ and her the remainder of Edgar's _ address.

(378) **RNR plus gapping.**

Sophy _ made _ sketches of _, and Edgar _ sculptures of _, each of the newly appointed laureates.

As for my second claim, that RNR does not involve anything akin to extraction, Postal (1994) shows that RNR freely violates island constraints on extraction. Postal also shows that what have, following Engdahl (1983), been taken to be parasitic gaps licensed by heavy NP shift, as in (379a-b) (Postal's (24c), itself from Authier (1991), and (32b) respectively), have very different properties from parasitic gaps licensed by regular extraction and can in fact be analysed as gaps created by noncoordinate RNR. (379b) involves both coordinate and noncoordinate RNR.

(379) a. We suggest _ to our employees, without actually requiring _ of them, that they wear a tie.

b. We could suggest _ to our employees, without actually requiring _ of them, and should suggest _ to our employees, without actually requiring _ of them, that they wear a tie.

For Postal, however, RNR does involve something akin to, though nonetheless different from, extraction. I wish to argue, however, that syntactically, RNR is illusory. This will explain:

'a property noted long ago by Wexler and Culicover (1980: 301) for standard (coordinate) RNR:

(152) "[A] raised [by RNR] node always behaves, vis-à-vis all constraints on analyzability, just as it would if it were in its original underlying position. Hence, whereas it is apparently possible to apply RNR to a constituent of a relative clause, if we then try to analyze this raised node, we find that it acts as though it were still within the relative clause."

This generalization was intended to cover cases like (153b).

(153) a. Mary buys t, and Bill knows a man who sells t, — [pictures of Elvis Presley],
b. Who does Mary buy and Bill (*know a man who) sells — [pictures of t]

Here, although RNR can yield a gap inside the relative clause, the RNR extractee acts with respect to L-extraction as if it had not been extracted by RNR.’ [Postal (1994: 110)]

The explanation is that “the RNR extractee” behaves as if it had not been extracted by RNR because it has not been extracted by RNR. The raised node “acts as though it were still within the relative clause” because it is still within the relative clause.

How can that be? My analysis rests on the distinction made in §5.2 between words, which are syntactic objects, and vocables, which are morphological-phonological objects (‘phonological words’). They are different objects, not merely different aspects of the same object, but normally occur in a one-to-one correspondence. (380a) shows a normal pattern of word—vocable correspondences, and (380b) shows the sort of pattern we get with RNR.150 Words are represented in roman, and vocables in italics.

(380) a. She will try to buy the book and he will try to sell the book
b. She will try to buy the book and he will try to sell the book

---

149 (i) is a slightly more effective rendering of this example.

150 It would be interesting to see how RNR affects the allomorphy of AN — whether it is sensitive to, say, the phonology of the following vocable, or to the phonology of the vocable associated with the following word. But the data, (i—viii) is mostly unclear.

(i) * {He read an} {book that had been recommended} {but she read the} {book that had been recommended}.
(ii) {He read a} {book that had been recommended} {but she read the} {book that had been recommended}.
(iii) ? {He read an} {article that had been recommended} {but she read the} {article that had been recommended}.
(iv) ? {He read a} {article that had been recommended} {but she read the} {article that had been recommended}.
(v) ? {He read an} {book that had been recommended} {and she read the} {book that had been recommended}.
(vi) ? {He read a} {book that had been recommended} {and she read the} {book that had been recommended}.
(vii) ? {He read an} {article that had been recommended} {and she read the} {article that had been recommended}.
(viii) ? {He read a} {article that had been recommended} {and she read the} {article that had been recommended}.
(381) is an informal statement of the rule that permits RNR.

(381) If (i) A, B, C and D are sequences of words in the sequence A < B < C < D or D < C < B < A, and (ii) every word in C either (a) is s-dependent of another word in C or (b) is s-dependent of a word in D or (c) is unsubordinate of every word in C, then (α) the vocables corresponding to D can also correspond to B, and (β) associations between words in B and their enunciations are exceptionally allowed to cross associations between vocables and words not in B.

(β) is an exception to rule (355) in §4.2.4.

This allows the patterns diagrammed schematically in (382a–b). The pattern in (382b) is a kind of mirror image of RNR — we might call it Left Node Raising (LNR), and the phenomenon as a whole we could therefore call Node Raising. Examples of LNR are (383a–d).

(382) a. \[ \begin{array}{cccc}
A & B & C & D \\
\mid & \mid & \mid & \mid \\
A & C & D \\
\end{array} \]

b. \[ \begin{array}{cccc}
D & C & B & A \\
\mid & \mid & \mid & \mid \\
D & C & A \\
\end{array} \]

(383) “Left-node-raising”.

a. *Young ___ boys wear pink and ___ girls wear blue.*

b. *recently ___ married men and ___ divorced women*

c. *Him making a mess angered her and apologizing appeased her.*

Like RNR, LNR can combine with gapping, (384), and complex coordination, (385).

(384) **LNR, gapping.**

*Young ___ boys *wear* pink and ___ girls ___ blue.*

(385) **LNR, complex symmetric coordination.**

*very ___ big black blocks and ___ little lilac lollipops*

\[151\text{ Something like ‘vocabulary deduplication’ would be apter but unpalatable to neologophobes.}\]
(381) as formulated does not require any sort of similarity between the words in sequence B and the word in sequence D. Thus, it predicts that node raising can apply to a structure like *She will see Edgar and he will contact Sophy* to yield, say, *She will see —Edgar and he will contact Sophy*. However, in this node raised structure the vocable Sophy corresponds to both the word Edgar and the word Sophy. I assume this contravenes the rule of grammar saying that a word Edgar is associated with a vocable Edgar. According to this analysis, then, words in sequences B and D must be associated with the same phonological-morphological tokens, but do not have to have the same syntax. This is borne out by (386) in which the vocable Anne is enunciation of an object, referring to a person, in sequence B and of an xcomp, referring to a name, in sequence D. (The braces indicate word sequences, not constituents. The words in sequence B whose enunciations do not occur at the same time as the words themselves are subscripted.)

(386) \{She met\}_A, \{Anne\}_B \{and he named his child\}_C, \{Anne\}_D.

My countenancing (386) does not mean that I necessarily would similarly predict that (387a—b) are grammatical.

(387) a. * She bought a round and he turned — round.
   b. * He tried to walk and she went for a — walk.

Vocables are morphological as well as phonological entities, so homonymy is not a sufficient condition for RNR: if vocable V corresponds to word X and word Y then the morphological properties V must have by virtue of its correspondence with X must be compatible with the morphological properties V must have by virtue of its correspondence with Y. In the case of (387a), it could be that the stem of ROUND\_Noun is distinct from the stem of ROUND\_prep even though the phonological properties of each stem are the same. (387a) would be bad because the vocable round is required to be at one and the same time an instance of two distinct stems, ROUND\_Noun's and ROUND\_prep's, a requirement which it is reasonable to presume cannot grammatically be satisfied. As for (387b), it may be that again we have distinct (but derivationally related) stems; for example, the stem of WALK\_v may be [wɔ:k] (the brackets indicating a morphological constituent), while the stem of WALK\_n may be a morpheme that contains the stem of of
**Walk** (i.e. [[wo:k]]). Since new lexemes are created by combining morphemes, and the meaning of the new lexeme is predictable, it seems quite plausible that morphemes as well as words are associated with meanings. A difference in meaning between morphemes is sufficient to recognize them as distinct morphemes. The morpheme *round* in *round-shy* contributes a different meaning to the whole from the meaning the morpheme *round* in *whip-round* contributes. Thus we have distinct morphemes. A similar argument holds for *walk-organizer* versus *tightrope-walk*. In contrast, to the extent that (386) is grammatical, we have evidence that the morpheme *Anne* corresponding to the noun that refers to the person and the morpheme *Anne* corresponding to the noun that refers to the name are the same, and that the difference in meaning of the two nouns is not shared by their morphology.

We move on to discuss other aspects of (381). Note that it is recursive in that A—B—C—D sequences may overlap, so that the C—D bit of one sequence corresponds to the A—B bit of another, giving sequences like \([A1=B1][C1=A2][D1=B2][C2=A3][D2=B3][C3][D3]\). We find such a pattern as this in (388).

(388) \{Sophy has written\}_{A1} \{numerous articles on Kartvelian\}_{B1} \{Edgar has read\}_{C1=A2}

\{numerous articles on Kartvelian\}_{D1=B2} \{Thalia has published\}_{C2=A3} \{numerous articles on Kartvelian\}_{D2=B3}

\{and Euphrosyne has reviewed\}_{C3} \{numerous articles on Kartvelian\}_{D3}.

The function of condition (ii) in (381) is to rule out examples like (389a—b). The italicized word contravenes condition (ii) because it is neither (a) s-dependent of another word in sequence C, nor (b) s-dependent of any word in sequence D, nor (c) un subordinate of every word in C. The variants in (390a—b) are bad both for the same reason and because they contravene word order rules — *black big blocks, *give tulips everyone.*

(389) a. *\{She gave\}_{A} \{everyone who called at their house\}_{B} \{roses and he gave\}_{C} \{everyone who called at their house\}_{D} tulips.

b. *She made big \{black\}_{D} \{blocks and little\}_{C} \{black\}_{B} \{balls\}_{A}.*
(390) a. * {She gave} \_A \{everyone who called at their house\}_B \{roses and he gave tulips\}_C \{everyone who called at their house\}_D.

b. * She made \{black\}_D \{big blocks and little\}_C \{black\}_B \{balls\}_A.

Note that in the light of this constraint, (379a–b) must be reanalysed as (391a–b), in which heavy NP shift applies.

(391) a. {We suggest to our employees} \_A \{that they wear a tie\}_B \{without actually requiring of them\}_C, \{that they wear a tie\}_D.

b. {We could suggest to our employees} \_A \{that they wear a tie\}_B \{without actually requiring of them\}_C, \{and should suggest to our employees\}_D \{that they wear a tie\}_D.

It is possible that in somelects there are syntactic constraints on sequence D just as there are on sequence C, for it is sometimes alleged that the raised node (i.e. sequence D) must be a constituent, though this may be wishful thinking, just like the false but ubiquitous claim that conjuncts must be constituents.\(^{152}\) Certainly I find (392) about as acceptable as an average RNR example.\(^{153}\)

(392) {She sold} \_A \{Sophy roses\}_B \{and he bought\}_C \{Sophy roses\}_D.

I’ll now survey sundry pieces of evidence that support various aspects of this analysis of node raising. Take (393a–d) (brought to my attention by Jasper Holmes). If a present tense, third person be has two subjects then we would expect its enunciation to have the form are, rather than is — She and he are/*is here. But if the vocable is is

\(^{152}\) For Bresnan (1974) the raised node must be a constituent. For Postal (1974: 125–129) the raised node must be a constituent and any constituent is eligible to be a raised node. On the basis of a very limited range of data, Grosu (1976) argues that for some speakers RNR can affect more than one constituent so long as these constituents were not sisters prior to the RNR. Abbot (1976) argues that there is no single-constituent constraint on RNR, and proposes that the unacceptability (“the relative ungrammaticality”) of some RNR sentences is due to “factors that make processing difficult or that produce stylistically inappropriate examples”; she also notes the variability of different speakers’ judgements on RNR and concludes that “[m]uch further work is required to determine what causes this variability and what factors operate generally as constraints on RNR”.

\(^{153}\) Analogous examples are asterisked by Abbot (1976), though she accepts Joan offered, and Mary actually gave, a gold Cadillac to Billy Schwartz. Steedman (1989) accepts (without comment) I will offer, and may give, this Cadillac to that man.
enunciation of two words, all is explained. (393b–c) are ungrammatical because *are* is enunciation of a BE with only a singular subject. (393d) is ungrammatical because the vocable *is* is enunciation of a BE with a plural subject.

(393) a.  {She knows he} {in here} {and he knows she} {is here}.

   b.  * {She knows he} {in here} {and he knows she} {are here}.

c.  * {She knows he} {in here} {and he knows they} {are here}.

d.  * {She knows he} {in here} {and he knows they} {is here}.

The assumption that sequence B is syntactically present means that it is only an illusion that (as Postal (1994) has shown) RNR violates island constraints. It is also an illusion that RNR violates such constraints as that requiring a complement of pronoun to be its s-dependent, and that requiring an unpromoted inject of a ditransitive to be adjacent to the verb. Cf. (394a–b).

(394) a.  {He read this} {book on recent advances in autosegmental phonology} {and she read that} {book on recent advances in autosegmental phonology}.

   b.  {What did he buy} {the child whose birthday it was} {and she sell} {the child whose birthday it was}?

If node raising were a form of coordination, we'd expect the 'raised node' to be ward of the supervener. That is, just as *the apples* is ward of *and* in *She'll core and peel the apples*, so we would expect *the apples* to be ward of *and* in *She'll try to core but fail to peel the apples*. But my analysis of node raising says that would be the wrong analysis, and some confirmation of this comes from 'LNR'. If we take (383a–c) and add correlatives in the position they'd take if the raised node was a ward, we find that the result is ungrammatical, (395a–c). They are ungrammatical because of discontinuity: the phrases *young boys, recently married* and *him making a mess* are interrupted by *both* which is not contained within these phrases.
(395) a. *{Young} {both boys wear pink and} {young} {girls wear blue}.
b. *{recently} {both married men and} {recently} {divorced women}.
c. *{Him} {both making a mess angered her and} {him} {apologizing appeased her}.

I predict, incidentally, that (396a–c) should be okay, because the correlative is to the left of all the LNR. They certainly are far better than (395a–c), and I find them acceptable, but — inexplicably — not as good as versions with the correlative absent.

(396) a. Both {young} {boys wear pink and} {young} {girls wear blue}.
b. both {recently} {married men and} {recently} {divorced women}.
c. Both {him} {making a mess angered her and} {him} {apologizing appeased her}.

Both RNR and LNR can occur, in combination with simple symmetric coordination, complex symmetric coordination or gapping, in a form where in each conjunct a dependent of the item shared by the conjuncts is left ‘abandoned’ apparently without a s-head within the conjunct. Examples of the possibilities are (397–402) (though such utterances are not doing their addressee any favours). The ‘abandoned’ words are italicized.

(397) RNR, simple symmetric coordination, abandonment.
a. I {do read long} {books about the history of medicine in certain regions of South East Asia}, {though would prefer to read short}, {books about the history of medicine in certain regions of South East Asia}.
b. {I believe that John} {climbed the mountain} {and Harry thinks that Mary} {climbed the mountain}. [from Pickering & Barry (1993)]
c. Remember {that he} {will be visiting} {and that she} {will be visiting}.

(398) LNR, simple symmetric coordination, abandonment.
a. (?Either) {friends} {(*either) of yours (*either) are fond of me or} {friends}
{of mine are fond of you}.
b. {Most} {boys wear pink and} {most} {girls wear blue}.
5.3.2 Some real and apparent problems with the node raising analysis

I turn now to some possible problems for the analysis of node raising I'm proposing. The most significant problem is Postal's claim that RNR obeys the Coordinate Structure Constraint — RNR must apply across the board. One of Postal's examples that illustrate his claim, (403) (Postal's (114b) with his annotation replaced by mine), is already ruled out by the condition on the continuity of sequence C. The offending word is italicized. (404a—b) (Postal's (113a—b)) are bad simply because of discontinuity — the raised node is in the wrong position to be the preceding conjunct of the conjunction.

(403)  * I purchased {pictures of} {the divinity you are trying to pay respect to} {and} {carvings of Zeus and Mary purchased videos of} {the divinity you are trying to pay respect to}.

(404) a. * {Frank offended} {his favourite uncle from Cleveland} {by not calling _ or recognizing Glen} {his favourite uncle from Cleveland}.

b. * {Frank claimed} {that 3 + 3 is _} {without really proving _ or believing that 2 + 2 is 5} {that 3 + 3 is 7}.
Of the remainder of Postal’s examples, I find (405a—c) (his (116b), (115), (114a)) acceptable, though (405b—c) are decidedly clumsy. This leaves just (406) (Postal’s (116a)) as the solitary unacceptable example not predicted ungrammatical by my analysis. Perhaps the problem here is that when one encounters the ‘gap’ one does not have a sufficiently clear idea about its syntactic relationship to sequence A — one can’t work out what the first conjunct of the coordination is until one gets to sequence D. In contrast, in the similar example (405b) it is clear that the gap sequence B is complement of with. My explanation predicts that use of a correlative ought to improve (406), since it will be clear what the first conjunct is. And indeed, I find that, making the necessary allowances for the stylistic grotesqueness of RNR, (407a—b) become as acceptable as (405a—c). And it is possible to find pretty unobjectionable CSC violations, such as (408a—b).

(405) a. Edmund recognized (that the managers were too lazy) (without criticizing his predecessor and without admitting publicly) (that the managers were too lazy).
   b. Everyone who lives near Sandra or works with (the guy she is married to) (tends to end up disliking those who respect) (the guy she is married to).
   c. I argued for (the idea Shirley suggested) (and Frank argued against both your proposal and) (the idea Shirley suggested).

(406) ? Edmund recognized that the products were too expensive and (that the managers were too lazy) (without admitting publicly) (that the managers were too lazy).

(407) a. Edmund recognized both that the products were too expensive and (that the managers were too lazy) (without admitting publicly) (that the managers were too lazy).
   b. Edmund recognized both that the products were too expensive and that (the managers were too lazy) (without admitting publicly that) (the managers were too lazy).

(408) a. I’ll meet either Sophy and the sister of (their respective maternal grandmothers’ butlers’ secret love child) (or Edgar and the brother of) (their respective maternal grandmothers’ butlers’ secret love child).
   b. I’ll meet both Edgar and (their respective maternal grandmothers’ butlers’ secret love child) (or either Sophy or) (their respective maternal grandmothers’ butlers’ secret love child).
It remains possible of course that for some speakers RNR does obey the CSC, but a mere intuition of unacceptability of the relevant examples is not sufficient to show this. It must also be shown that considerations of style and processability are not responsible for such intuitions. Naturally this cuts both ways: before we can conclude that (405–408) are grammatical for all speakers we would have to show that it is considerations of style and processability that will account for their unacceptability for some speakers.

A second problem is (409), which Postal finds acceptable. He suggests that this involves rightward extraction by RNR followed by standard leftward extraction of the RNR-extracted material out of the parenthetical clause. This is incompatible with my analysis of RNR, which says, in effect, that material apparently RNR-extracted is in fact in situ.

(409) a. Sidney is, I suspected __ long before becoming certain __, living in Australia. [Postal's (148c)]

b. Traces are, some who have long held __ have finally succeeded in convincing me __, lexical in that sense. [Postal's (149d)]

However, his grounds for thinking that there must be RNR in these examples are not that strong. First, he claims that parasitic gaps licensed by standard extraction must be NPs and must be licensed by a nonparasitic NP gap. So if parenthetical extraction is standard extraction it should be able to strand prepositions, as in Sophy will, she convinced him (*of), be arriving soon. This I dealt with in §2.3.3, showing that such examples can involve regular extraction.

Second, he claims that genuine parasitic gaps (created by leftward extraction) cannot occupy positions incompatible with definite NPs.\textsuperscript{154} The complement of a verb like hold in (409b) can’t be a definite NP, so (409b) can’t involve leftward extraction. But the data Postal wishes to cover by this claim is possibly also covered by a requirement that parasitic gaps be objects or injects. The data in question is (410–412) (Postal’s (65), (67), (66)). In (410b) the parasitic gap is an xcomp, so would therefore be ungrammatical. Similarly, I see no obstacle to saying the complement of on in (411a–b) is an xcomp, making (411b) ungrammatical because the parasitic extractee isn’t an ob-

\textsuperscript{154} Judging by Postal’s examples, the does not always signal definiteness.
ject or inject.\textsuperscript{155} As for (412a–b), I don’t know what the function of the extracted/postverbal NP \textit{(guns/them/what)} is with respect to \textit{are/were} (\(-\ xcomp\) perhaps?), but there are reasons for thinking it isn’t an object (because in \textit{What is there here} it is \textit{there} that is object of \textit{is}).\textsuperscript{156}

(410) a. He named his camel Ernie/*/it.
    b. * What did he name his dog ___ after naming his Camel ___?

(411) a. They touched (*him on) it.
    b. Which of his arms did they have to immobilize ___ after accidentally touching (*him on) ___?

(412) a. There are guns/Them in the cabinet.
    b. * What did he look for ___ in the closet without knowing there were ___ on the table?

Third, Postal claims that genuine parasitic gaps cannot occur in inherently unpassivizable positions, but that that is what happens in (409a–b). The validity of Postal’s claim depends on when unpassivizability is inherent. (413a) is ungrammatical but (413b–d) aren’t.

(413) a. * Ophelia was pretended to to be mad.
    b. Who did Hamlet pretend to to be fond of?
    c. Who did Hamlet drive mad after pretending to to be mad?
    d. Who did Hamlet drive mad after pretending to to be fond of?

Postal’s claim rests on some very iffy data about prepositional passives plus the contrast between (414a–b) (where it is implied rather than asserted that (414a) is held to be okay).

(414) a. Who did she watch faint?
    b. * Who did they want to operate on after watching faint?

\textsuperscript{155} But my failure to see an obstacle should be seen in the light of my not having any views on the valency patterns of prepositions.

\textsuperscript{156} [I am now inclined to think that \textit{there} is inject and \textit{guns/them/what} is object, so I have not succeeded in countering entirely Postal’s argument, though I would still try to maintain the thrust of my case, which is that the constraint on parasitic gaps is a constraint on something other than definite NPs that allows through parasitic gaps in (409a–b).]

[235]
These are interesting data, and do suggest some unaccounted for and ill-understood connection between parasitic gaps and passivizability, but since I've already contested in §2.3.3 the claim that clausal complements won't passivize, I therefore contest that such a connection does not show that there are no parasitic gaps in (409a—b). In summary, then, none of Postal's arguments for why (409a—b) must involve RNR are really persuasive.

A third problem with my analysis of RNR is presented by (415a—c). They look as if RNR has applied, and then ellipsis.

(415) **RNR plus ellipsis.**

a. I would like to drink __, but don't __, tea that's been stewing for three hours in a mould-filled pot.

b. I like to drink __, but try not to __, tea that's been stewing for three hours in a mould-filled pot.

c. I like __, but try not to __, smoking cigarettes over dinner. [i.e. “try not to like”]

I am forced to analyse these as not involving RNR. One possibility is that but don't/but try not to is parenthetical or in some other way subordinated rather than coordinated. Another possibility is that something like heavy NP shift has occurred, giving a bracketing like that in (416). The need for similar bracketings to be possible is shown by nonelliptical examples like (417).

(416) [[I] [like [to drink]] but [try not to] [tea that's been stewing for three hours in a mould-filled pot]].

(417) [[I] need [to drink] [very badly] [tea that's been stewing for three hours in a mould-filled pot]].

Lastly, a fourth problem with my analysis of RNR (and perhaps symmetric coordination too) is that it says that some things that sound and feel like RNR aren't, and some things that don't sound and feel like RNR are. An example of the latter is (418). Perhaps here we could argue that the propinquity of the two word-sequences that corre-
spond to the same vocable-sequence makes (418) an unusually innocuous example of RNR.

(418) \textit{She wants him to core and to peel the apples.}
She wants him [[to [core \textit{the apples}]] and [to [peel the apples]]].

Examples of things which sound and feel like RNR but are predicted not to be are (419a–b). In (419a), \textit{get} is not node raised but rather is simply the ward of \textit{and} and s-dependent of the coordinatees \textit{should} and \textit{would}. In (419b), \textit{the} is ward of \textit{and} and s-dependent of the coordinatees \textit{core} and \textit{peel}. Why (419a–b) should sound so much more like RNR than (420a–b) do is unexplained.

(419) a. [[[She should] and [[he will] [get up first thing in the morning]]].
   b. She’ll [[[skillfully] core] and [[carefully] peel] [the apples]].

(420) a. She and he will get up first thing in the morning.
   b. She’ll skillfully core the apples and peel the oranges.

What is even more problematic is that (421a) needn’t be analysed as involving RNR, yet the grammaticality of the verb agreement in (421b–e) indicates that (421a) must involve RNR, as in (422).

(421) a. [[[She] [quickly] and [he] [slowly]] gets richer.
   b. * She quickly and he slowly get richer.
   c. * She quickly and they slowly get richer.
   d. * She and he slowly gets richer.
   e. She and he slowly get richer.

(422) [[[She quickly \textit{gets richer}] and [he slowly gets richer]].

(422) is not ruled out, and nor are RNR versions of (419a–b), shown in (423a–b). The problem is to prevent (421), and perhaps (419a–b) into the bargain, from being analysable as coordination without RNR. Presumably there are further constraints on wards that are yet to be worked out.
(423) a. [[[She] should [get up first thing in the morning]]] and [[[he] will [get up first thing in the morning]]].
   b. She'll [[[skillfully] core [the apples]]] and [[[carefully] peel [the apples]]].

5.3.3 Morphological coordination

The problem with morphological coordination, such as (424a–b), is that it appears that the coordinatees of and are not words but prefixes. For reasons explained at length in §5.5, it ought to be impossible for morphological objects to participate in syntactic structure, so morphological coordination might seem to present a worrying counterexample.

(424) a. *pre- and post-war*
   b. *centrifugal and -petal forces*

I see three solutions. First, (424a–b) may be the result of lexeme-creation: *[pre- and post-war]* and *[centrifugal and -petal]* would be the morphological-phonological form associated with instances of the lexemes PRe-AND-POSt-WaR and CeNTRIfuGAL-AND-PETAL respectively. As WG has so far had little to say about lexeme-creation, it is hard to evaluate that proposal. Second, the syntactic structure corresponding to the morphological-phonological structures in (424a–b) could be those in (425a–b).

(425) a. *pre-war and post-war*
   b. *centrifugal and centripetal forces*

In this case, we would posit a rule that allows a morpheme to be ‘deleted’, subject to the conditions that there be sufficient syntactic and semantic parallelism and the undeleted morphological remnant have the phonological properties sufficient for it to be an independent vocable.

This latter solution might also apply to RNR. (426a–b) would involve RNR plus ‘morpheme deletion’, and (426c) would just involve morpheme deletion. (Subscript italic denotes deleted morphemes and subscript roman denotes words not simultaneous with their associated vocables.)
(426) a. \{She had ante-natal \} \{depression\} \{and he had post-natal\} \{depression\}.
b. \{He was intro-duced\} \{by Euphrosyne\} \{and she was seduced\} \{by Euphrosyne\}.
c. This plane is sub-sonic but that plane is supersonic.

Alternatively, it may be that my earlier account of RNR was wrong: rather than involving one sequence of vocables associated with two sequences of words, it could be that the vocables that should be associated with the first sequence of words are wholly deleted, as in the examples in §5.3.1, or partly deleted, as in (426), which would be reanalysed as (427a–c).

(427) a. She had ante-natal depression and he had post-natal depression.
b. He was intro-duced by Euphrosyne and she was seduced by Euphrosyne.
c. This plane is sub-sonic but that plane is supersonic.

This reanalysis of RNR would deal with (424a–b), but has the drawback that a way has to be found to formulate the rule that the deletion happens to a morphological-phonological string when it is parallel to another string that is in all respects similar. The earlier RNR analysis could sidestep this problem because there was no deletion involved. So perhaps the virtues of both analyses can be combined in the following way. Assume the analysis of RNR given in §5.3.1, but in addition assume that it is possible for some rather than all of a vocable to be associated with more than one word, and that if a part of a vocable is associated with more than one word then it is, exceptionally, permitted to not concatenate in the normal way, so long as any morphemes not properly concatenated are simultaneous with the word that the vocable that the morphemes are part of is associated with. Applied to (426a–c) this gives the following patterns of correspondence between syntax and morphology-phonology.

(428) a. She had ante- and he had post-natal depression.

\[
\text{She had antenatal depression and he had postnatal depression.}
\]
b. He was intro- and she was se-duced by Euphrosyne.

\[
\text{He was introduced by Euphrosyne and she was seduced by Euphrosyne.}
\]
c. This plane is sub- 

but that plane is super- sonic.

This plane is subsonic but that plane is supersonic.

(424a – b) would then be straightforward ‘node raising’, as in (429a – b). This is the third and most satisfactory solution to (424a – b).

(429) a. pre- and post- war

prewar and postwar

b. centri- fugal and petal forces

centrifugal and centripetal forces

5.4 Further evidence for coenunciation

As advertised earlier, this section presents three recalcitrant problems for which coenunciation offers an appealing solution. The success of these solutions lends support to a model of grammar that allows for coenunciation.

5.4.1 Interrogatives

The analysis of interrogative and relative clauses presented in Chapter 2, which essentially follows that of Hudson (1990) gives the partial dependency structure in (430).

(430) of

of

of

of

of

of

Which students have failed is unclear.

Hudson (1990, citing Bob Borsley p.c.) mentions this sentence as a problem for the kind of analysis shown for it, in which which is subject of both have and is. The problem is that have shows plural agreement but is shows singular agreement. There is the odd construction in which plural nouns are subject of singular verbs, as in (431a – b), and in which singular nouns are subject of plural verbs, as in (431c), but in the case of the former, the verb must be an equative be, while in the case of the latter, a singular
verb is always a possible variant, (431d). In (430), neither verb is an equative be, and is doesn’t alternate with are. So even the various minor subregularities of subject-verb agreement can’t accommodate the problem of (430).

(431) a. Your trousers is the problem.
    b. Your trousers is what’s bothering her.
    c. Parliament have debated the matter.
    d. Parliament has debated the matter.

A solution put forward by Dick Hudson in recent unpublished work (1995d) is that at least some interrogative pronouns, including the one in (430), actually involve co-enunciation of an interrogative pronoun and a relative pronoun, such that the relative pronoun is a complement of the interrogative.157 (432) diagrams this; the relative pronoun is shown as belonging to lexeme ‘WHICH’, and the interrogative is shown as belonging to lexeme ‘WH’.158

\[
\begin{array}{c}
| s \relmid o f \relmid s \relmid o f \relmid s \relmid o f \relmid of \relmid x | \\
| of \ c \relmid of \relmid of \ c | \\
| \text{WH WHICH STUDENTS HAVE BE UNCLEAR} | \\
| \text{Which students have failed is unclear.} |
\end{array}
\]

In this structure, the WHICH agrees in number with its complement and is therefore plural; hence have shows plural agreement with its plural subject. The WH is singular, so is shows singular agreement with its subject.

For different reasons, McCawley (1989) advocates an analysis of relative pronouns that is in some ways similar to this one. Following his arguments in McCawley (1981)

157 The only viable alternative to this suggestion that I’ve been able to think of is simply to stipulate various additional exceptions to the rules of subject-verb agreement.

158 To this can be added further evidence for a coenunciational analysis of interrogatives and of wh-ever words. As mentioned in §4.2.3, a kind of pied piping is possible in free relatives and in interrogative clauses where the interrogative word is within the subject NP, so long as the relative or interrogative word is at the extreme left edge of the clause. Examples showing this (repeated from §4.2.3) are (i–iv).

(i) Whoever’s manners are reprehensible is not to be allowed entrance to the club.
(ii) * The manners of whoever are reprehensible is not to be allowed entrance to the club.
(iii) (I wonder) whose mother (‘s personal guru) arrived late
(iv) * (I wonder) (the personal guru of) the mother of whom arrived late
he claims that restrictive relative clauses modify an N' rather than an NP. Since compound pronouns, e.g. someone, can be modified by a restrictive relative clause, McCawley concludes that they correspond to more than one node in syntactic structure, as in (433) (adapted from McCawley's (27)).

(433) \[([\text{some-}]_o \ [([\text{-one}]_n \ [\text{that I admire}]_s]_n]_N)_{NP} \]

McCawley appears to intend this analysis to be contingent on the morphological complexity of the pronoun, and comments that he knows of no grounds for choosing between an analysis like (433) (where sub-word-level morphemes are terminal syntactic

These facts are explicable on the assumption that free relative -ever words and at least within the subject NP interrogatives in fact comprise two terminal nodes, which, because they coenunc with each other, must be adjacent. The permissible structures are shown in (v−vii), where the constituency is as indicated by the bracketing, where 'WHO' is complement of 'EVER' and 'WH', and where it is 'WHO' rather than 'EVER'/'WH' that is involved in pied piping. (I leave open the question of whether it is 'WHO' or 'EVER'/'WH' that is subject of is in (v), complement of wonder in (vi) and subject of arrived in (vii); if it is 'WHO', then 'EVER'/'WH' would be supervener of 'WHO' and would work like subjunctions (see §4.6). (NB In (v−x) only relevant coenunciations are indicated.)

(v) Whover's manners are ... is ...
[EVER [WHO'S MANNERS ARE ...]] IS ...

(vi) i wonder whose mother's personal guru arrived late.
I WONDER [WH [WHO'S MOTHER'S PERSONAL GURU ARRIVED LATE]]

(vii) Whose mother's personal guru arrived late?
[WH [WHO'S MOTHER'S PERSONAL GURU]] ARRIVED LATE

The ungrammatical counterparts of (v−vii) are shown in (viii−x). They are ungrammatical because 'EVER'/'WH' is not adjacent to 'WH', so they cannot fulfill the requirement that they must be co-enunciate with each other.

(viii) * The manners of whoever are ... is ...
[EVER [THE MANNERS OF WHO ARE ...]] IS ...

(ix) * I wonder the personal guru of the mother of whom arrived late.
I WONDER [WH [THE PERSONAL GURU OF THE MOTHER OF WHO ARRIVED LATE]]

(x) * The personal guru of the mother of whom arrived late?
[WH [THE PERSONAL GURU OF THE MOTHER OF WHO]] ARRIVED LATE

Note that although (xi) is grammatical, this is a quizshow question. In my lect, whom is the enunciation of an instance of the lexeme Who if and only if the who is complement of a pied-piped preposition.

(xi) The personal guru of the mother of who arrived late?
Translated into WG terms, that would mean that they must modify a common noun. But that requirement would undergenerate. It remains unclear what exactly the constraints on restrictive relative clauses are. What is more persuasive evidence in favour of McCawley's analysis of compound pronouns is that the head of a bare relative clause can only be a relative or interrogative pronoun or a common noun, and if it's a common noun then it must be extractee, rather than subject, of the verb in the relative clause. In this respect, compound pronouns behave exactly like common nouns: She met someone you know, *She met someone knows you.
nodes) and (434a) below (where the compound pronoun constitutes two terminal nodes). There are grounds for choosing between them, and they would lead to the rejection of (433): the non-compound pronouns *aught/owt* and *nought/nowt* are in all syntactic respects identical to *anything* and *nothing*.

Since interrogative pronouns can be modified by restrictive relative clauses, as in *Who that you admire should we invite?*, it follows for McCawley that they too should receive an analysis like (433). He suggests two possible analyses, (435a—b).\(^\text{160}\)

\[(435)\]
\[
\begin{array}{l}
\text{a. } \text{NP} \\
\quad \text{Det} \\
\quad \text{N} \\
\quad \text{N} \\
\quad \text{S} \\
\quad \text{who} \\
\quad \text{that you know} \\
\end{array}
\]
\[
\begin{array}{l}
\text{b. } \text{NP} \\
\quad \text{Det} \\
\quad \text{N} \\
\quad \text{N} \\
\quad \text{S} \\
\quad \text{wh-some one} \\
\quad \text{that you know} \\
\quad \text{who} \\
\quad \text{that you know} \\
\end{array}
\]

In (435a), *who* has two mothers. (435b) is an analysis in the style of Autolexical syntax (Sadock 1985, 1991). Both analyses are in the spirit of my coenunciation model, but neither is strictly compatible with it. (435b) is very close to a coenunciation analysis, but gives morphological structure syntactic properties (*who* is within an NP). (435a) is incompatible with WG because WG doesn’t allow multiple mothers. Nonetheless, the translation into a coenunciative analysis is easy to make.

**5.4.2 'Underlying' demoted subjects of unactives**

There are various grounds for thinking that the demoted ‘underlying’ subject of unactive verbs (i.e. passives and mediopassives) is lurking unseen somewhere in the grammatical structure of the sentence. For example, in (436a) the participant who is naked

\[\text{\textsuperscript{160} I’ve had to rectilinearize the standard tree diagrams, due to the constraints of my word processing software.}\]
can be identified with Sophy, who is thought about, or with the person who thinks about her, and in (436b–c), the participant who is willing can be identified with Sophy or with her kisser.

(436) a. Thought about while naked, Sophy felt no qualms about abandoning herself to her passions.
   b. Willingly kissed, Sophy felt no qualms about abandoning herself to her passions.
   c. Kissed willingly, Sophy felt no qualms about abandoning herself to her passions.

Intuitively, rules identify the semantic ‘subject’ (and indeed possibly the syntactic subject) of naked and willingly with either the demoted subject or the promoted subject of the verb (thought, kissed). The problem, under a monostratal approach, is that there is no obvious way to capture the notion ‘demoted subject’. In the analysis of Rosta (1995a), adapted from that of Hudson (1990), there is a semantic role, called ‘Er’, which is always associated with subjects of active verbs, or at least of active transitive verbs, irrespective of what other semantic role the active subject bears. Rules identify the semantic or syntactic subject of willingly and its ilk with either the referent of the subject of the verb (the kissee, in these examples) or with the er of the meaning of the verb (the kisser).

The drawbacks of this analysis are firstly that in rules for the grammar of controlled adjuncts the disjunction “referent of the subject of W or er of the sense of W” keeps on cropping up. This suggests that some generalization about a common property they share is being missed. Secondly, the Er role amounts to a kind of betrayal of semantics, in that there is no claim that all active subjects have anything in common with each other semantically; rather, ers serve the essentially syntactic purpose of capturing the notion of underlying subject, in order to account for controlled adjuncts, for the meaning of the passive by-phrase, and so on. I therefore conclude that an analysis which offers us a syntactic representation of underlying, demoted subjects is extremely desirable. Coenunciation affords such an opportunity.

---

161 This phenomenon was observed by Lakoff (1970) (cited in Lakoff 1977).
Before we look at passives, consider the analysis, which I here have space only to sketch with the utmost brevity, of ‘esmerges’ — words whose extractee and whose subject merge. Esmerges are typically either an infinitival for or to. Examples are given in (437a–d), (437c–d) showing dependency structures for infinitival for or to respectively (non-s-dependents are in italics, and the dependents of the esmerge are in boldface and capitals).

(437) a. The book is (for her) to read several chapters of.
    b. He bought it (for her) to read several chapters of.

c. The book is easy for her to read several chapters of.

d. The book is easy to read several chapters of without taking a break.

When a to is an esmerge, as in (437d), the default merger of its object and its subject is overridden: the subject of this to is the (book) and its object is absent. The default mer-
ger of its object and the subject of its xcomp is not overridden, so in this example the object of to is merged with the subject of read. Exceptionally, the esmerge's subject merges with its extractee. Loosely speaking, in (437d) the (book) is 'promoted' subject of to read (taken as a unit), while the 'demoted' subject is not expressed, unlike in (437c), where the subject of to read is not demoted and is expressed by her. This split in subjects is possible only because to read is two words, not a single word. Thus, the subject of to merges with the object of easy (of is and bought in (437a–b) respectively), as befits an xcomp, or, more generally, an xfunc, while the unexpressed subject of read, which is merged with the object of to, can merge with the subject of the controlled adjunct without and its xcomp taking, with the result that the same participant is the reader of the book and the taker of a break.

If a passive is in fact a coenunciation of two words with the same enunciation, we can use essentially the same method as is used with esmerges for getting at demoted subjects. As mentioned later in §6.4–5, I take a passive to be a coenunciation of a variety of auxiliary have that has the same enunciation as its perfect xcomp. The auxiliary have in question is an 'unactive'. The properties of unactive auxes are as follows. As with esmerges, when the auxiliary is an unactive the default merger of its object and its subject is overridden. Again, as with esmerges, the default merger of its object and the subject of its xcomp is not overridden. The distinctive property of unactives is that their subject merges with, by default, the inject of its xcomp. For regular passives that don't involve preposition stranding, it is the inject that is promoted. Thus, most passives have the structure shown in (438a–b).

\[
\begin{array}{cccccc}
\text{of} & \text{of} & \text{x} & \text{of} & \text{of} & \circ \\
\text{i} & \text{of} & \text{of} & \text{of} & \text{of} & \text{of} \\
\text{of} & \text{of} & \text{of} & \text{of} & \text{of} & \text{of} \\
\text{HER} & \text{HAVE} & \text{ADORÉ} & \text{unactive} & \text{perfect} & 8
\end{array}
\]

(438) a. Her adored by everyone, nobody criticizes her work.

---

162 'Xcomp' is a terminological abbreviation for "xfunc of W that is also a complement of W". In She took it off to cover the baby with both off and to are xfuncs, but only off is an xcomp (i.e. only off is a complement). In He bought it for her to read several chapters of, for is xfunc but not complement of bought, and in He bought it to read several chapters of to is xfunc but not complement of bought.

163 Rosta (1995a) explains how unexpressed dependents successfully serve the purpose of merging semantic 'arguments' of different 'predicates'.

164 There is insufficient space to elaborate on this analysis here.
The referent of by is equated with the referent of its complement, everyone. Either (439a–b) or (440a–b) are responsible for the referent of everyone being understood as the adorer.

(439) a. Every by passive is a dependent of an unactive HAVE.
    b. For every by passive, X, the referent of X is the referent of the object of the head of X.

(440) a. Every by passive is a dependent of the xcomp of an unactive HAVE.
    b. For every by passive, X, the referent of X is the referent of the subject of the head of X.

As for kissed willingly, there are various ways of dealing with the ambiguity. First, the alternative structures could be as in (441a–b), with the indicated interpretations determined by (441c). Second, the alternative structures could be as in (442a–b), with the indicated interpretations determined by (442c). Third, the alternative structures could be as in (443a–b). These rely on the coenunciative analysis of adverbs, and on the ly being subject-raising. The advantage of such an analysis is that no rule analogous to (441c) or (442c) is required.165

(441) a. Sophy was kissed willingly by Edgar

[= Sophy was willing]

165 But there are also drawbacks to the analysis, since if it were the case that adverbs (lys) can't have subjects, then their failure to ever function as xcomps or as predicative adjuncts of common nouns or compound pronouns would perhaps be thereby explained at least in part.
HAVE KISS WILLINGLY

b. Sophy was kissed willingly by Edgar

[= Edgar was willing]

c. The referent of the sense of adjuncts like willingly is the referent of the subject of the head of the adjunct.

HAVE KISS WILLINGLY

b. Sophy was kissed willingly by Edgar

[= Sophy was willing]

c. The subject of adjuncts like willingly is the referent of the subject of the head of the adjunct.

HAVE KISS WILLINGLY

b. Sophy was kissed willingly by Edgar

[= Sophy was willing]

[248]
Some but not all speakers find this ambiguity of the sort that we get with *willingly* and controlled adjuncts not only with passives, but also with mediopassives.\(^\text{166}\) For those speakers who don't find any ambiguity, it is always the promoted subject that is semantically linked to the adjuncts. This, among various other facts, some mentioned in Rosta (1995a), suggest that two syntactically somewhat different, but semantically rather similar, constructions go by the name of 'mediopassive' or 'middle' in the literature.\(^\text{167}\) For speakers who don't find the ambiguity, all mediopassives are merely osmerges — words whose objects and subjects merge; they are 'unaccusativized', one might say. For other speakers, however, mediopassives can also have a syntactic structure exactly analogous to passives; for example, these speakers accept prepositional mediopassives like *Paper cups dispose of easily*, while the former group of speakers, all of whose mediopassives are osmerges, will accept only *Paper cups discard easily*. For the latter group of speakers the difference between passives and mediopassives are largely only semantic.

Now, suppose that all verbs are infinitival xcomps of something like an auxiliary or infinitival *to*, possibly with coenunciation involved. (The evidence of double depictives in §5.5 leads to the same conclusion.) If our analysis of passives as involving an unactive auxiliary have is correct, then since in some lects mediopassives behave syntactically like passives, it follows that all mediopassives will involve a verb that is xcomp of an unactive auxiliary-like word, as, for example, in (444). In this example, the unactive and its EXORCIZE xcomp share the same enunciation, *exorcize*. The empty, understood subject of the EXORCIZE is merged with the object of the unactive.

\(^\text{166}\) See Rosta (1995a: 140 n9).

\(^\text{167}\) See Rosta (1995a) for a survey of names for the construction.
Just as any of the traditionally recognized forms of verbs, including bare infinitive xcomps of verbs, can have double depictives (see §5.5), so any of the same range of verb forms can be mediopassive. Both double depictives and mediopasses therefore support the conclusion that all verbs are xcomps of a possibly coenunciate aux-like word.

5.4.3 Gerunds

The problem with gerunds, as in *Him kissing her caused no end of a fuss*, is that as dependents they behave like nouns, while as heads they behave like verbs. Hudson (1990) offers an ingenious analysis that goes a long way to solving the problem, and foreshadows the use of coenunciation, with which his analysis is (arguably but probably) compatible. He treats the *-ing* as a nominal clitic that takes an infinitive complement, and gives (445a–b) the analyses shown. (In the light of the present coenunciation model, the nominal *ing* and its infinitive complement would instead be analysed as associated with the same enunciation.)

This analysis solves the gerund problem because it is the nominal *ing* that acts as a dependent of words outside the ‘gerund phrase’, while it is the *ing*’s complement, the infinitive, that functions as a head within the gerund phrase. We return to gerunds in §§6.3–4.
5.5 The nature of terminal nodes in s-dependency structure

Dick has pointed out to me (446), which I present with the partial dependency structure he gives it.

\[
\text{of} \quad \text{of} \\
\text{of} \quad \text{of} \\
\text{COMBIEN IL DE LE MANGE} \\
\text{Combien il en mange?}
\]

(446)

If we continue to assume that terminal nodes in s-dependency structure are words, then the dependency structure in (446) can give rise to either of the s-dependency structures shown in (447a–b) by stemma.

\[
\text{COMBIEN IL DE LE MANGE} \\
\text{Combien il en mange?}
\]

(447) a.

\[
\text{COMBIEN IL DE LE MANGE} \\
\text{Combien il en mange?}
\]

b.

As (447a–b) show, the dependency structure in (446) inevitably gives rise to tangling.

But suppose terminal nodes in s-dependency structure are not words but vocables. Dependency structure could continue to constrain s-dependency structure by means of a rule like (448).

\[
\forall a \forall b, a \text{ is s-dependent of } b: \exists c \exists d, a \text{ is enunciation of } c, b \text{ is enunciation of } d, c \text{ is dependent of } d
\]

(448)

This allows (446) to be associated with either of the s-dependency structures shown in (449a–b) both by stemma and by bracketing.

\[
\text{COMBIEN IL DE LE MANGE} \\
\text{Combien il en mange?}
\]

(449) a.

\[
\text{COMBIEN IL DE LE MANGE} \\
\text{Combien il en mange?}
\]

b.
The s-dependencies in (449b) are licensed by either of the dependency structures shown in (450a–b).

(450) a.  
```
[Combien [Il [en mange]]?

COMBIEN IL DE LE MANGE

of e
of k

b.  
```
```
[Combien [Il [en mange]]?

COMBIEN IL DE LE MANGE

of e
of c

The reader should be warned that I shall be rejecting this solution, but before I explain why, let's consider its merits and implications further.

From this solution of the tangling in (447a–b) emerges a view of a merger of s-dependency structure and phonological/morphological (P-M) structure; roughly, s-dependency structure would be above the vocable level, and P-M structure would be below the vocable level. Such a merger has a number of attractions. For example, sequentiality — linear precedence — is a property of P-M structure and of s-dependency structure, but not of dependency structure. And both P-M structure and s-dependency structure are constituency-based, whereas dependency structure is not. These patterns of similarity and dissimilarity emerge more naturally if vocable structure and s-dependency structure are the same.

In this case, the function of words and the dependency structure they participate in is to constrain P-M/s-dependency structure and to build semantic structure. Coenunciation will serve principally to explain semantic structure and lexical blocking. Since there is no reason to think that words are sequenced, there is no reason to think words are inherently or necessarily located at a specific time; thus words are more like signs — links between sounds and meanings — than like actions. For the semantics of deixis, the WG analysis of which has relied on words being actions, we must either make the senses of deictic words involve not the words themselves but their enunciations, as in (451a), or stipulate, as in (451b–c), that words share certain properties, such as their time, with their enunciations.

---

168 At least, constituency plays at least some role in morphological and phonological structures.
(451) a. If \( W \) is an instance of \( \text{Now} \) then the referent of \( W \) is the time of the enunciation of \( W \).

b. If \( W \) is an instance of \( \text{Word} \) then the time of \( W \) is the time of the enunciation of \( W \).

c. If \( W \) is an instance of \( \text{Now} \) then the referent of \( W \) is the time of \( W \).

Though the tentative conclusions reached so far in this section are appealing, we must now consider a variety of data the explanations for which rely on assuming that s-dependencies hold between words rather than vocables, and that therefore those tentative conclusions are incorrect. I present the data in increasing order of dubitability.

\textbf{Possessive 's.} (452a–d) present alternative analyses of two ungrammatical sentences. In the analysis shown in (452a–b), s-dependencies hold between vocables; the resulting s-dependencies do not tangle, and hence the ungrammaticality remains unexplained. In the analysis shown in (452c–d), s-dependencies hold between words; the resulting s-dependencies tangle, and thereby the ungrammaticality is explained.

\begin{itemize}
  \item \textbf{a.} * [Singe [[[the] king's [of [Spain]] [beard]]]].
    \begin{align*}
    \text{SINGE} & \quad \text{THE K\textsc{ing}'s} \quad \text{OF SPAIN} \quad \text{BEARD} \\
    \text{.} \quad \text{.} \quad \text{of C} \quad \text{of C} \\
    \text{.} \quad \text{of C} \quad \text{.} \\
    \text{.} \quad \text{of C} \quad \text{.} \\
    \text{of C} \quad \text{.} \quad \text{.}
    \end{align*}

  \item \textbf{b.} * [Singe [kings' [of [Spain]] [beard]]].
    \begin{align*}
    \text{SINGE} & \quad \text{KING'S} \quad \text{OF SPAIN} \quad \text{BEARD} \\
    \text{.} \quad \text{.} \quad \text{of C} \quad \text{.} \\
    \text{.} \quad \text{of C} \quad \text{.} \\
    \text{.} \quad \text{of C} \quad \text{.} \\
    \text{of C} \quad \text{.} \quad \text{.}
    \end{align*}

  \item \textbf{c.} * [Singe [[[the] king's [of [Spain]] [beard]]]].
    \begin{align*}
    \text{SINGE} & \quad \text{THE K\textsc{ing}'s} \quad \text{OF SPAIN} \quad \text{BEARD} \\
    \text{.} \quad \text{.} \quad \text{of C} \quad \text{of C} \\
    \text{.} \quad \text{of C} \quad \text{.} \\
    \text{.} \quad \text{of C} \quad \text{.} \\
    \text{of C} \quad \text{.} \quad \text{.}
    \end{align*}
\end{itemize}
-ly adverbs. Many adjectives subcategorize for a complement, e.g. *happy in (453a). In most cases, when an -ly adverb is formed from the adjective, the adverb does not have a complement — (453b).

(453) a. She was happy that it was sunny.
   b. * She smiled happily that it was sunny.

We should not necessarily expect the adverb to lack the complementation of its adjective counterpart, since changing word class of a word does not necessarily alter its complementation. as (454a–d) show.

(454) a. She was reluctant to leave.
   b. her reluctance to leave
   c. She was unwilling for him to leave.
   d. her unwillingness for him to leave

Alongside (453a–b), we have cases where the adverb does keep the complementation of its adjective counterpart, as (455a–d) show.

(455) a. Her writing is different from his.
   b. Her writing is similar to his.
   c. She writes differently from him.
   d. She writes similarly to him.

The ungrammaticality of (453b) is easily accounted for by a stipulation that by default adverbs don’t have complements. This stipulation could be overridden for similarly and differently. It is not clear how these two adverbs come to share the complementation of
their adjective counterparts, but such a device is anyway presumably needed for the adjective/noun alternations in (454a–d).

There is, however, an alternative and probably more explanatory way of dealing with -ly adverbs. Suppose that happily is actually a coenunciation of an instance of HAPPY and an instance of the lexeme LY. The instance of HAPPY is an adjective and a complement of the instance of LY, which is an adverb. And suppose that all -ly adverbs are coenunciates in this way. Now, if we assume (i) that the LY exceptionally follows its complement, and (ii) that s-dependencies hold between words rather than vocables, then we predict the failure of happily and its like to have complements, since if they had complements tangling would arise. (456) diagrams the relevant bit of (453b).

Contrary to what might be expected, the grammaticality of (455c–d) indirectly supports this analysis. As we see from (457a–b), the complement of DIFFERENTS and SIMILARS extraposes across the s-head of the DIFFERENT or SIMILAR.

We can offer analogous structures for (455c–d), shown in (458a–b).
A remaining problem is *immediately*, which actually has a complement of a type its adjective counterpart doesn’t; cf. (459a—b).

(459) a. She entered immediately the door opened.

b. * Her entrance was immediate the door opened.

At present I’ve no suggestion for how (459a) should be analysed.

It is tempting to go one step further, and treat all adverbs as a coenunciation of an adjective and a *ly*, with certain adjectives, such as the one coenounced by *seldom*, belonging to a class of adjectives that (i) are always a complement of a *ly*, and (ii) have an enunciation that doesn’t end in *-ly*. The class of adverbs could then be equated with the lexeme *ly*. We would then have an explanation for why exactly the same range of degree modifiers modify apparent adverbs as modify adjectives (*very/too/how/so/this/that/more/quite happy/happily*): they modify not the adverb but the coenounced adjective. This would be in contrast to, say, nouns whose morphology consists of the stem of an adjective and the suffix *-ness*: these nouns have the modifiers that other nouns can have, not the modifiers adjectives can have. We would therefore not want to say that such nouns are a coenunciation of an adjective and a *‘ness’*.

In a more conservative analysis where adverbs aren’t coenunciate, we could put adjectives and adverbs in some more general category, such as ‘ad-word’, and then state the permitted range of degree modifiers in terms of ‘ad-words’. But it should be noted that this is the only characteristic adverbs and adjectives share: as dependents, they behave completely differently, so ‘ad-word’ would be a rather ad hoc category.
Double depletives. Depictive adjuncts can be 'object-related', as in (460a), or 'subject-related', as in (460b). (‘dP’ stands for ‘depictive’.)

\[
\text{He ate the meat raw.}
\]

\[
\text{He ate the meat naked.}
\]

A verb can have both an object-related depictive and a subject-related depictive, but in such a case the subject-related depictive cannot be between the verb and the object-related depictive. (461a–b) show this.

\[
\text{[He ate [the [meat]] [raw] [naked]].}
\]

\[
\text{[He ate [the [meat]] [naked] [raw]].}
\]

Intuitively, (461b) looks like it is bad because of some tangling somewhere, and indeed the subject dependencies do tangle, as (462a) shows. But, first, the diagram for (461b) shows it to be tangle-free. And, second, as Dick Hudson has pointed out to me, the subject dependencies don’t tangle in (462b), which is nonetheless ungrammatical.170

\[
\text{He ate the meat naked raw.}
\]

\[
\text{Which meat did he eat naked raw?}
\]

---

170 This put paid to an earlier attempt of mine to explain the double depictive problem in terms of processing. The idea was that potential subjects are placed on a stack, and when a word requiring a subject turns up, the top word on the stack can either be discarded and removed from the stack, or be made the subject and left on the stack.
The ungrammaticality of (461b) and (462b) is explicable if we assume the following:

(A) s-dependencies hold between words;
(B) every verb (not counting modals) is an infinitive that is xcomp of either (i) a TO, (ii) something like a TO, but which the infinitive coenounces with, or, maybe, (iii) a modal;
(C) exactly as with xcomps, the subject of a depictive of W is the object of W.\footnote{More neatly, we can define a dependency type ‘xfunc’, such that the subject of an xfunc of W is the object of W. Then we can define ‘xcomp’ as a combination of Xfunc and Complement: X is xcomp of Y iff X is xfunc of Y and X is complement of Y. As for Depictive, X is xfunc of Y and adjunct of Y if (but not only if) X is depictive of Y. On the whole I leave for future research the issue of how to properly formalize cross-classificatory definitions of dependency types, and of which more specific dependency types should be cross-classified in terms of which more general dependency types.}

To assume (C) is rather attractive, since it gets us close to a uniform analysis of ‘subject sharing’ (‘control’): if X gets its subject by virtue of being a dependent of Y, then the subject of X is object of Y.\footnote{There is not enough space in the thesis to cover this sort of stuff in greater detail.} The pros and cons of assuming (B) are taken up later in this chapter. Given assumptions (A—C), we get the structures shown in (463a—d); the s-dependencies of (463c—d) inevitably tangle and hence lead to ungrammaticality.

\footnote{More neatly, we can define a dependency type ‘xfunc’, such that the subject of an xfunc of W is the object of W. Then we can define ‘xcomp’ as a combination of Xfunc and Complement: X is xcomp of Y iff X is xfunc of Y and X is complement of Y. As for Depictive, X is xfunc of Y and adjunct of Y if (but not only if) X is depictive of Y. On the whole I leave for future research the issue of how to properly formalize cross-classificatory definitions of dependency types, and of which more specific dependency types should be cross-classified in terms of which more general dependency types.}

\footnote{There is not enough space in the thesis to cover this sort of stuff in greater detail.}
On balance, I think the evidence of possessive 's, of adverbs and of double depictives leads us to favour an analysis in which words rather than their enunciations are terminal nodes in s-dependency structure. We ought therefore to return to *Combien il en mange*, which was what led us initially to consider the contrary view. I see three solut-
ions. The first is that unlike in English, in French s-dependencies are between vocables. That may be in fact be right, but since the French data would then have no bearing on English, and our object of study here is English, I shan’t consider this solution further.

The second solution is to allow words to be simultaneous. The effect of this is to locate words in a two-dimensional rather than one-dimensional space. Under such circumstances, bracketability — or boxability — becomes much easier. In (464a—b) words are horizontally sequenced according to position in time; they are not vertically sequenced. Each box represents a phrase. As the diagrams show, there is no discontinuity. In fact, without further constraints, simultaneous words can never give rise to discontinuity.

This notation makes *(Combien il en mange) grandes avec cuillères boxable, as in (465), so obviously some extra constraints of some sort are required.

The most natural constraint that would rule out (465) would be to require every s-head to be adjacent to at least one of its s-dependents. But this would also rule out (464b), and wrongly predict Beaucoup, il en mange to be ungrammatical, since mange is not adjacent to any of its s-dependents. Given that the most natural constraint makes wrong predictions, it seems that this third solution of allowing simultaneous words is probably not worth pursuing.

[260]
Dick Hudson (p.c.) prefers a third solution, which is that “word structure rules take priority over sentence structure rules,” a view shared by Autolexical Syntax. But we have just seen that this is wrong: it would make wrong predictions about possessive 's, about -ly adverbs and about double depictives. However, it is interesting to note that in each of these constructions the tangling s-dependency is between a coenounced word and its s-dependent, never between a coenounced word and its s-head. For convenience, the relevant structures are repeated in (466a–c).

(466)

a. *Singe kings' of England beards

b. *smile happily that she was

c. *eat meat naked raw

Now, contrast this with the French clitic example, (467). Here, there is tangling between the coenounced de and its s-head the BEAUCOUP.

(467) DE LE AVOIR MANGER BEAUCOUP

Sophy en a mangé beaucoup

The contrast between the ungrammatical and tangling (466a–c) and the grammatical but apparently tangling (467) looks as if it might be significant. Could it be that word structure rules take priority over sentence structure rules in that they can relieve a coenounced word of the need for an s-head? Let’s try an alternative rule to capture this: (468).

(468) If X is a dependent of Y, and Y is a subordinate of Z, and X and Z share their enunciation, then X is an s-dependent of Z.

Applying this to (467) we find that the de is a dependent of the BEAUCOUP, and the BEAUCOUP is a dependent of the AVOIR. Therefore the de is an s-dependent of the AVOIR.

Cf. Hudson (1990: §11.9) on gerunds.
An objection to (468) is that it is an exception to the otherwise exceptionless rule that if \(X\) is an \(s\)-dependent of \(Y\) then \(X\) is a dependent of \(Y\). Further, (468) is anyway tantamount to (469); the only difference is that (469) is not open to the objection that it introduces an exception to the correspondence between dependency and \(s\)-dependency.

(469) \( \) If \(X\) is a dependent of \(Y\), and \(Y\) is a subordinate of \(Z\), and \(X\) and \(Z\) share their enunciation, then \(X\) is a dependent of \(Z\).

(469), then, is the fourth, final, most conservative, and, I opine, best solution. Under the specified structural conditions it adds extra dependencies that will license sufficient \(s\)-dependencies for structures to be free of tangling: for *Combien il en mange* this requires making the DE a dependent of the MANGE, as shown in (470a–b).

\[
\text{a. Combiend il en mange?}
\]

\[
\text{b. Combiend il en mange?}
\]

Phenomena such as clitic climbing (as in Italian *Lei lo vuole vedere*, literally “She him wants to see” — “She wants to see him”) suggest that the clitics are raised or extracted,\(^{174}\) supporting the view that the displaced word is a dependent of the word it shares its enunciation with — loosely, that the clitic is a dependent of its host.

\(^{174}\) Volino (1990) argues reasonably that the Italian construction involves not extraction but rather something more like a variety of raising.
6

Where coenunciation is

6.1 Coenunciation and pronouns
So far a coenunciation analysis has been suggested for several areas of English, includ­ing contracted auxiliaries, possessive 's, interrogatives, adverbs and verbs. Our main purpose in the above discussion has been to demonstrate the existence of coenunciation. For the rest of this chapter the emphasis changes to an application of coenunciative ana­lyses to various further areas of grammar. We may thereby get an impression of how pervasive coenunciation may be.

We’ll look first at some of Dick Hudson’s speculative but ingenious suggestions of candidates for a coenunciation analysis (here presented using his form of presentation).\(^{175}\) (471) illustrates his form of presentation and is one of his suggestions for a way to handle ‘zero words’. I give (471) here for want of an apter place.

\begin{equation}
(471) \quad \text{HAD} + \text{BETTER} = \text{better}
\end{equation}

\(^{175}\) Two last-minute additions from discussion on wordgrammar@ucl.ac.uk: \text{FOR} + \text{EACH} = \text{per} \\
(Dick Hudson, again) and \text{TO} + \text{HOME} = \text{home} (Nik Gisborne).
6.1.1 Quantifiers and indefinite determiners

• **Another.**

(472)  

\[ \text{some more beer, five more people, another beer, another five people} \]

\[ \text{AN} + \text{MORE}_{\text{extra}} = \text{another} \]

Cf. *some extra beer, an extra five people, an extra beer.*

• **Other.**

(473)  

\[ \text{someone else,} \]

\[ \text{* someone else than John} \]

\[ \text{* someone other} \]

\[ \text{someone other than John} \]

\[ \text{MORE} + \text{ELSE} = \text{other} \]

• **One.**

(474)  

\[ \text{a big one} \]

\[ \text{one that's big} \]

\[ \text{* a one that's big} \]

\[ \text{AN} + \text{ONE} = \text{one} \]

Superficially, it looks as if an AN can have a ONE complement only as long as it is not immediately followed by the ONE: this restriction comes out as much less strange if we see it not as a quirk of syntax, which tends to be fairly regular, but as a quirk of morphology, which tends to be fairly irregular. In this case, we can say that the enunciation of an AN whose immediately following complement is a ONE is the enunciation of its complement. *She read just such a one* and *She read such a one as this* present a slight problem for that view, but it could be that the SUCH actually comes between the AN and the ONE, with the SUCH and the AN coenounced, as in (475a), or, if articles coenounce with the word following them, with the SUCH, the AN and the ONE coenounced, as in (475b).
6.1.2 Deictics

Another of Dick’s suggestions is (476).

(476) today, tonight, this morning, this evening

\[ \text{THIS + DAY} = \text{today} \]
\[ \text{THIS + NIGHT} = \text{tonight} \]

As he notes, coenunciation offers an effective way of explaining the otherwise thorny problem of lexical blocking. How come we can’t say this day (as a deictic adjunct)? — The answer is that we can, but the THIS and the DAY share the same enunciation, today. This coenunciation has to apply whenever the DAY follows the THIS; if they’re not adjacent, as in I’ll do it this very day, then there’s no coenunciation. To (476) we can add (477).

(477) last week, last year, yesterday, next week, next year, tomorrow

\[ \text{LAST + DAY} = \text{yesterday} \]
\[ \text{NEXT + DAY} = \text{tomorrow} \]

6.1.3 Possessives

In this section we’ll consider (i) how the determiner ’s, which for convenience I shall here call ‘Poss-S’, is a coenunciate, (ii) the syntax of poss-S, (iii) the relationship between poss-S and other possessive determiners (the suggestion being that they too are coenunciates), and (iv) the relationship between poss-S and other definite determiners.

Poss-S’s morphological manifestation, ’s, the morpheme ‘Z’, forms a single vocable with the morphological form of the word the poss-S coenounces with, as the test of repetition shows: Sophy’s...Sophy’s...Sophy’s vs. †Sophy’s...’s...’s. As is well known, poss-S is not always unambiguously morphologically manifest: in all lects the morpheme Z is phonologically, though not orthographically, ambiguous between poss-S and
the plural marker (dogs’(*s) — cf. children’*(s)), and in some lects the morphological form of poss-S is suppressed according to various phonological criteria (Hopkins’%(*s), Augustus’%(*s), Katz’%(*s)); see Zwicky (1987). Oddly, though, a sequence of two instances of the Z morpheme is marginally possible provided that they belong to a vocable that is an enunciation of two instances of poss-S:

(478) a. ? Your mother’s eyes are brown, but Sophy’s’s are green. [i.e. “Sophy’s mother’s are green”]
b. ? Your mother’s eyes are brown, but hers’s are green. [i.e. “her mother’s are green”]
c. Your mother’s eyes are brown but mine’s are green. [i.e. “my mother’s are green”]

For what it’s worth, my intuitions about (478a—c) are that they violate the spirit of the law (i.e. the grammar) but not the letter of the law, because the grammar has failed to provide for the rare possibility of their being said.176

There are good syntactic and semantic grounds for taking poss-S to be ditransitive (i.e. having both an inject and an object). The semantic grounds are that in both give Sophy chocolate and Sophy’s chocolate the semantic roles associated with Sophy and chocolate are the same: the referent of Sophy is possessor of the referent of chocolate. The syntactic grounds are that in Sophy’s chocolate both Sophy and chocolate are complements: we can tell this because both are subcategorized for and because the dependency between Sophy and the poss-S is obligatory (He ate *(Sophy)’s chocolate). Since poss-S has two complements, they must be distinct, chiefly because there must be some way to tell them apart, firstly in order to assign each distinct semantic roles, and secondarily in order to state the rule that if a word has more than one dependent of a specific type then the dependents must be conjoined (Sophy *(and) Edgar’s character *(and) chocolate). The choice of complement-types to distinguish the two complements of poss-S is rather limited, as will be discussed in a later chapter. The three initially most plausible analysis are presented in (479a—c).

---

176 A similar example I encountered was Is this book one of yours?, meaning not “Does this book belong to you?” but “Does this book belong to one of you?”. I actually find that ungrammatical, and prefer Is this book one of you’s?, but at any rate it is clear that the grammar offers no unproblematic options to the speaker in cases like these.
In (479a) Sophy is a ‘measure modifier’ of poss-S, as, say, in three grammes more chocolate or three more chocolates the three is a ‘measure modifier’ of the more. In favour of this analysis is the position of Sophy, since measure modifiers precede their heads, but against it there is the utter difference between the semantic relationship between Sophy and poss-S, on the one hand, and on the other hand the semantic relationship between a word and its measure modifier. Furthermore, as (480a—b) show, poss-S can have a measure modifier (three is measure modifier of poss-S), though this does not clinch the case against (479a), since it could be that both three and Sophy are measure modifiers of poss-S, but Sophy, but not three, is also a complement of poss-S.

(480) a. Edgar is three years Sophy’s junior.

b. Three years Sophy’s junior, Edgar felt constrained to exhibit the deference his age demanded.

In favour of (479b) we have, first, the pre-head position of Sophy, which is the characteristic position of subjects, and, second, the chance to develop an extremely neat

---

177 These are curious examples. While determiners functioning predicatively can have measure modifiers, as in (i—iii), analogous to (iv), these express duration of the situation during which the referent of the subject of the determiner (Edgar, in these examples) belongs to the category expressed by the determiner, whereas in three years her junior the three expresses “how junior”, not “junior for how long”. It is as if the three is not a measure modifier of the poss-S but rather is a lexically selected measure modifier of the junior, even though that would result in the tangling illustrated in (v). I leave the search for an explanation to the future.

(i) Three years a brigadier, Edgar was leading a fulfilled life.
(ii) Three years Sophy’s lover, Edgar was leading a fulfilled life.
(iii) Three years (the) king, Edgar was leading a fulfilled life.
(iv) Three years married, Edgar was leading a fulfilled life.
(v) three years Sophy poss-S junior
analysis of the semantics of phrases like *Sophy’s destruction of the city*, wherein *Sophy* would be subject of *destruction*. But alluring though (479b) is, there is in this case clinching evidence against it: in (481a–d) the poss-S has a subject, and it is not *Sophy*.

(481) a. He considers it *Sophy’s* best idea

b. With Edgar *Sophy’s* husband, we hesitated to tell him of her indiscretions.

c. *Sophy’s* best idea, it nonetheless failed to win support.

d. Edgar *Sophy’s* husband, we hesitated to mention to him her indiscretions.

So we come to (479c), with poss-S as ditransitive. As has already been said, the semantics supports this analysis. The syntactic evidence against is the pre-head position of the putative inject: with all other ditransitives, if X is an inject of (and, more generally, a complement of) and an s-dependent of Y then X will follow Y. But this is easily overridden by a simple stipulation, and in other respects the word order of the putative inject supports the idea that it is genuinely an inject: like all ditransitives’ injects that are also s-dependents of the ditransitive, the inject of poss-S is a tail of poss-S. (‘Tail’ is defined below in (482b).) On the other hand, so long as the preceding dependent of poss-S was a complement of poss-S it would still end up as a tail of poss-S even if it wasn’t an inject of poss-S, since complements of determiners are tails of the determiners. (482a–d) give the relevant rules.

(482) a. Every word X that is a dependent of Y and that is not separated from Y by anything other than subordinates of X, is a ‘propinquant’ of Y.

b. Every tail of W is a complement of W, a dependent of W, and a propinquant of W.

---

178 O.E.D.: “near, neighbouring, adjacent”.
c. If X is an inject and s-dependent of Y, and Y is a ditransitive, then X is a tail of Y.

d. If X is a complement of Y and Y is a determiner then X is a tail of Y.

We move on now to the relationship between poss-S and possessive personal pronouns/determiners. Essentially we want to explain three things. First, why is it that poss-S and possessive pronouns behave alike, form a natural class, and, second, why is it that possessive pronouns don’t form a natural class — that is, in all respects in which possessive pronouns behave alike, so poss-S shares the same behaviour. Third, why is it that the inject can be any noun whatever except an intransitive personal pronoun?

(483a) shows a sample of the range of possible injects of transitive poss-S. (483b) shows the intransitive personal pronouns that apparently definitely can’t be injects of transitive poss-S, and (483c) shows what we say instead. (483d) lists uncertain cases where the morphology is unclear. The same pattern holds for (483e–g), except here the poss-S and possessives are intransitive.

(483) a. this’s cover, these’s cover, that’s cover, those’s cover, books’ covers, Sophy’s name, us gymnasts’ technique, you gymnasts’ technique, the book’s cover, this book’s cover, these books’ covers, that book’s cover, those books’ covers, ...

b. *me’s name, *us’s name, *you’s name, *her’s name, *him’s name, *them’s name

c. my name, our name, your name, her name, his name, their name

d. ? who’s name [<whose>], it’s name [<its>], one’s name

e. this’s, these’s, that’s, those’s, books’, Sophy’s, us gymnasts’, you gymnasts’, the book’s, this book’s, these books’, that book’s, those books’, ...

f. *me’s, *us’s, *you’s, *him’s, *them’s

g. mine, ours, yours, his, theirs

h. ? who’s [<whose>], it’s [<its>], one’s, her’s [<hers>]

Some examples of properties peculiar to poss-S and possessive pronouns are illustrated in (484–485). Only poss-S and possessive pronouns can have a common noun own as their object; such an own is always an object of poss-S or a possessive pro-
Similarly, only poss-S and possessive pronouns can have an every as their object, (485a), only poss-S and possessive pronouns can have an each-and-every as their object, (485b), and every each-and-every is an object of poss-S or a possessive pronoun. Incidentally, it is relevant to our general topic of word-enunciation relations to note that there is a lexeme each-and-every the instances of which are common nouns requiring a common noun complement. That we are not dealing with a coordination of an each and an every is shown by the ungrammaticality of (485c—d).

(484) a. my/our/your/his/her/its/their/whose/one’s own (opinion)
    b. Sophy’s own (opinion)
    c. * the own (opinion)
    d. * an own (opinion)

(485) a. my/our/your/his/her/its/their/whose/one’s/Sophy’s every *(wish)
    b. my/our/your/his/her/its/their/whose/one’s/Sophy’s each and every *(wish)
    c. * my/our/your/his/her/its/their/whose/one’s/Sophy’s every and each (wish)
    d. * my/our/your/his/her/its/their/whose/one’s/Sophy’s each (wish)

The facts presented in (483a—g) and (484—485) are not in themselves particularly tricky to deal with. The generalizations pertaining to (484—485) could be captured by invoking a category ‘Possessive’ containing both poss-S and the possessive personal pronouns. (483a—g) can be accounted for easily by stipulating that possessive personal pronouns can’t be injects of poss-S. But the intuitive explanation for (483a—g) is that it involves lexical blocking: for example, *me’s is preempted by my/mine. And we’ve also seen that lexical blocking is best analysed as coenunciation. So, I propose that possessive personal pronouns and determiners in fact involve coenunciation.

The analysis of this own as a common noun (which has another common noun as its complement) is from Hudson (1990). A number of the observations made below about common nouns like these that have common noun complements and occur as objects only of certain determiners also come from Hudson (1990).

An own *(opinion) is marginally acceptable, on analogy with an own-label product, an own goal, but these, and an own opinion might better be analysed as compounds, given the ungrammaticality of *an own firmly held opinion versus her own firmly held opinion.

This is how Hudson (1990) does things.

Apparently, older speakers accept Let’s meet at John’s but not %Let’s meet at mine. My analysis certainly doesn’t lead us to expect such a difference. Evidently the variety of intransitive poss-S that refers to someone’s home mustn’t have a personal pronoun as its inject. As we would expect, this peculiarly arbitrary restriction has disappeared from the lects of younger speakers.
These coenunciations are virtually entirely regular coenunciations of a pronoun and a poss-S, the only irregularity being morphological. The default morphological rule is, roughly, (486a), and (486b–h) comprise the entirety of the rules necessary to account for so-called possessive pronouns and determiners.\footnote{The rules are only semi-formalized, since the necessary formal machinery for handling matters morphological is insufficiently well worked-out for WG.}

(486) a. If X is enunciation of Y and of S, and if S is poss-S, then X is an instance of a structure consisting of a concatenation of the morphological structure proper to Y with the morpheme Z.

b. If X is enunciation of Y and of S, if S is transitive poss-S, if Y is inject of S, and if Y is a ME then X is my.

c. If X is enunciation of Y and of S, if S is intransitive poss-S, if Y is inject of S, and if Y is a ME then X is mine.

d. If X is enunciation of Y and of S, if S is poss-S, if Y is inject of S, and if Y is a HIM then X is his.

e. If X is enunciation of Y and of S, if S is transitive poss-S, if Y is inject of S, and if Y is a HER, an US, a YOU or a THEM then X is an instance of the morphological structure proper to Y.

f. If X is enunciation of Y and of S, if S is poss-S, if Y is inject of S, and if Y is an US, then the morphological structure proper to Y is our.

g. If X is enunciation of Y and of S, if S is poss-S, if Y is inject of S, and if Y is a YOU, then the morphological structure proper to Y is your.

h. If X is enunciation of Y and of S, if S is poss-S, if Y is inject of S, and if Y is a THEM then the morphological structure proper to Y is their.

(486b–d) are fairly straightforward. The effect of (486e) is to prevent the enunciation of her/our/your/their in her/our/your/their face from ending in an instance of the morpheme Z. The effect of (486f–h) is to override the default morphology from us/you/them to our/your/their.

The relevant points about the coenunciative nature of 'possessive pronouns' have been made, but I shall now briefly consider where this leaves the system of definite pronouns. We can start by noting that the categories of possessives in general, which had been required to state the necessary generalizations about (484–485), and possessive
personal pronouns/determiners in particular, which was necessary to stipulate the apparent restrictions shown in (483a–g), disappear; all we have is poss-S. The data in (484–485) illustrates properties peculiar to poss-S.

Ignoring indefinite, relative and reflexive pronouns, I tentatively offer (487) as a classification of pronouns. The categories, which cross-classify, are shown in capitals, and examples of instances of the intersecting categories are shown in italics. Thou idiot and you idiot are possible only as vocatives, and therefore don’t properly belong in this paradigm; also, the you in you idiot would be singular, but in other respects you behaves as a plural even when it refers to a singular — cf. O Sophy, you are generous to yourself.

(487)

<table>
<thead>
<tr>
<th></th>
<th>SINGULAR</th>
<th>SINGULAR</th>
<th>PLURAL</th>
<th>PLURAL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>INTRANSITIVE</td>
<td>TRANSITIVE</td>
<td>INTRANSITIVE</td>
<td>TRANSITIVE</td>
</tr>
<tr>
<td>DEFINITE DEMONSTRATIVE</td>
<td>PROXIMAL</td>
<td>this</td>
<td>this book</td>
<td>these</td>
</tr>
<tr>
<td>DEFINITE DEMONSTRATIVE</td>
<td>DISTAL</td>
<td>that</td>
<td>that book</td>
<td>those</td>
</tr>
<tr>
<td>DEFINITE IT-THE</td>
<td>PERSONAL</td>
<td>it</td>
<td>the book</td>
<td>them</td>
</tr>
<tr>
<td>DEFINITE IT-THE</td>
<td>DITRANSITIVE</td>
<td>X's book</td>
<td>X's</td>
<td>X's books</td>
</tr>
<tr>
<td>HIM</td>
<td>PERSONAL</td>
<td>him</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HER</td>
<td>PERSONAL</td>
<td>her</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ME-US</td>
<td>PERSONAL</td>
<td>me</td>
<td>us</td>
<td>us people</td>
</tr>
<tr>
<td>YOU</td>
<td>PERSONAL</td>
<td>thee</td>
<td>you</td>
<td>you idiots</td>
</tr>
</tbody>
</table>

The motivation for this classification is to minimize the number of basic categories (i.e. those not defined by intersections) while allowing the necessary generalizations to be captured. Here are some justifications and some explanations of the classification.

**• It** is analysed as the intransitive (that is, objectless) counterpart of the: they are instances of the same lexeme but happen to have different phonology. This move simplifies the range of basic categories, and also brings out a pattern shared with demonstratives: it (an intransitive personal IT-THE) and intransitive demonstratives can refer to humans only when subject of an equative be, (488a–b). Otherwise they refer to non-humans, (489a–b). But when the demonstrative is transitive (i.e. has an object) or we

---

134 This idea is either the same as or similar to one once suggested by Dick Hudson (p.c.).
have *the* (a transitive personal *it*-**the**) instead of *it*, the pronoun can always refer to humans, (490a—b).

(488)  

a. It's Sophy.  
b. This/that is Sophy.

(489)  

a. ! He married it.  
b. ! He married this/that.

(490)  

a. He married the professor.  
b. He married this/that professor.

One fault of (487)'s classificatory system is that this generalization can't be stated in terms of an intersection of categories: it applies to all definites that aren't ditransitives. That, I suppose, is at least learnable, since the exceptional behaviour of ditransitive definites in this respect (cf. *She likes Edgar's mother and he likes hers*, where *hers* refers to a human), is clearly learnable from positive evidence (as is the exceptional status of the (488) pattern with respect to the general rule of (489)).

Interestingly, in this respect *What* is similar to *It*-**the**, as Dick Hudson has pointed out to me. A transitive *what* can refer to a human, (491a), and can be plural, (491b), but an intransitive *what* cannot, (491c–d) (if not ungrammatical as a plural it is at least strikingly odd). (I cannot conceive of a plausible example to test whether an intransitive *what* subject of an equative *be* can refer to a human.)

(491)  

a. What student has become her friend?  
b. What students have become her friends?  
c. ! What has become her friend?  
d. ? What have become her interests?

To refer to humans, instead of an intransitive *what* we of course use *Who*.

(492)  

a. Who has become her friend?  
b. Who have become her friends?

*Which* is free from the restrictions on *What*.
(493) a. Which student has become her friend?
b. Which students have become her friends?
c. Which has become her friend?
d. Which have become her friends?

I have no full explanation of this to offer, though noting the phonological resemblance between *what* and *that*, Dick Hudson has suggested that *what* involves a coenunciation of something with a *THAT*. Ideally a full explanation will explain why a relative *what* can only be a free relative pronoun, while a relative *who* or a *which* can only be an appositive relative pronoun.

- Only an *IT-THE* can have a *SOLE* or an *ONLY* as its object.

(494) a. the sole/only problem
   b. Sophy's/her sole/only problem
   c. * a sole/an only problem
   d. * this/that sole/only problem

Further support for the analysis of poss-S as an instance of the lexeme *IT-THE* is presented in the next section.

- As discussed earlier, only a ditransitive *IT-THE* can have an *OWN*, an *EVERY* or an *EACH-AND-EVERY* as its object.

(494) a. Sophy's own/every/each and every suggestion was dismissed by the committee.
b. * The own/every/each and every suggestion was dismissed by the committee.

- Only a personal *IT-THE* can have an *ODD* or a *SAME* as its object.\(^{185}\)

---

\(^{185}\) This is observed by Hudson (1990).
(494) a. She read the odd/same book.
   b. * She read Edgar’s odd/same book.
   c. * She read an odd/a same book.

Also only a personal it-the can have a relative which as its object — %the which — though possibly any it-the can: ?his which. (This construction is decidedly moribund.)

• Only a definite can have a very as a dependent.

(494) a. Sophy’s/her very name
   b. the very day
   c. this/that very day
   d. * a very day
   e. * us/you very people

Contrary to Hudson’s (1990) analysis of this very, it is not like own, sole, only and their ilk; that is, it is not a countable common noun whose complement is a common noun. Our evidence is Sophy’s/her very own book and the very same day, and perhaps ?the very odd cigarette. Since every own is an object of a ditransitive it-the and every same is an object of a personal it-the, it can’t be that the very is object of the definite. However, it’s not clear what the function of the very is. Since it comes between the definite and its object, it would appear not to be an object. And it can’t be an inject, since in Sophy’s/his very own the inject of the definite is a sophy/him. Perhaps the very is both a complement, which would explain its position, and some kind of adjunct, which would explain why it doesn’t seem to be one of the run-of-the-mill varieties of complement.

• The inject of a ditransitive verb can, in some lects, including mine, follow the object only if both are intransitive personal pronouns. (In other lects (494a) is ungrammatical.)
(494)  
a. % Give it me. ["Give the book to me"]
   b. * Give the book me. ["Give the book to me"]
   c. * Give his me. ["Give his book to me"]
   d. * Give this me. ["Give this book to me"]
   e. * Give it her mother. ["Give the book to her mother"]
   f. * Give it his. ["Give the book to his mother"]
   g. * He gave it you/us people. ["Give the book to you/us people”]

I'm not sure whether this test shows reflexives to be personals: *She gave it herself for Christmas.

- Only intransitive personal pronouns won't particle-shift, (495a—c) or heavy-NP shift, (496a—c).  \(^{186}\)

(495)  
a. * Put away it.
   b. Put away this/that.
   c. Put away the book.

(496)  
a. * Announce today it.
   b. Announce today this/that.
   c. Announce today every piece of news you deem relevant.

By this test, reflexives are personal pronouns: *?She's always putting down herself, *Comfort today yourself.  \(^{187}\)

---

\(^{186}\) Since I analyse particles as xcomps that are intransitive prepositions, the unmarked pattern is object before particle. Therefore the rule is that intransitive personals won’t particle-shift, and not that intransitive personals must particle shift (as would be the case if the default was that particles precede objects).

\(^{187}\) But see Rosta (1992) for fuller discussion of reflexives and particle shift.
6.2 Numerics, comparatives, superlatives

6.2.1 Numerics
Also suggested by Dick Hudson are coenunciative analyses of multiples, as in (497), and ordinals.

(497) once, twice, three times, four times
ONE + TIMES = once
TWO + TIMES = twice

To account for why we say *five thousand and twenty-two times* and not *five thousand and twenty-twice*, we must say that then *twice* arises only when a dependency holds between a TWO and a TIMES. Either *five thousand and twenty-two* is one word, but (obviously) not a TWO, or it is several words, but the TWO is not the root of the whole phrase.

As for ordinals, consider the following.

(498) first, second, third, fourth, fifth, sixth, seventh, eighth, ninth, tenth, five hundred and twenty-first, five hundred and twenty-seventh

Going by meaning, we’d want the following bracketing: [*five hundred and twenty-seventh*]. That is, the number, or the root thereof, is a complement of a word expressing ordinality. A possible analysis is to posit a lexeme TH which shares its enunciation with the preceding word (irrespective of whether the preceding word depends on it). The morphology of the enunciation is by default composed of the morphology of the number, plus the morpheme *th*, with some modification for *fifth* and *twelfth*, and total or near-total suppletion for *first*, *second*, *third*.

The relationship between ordinals and fractions is rather interesting. On analogy with ordinals we could posit a fractional variant of TH, with certain morphological differences. TWO + fractional TH = *half* when TWO is a dependent of TH, but *second* otherwise — cf. *Divide it into twenty-seCONDS*. FOUR + fractional TH = *quarter* or *fourth*, when FOUR is dependent of TH, but *fourth* otherwise — *Divide it into twenty-fourths*. If you slice a cake into twenty one pieces you probably end up with *twenty-firsts*.
6.2.2 Comparatives and superlatives

A good case can be made for analysing comparative and superlative forms of adjectives as coenunciations of the adjective with a preceding MORE or MOST.188

(499) MORE + BLACK = blacker
MORE + GOOD = better
MORE + BAD = worse
MOST + BLACK = blackest
MOST + GOOD = best
MOST + BAD = worst

In a sense this analysis is actually supported by non-standard more blacker, more better, more worse. In both standard and these non-standard lects, the morphological form of the enunciation of the adjective changes when the adjective is modified by a MORE, while only in the pattern of standard English does the MORE share its enunciation with the adjective.

As well as the morphologically irregular coenunciations with GOOD and BAD, there are also those in (500).189

(500) MORE + FEW = fewer (and/or less?)
MOST + FEW = fewest
MORE + LITTLE = less
MOST + LITTLE = least
MORE + MANY = more
MOST + MANY = most
MORE + MUCH = more
MOST + MUCH = most

---

188 As Dick Hudson has pointed out, only a certain type of MORE does not coenounce at all:
(i) It's more black than grey.
(ii) It's more bad than good.
This type of MORE is paraphrasable by (iii−v).
(iii) It's black rather than grey.
(iv) It's black more than grey.
(v) It's not so much grey as black.

189 The alternation little milk/less milk suggests that less is MORE + LITTLE, but in this case it is surprising that though less happy is fine, little happy is not. We must resort to stipulating that any LITTLE that modifies any adjective must have a MORE dependent with which it shares its enunciation.
The morphologically invisible coenunciation of MANY and MUCH is found also when they coenounce with ENOUGH.

\[(501)\]
MANY + ENOUGH = enough
MUCH + ENOUGH = enough
numerous enough
* many enough
few enough
* much enough
little enough

The evidence for the equivalent behaviour of MORES and comparative adjectives, and of MOSTS and superlatives, is pretty overwhelming. Both can occur as complement of a THE, \((502a-d)\). Both can have a complement that is an OF, \((502e-h)\). The OF complement of a MORE or comparative must be definite, \((502e-f)\), while the OF complement of a MOST or superlative can be definite or indefinite, \((502g-h)\).

\[(502)\]
\begin{align*}
a. & \text{ She read the more pleasant.} \\
b. & \text{ She read the shorter.} \\
c. & \text{ She read the most pleasant.} \\
d. & \text{ She read the shortest.} \\
e. & \text{ She read the more pleasant of *(the) books.} \\
f. & \text{ She read the shorter of *(the) books.} \\
g. & \text{ She read the most pleasant of (the) books.} \\
h. & \text{ She read the shortest of (the) books.} \\
\end{align*}

Another way in which MORES and comparatives are alike is that they allow the same range of premodifiers, \((503a-b)\), most which are not allowed by regular adjectives, \((503c)\), while the modifiers that are allowed by regular adjectives, \((503e)\), are not allowed by MORES or comparatives, \((503e-f)\).
a. (all/any/much/none) the closer, three miles closer, much closer, no closer, any closer, a lot closer, loads closer, tons closer, lots closer, little closer, a little closer, a bit closer, a tad closer, ...

b. (all/any/much/none) the more distant, three miles more distant, much more distant, no more distant, any more distant, a lot more distant, loads more distant, tons more distant, lots more distant, little more distant, a little more distant, a bit more distant, a tad more distant, ...

c. *((all/*any/*much/*none) the close, *three miles close, *much close, *no close, *any close, *a lot close, *loads close, *tons close, *lots close, *little distant ...

d. so close, how close, this close, that close, very close, quite close, dead close, well close, megaclose, extraordinarily close, ...


f. *so more distant, *how more distant, *this more distant, *that more distant, *very more distant, *quite more distant, *dead more distant, *well more distant, *mega more distant, *extraordinarily more distant, ...

They are also alike in in that they have a THAN complement, (504a—b), which can extrapose across a noun, (504c—d).

a. That is more insightful than this.

b. That is shorter than this.

c. She read a more insightful book than this.

d. She read a shorter book than this.

These THAN complements also serve to further demonstrate the difference between comparatives and regular adjectives. As (505a) shows, a regular adjective cannot, when attributive, have a complement. (505b—c), though, are grammatical because it is not the attributive adjective but the MORE/comparative that has a complement. (505d) is ungrammatical because at is complement of an attributive adjective.
(505) a. * every good at maths student

b. every better than average student

c. every more astute than average student

d. * every better at maths than average student

For some reason, (506a) is ungrammatical, but since (506b) is grammatical and (506c) isn’t, I take it that my point holds good.

(506) a. * a better than Sophy student

b. a better student than Sophy

c. * a better student at maths

---

190 A better student at maths in speech possibly has the status of an ‘acceptable performance error’, just as a good student at maths does. The relevant point is that the degree of acceptability of these two phrases is the same.
Mosts and superlatives are also alike in that they allow an xcomp, as in (507a–e).

(507) a. the best person to go  
b. the most charming person to have graced this podium in a long while  
c. the most comprehensive report ever to be compiled  
d. the most accurate to be published so far  
e. the shortest to receive acclaim for several years

The xcomps are the tos. There are several reasons to conclude that it's the most that licenses the xcomp. That it is not the adjective that licenses the xcomp can be seen from the ungrammaticality of (508a–b).

(508) a. * This is accurate to be published (so far).  
b. * This is short to receive acclaim (for several years).

We can tell it's not the noun that licenses the xcomp because, firstly, the noun is omissible, as (507d–e) show. Secondly, the xcomps in (507a–e) are of a non-extractional variety; that is, their subject is not their extractee. Though nouns can have an adjunct that is a non-extractional to, as in (509a–b), these have a meaning paraphrasable by (510–511), whereas (507b–e) are paraphrasable by (512a–e), which are markedly different from (510–511). The tos in (509a–b) refer to future, potential situations, while the tos in (507b–e) refer to past or perhaps present situations.

(509) a. The man to save us will appear.  
b. She should be the person to go.

(510) a. The man who will save us will appear.  
b. She should be the person who will go.

(511) a. The man who is to save us will appear.  
b. She should be the person who is to go.
(512) a. the person such that it would be best if they went
    b. the most charming person that has graced this podium (in a long while)
    c. the most comprehensive report that’s (ever) been compiled
    d. the most accurate that’s been published (so far)
    e. the shortest that’s received acclaim (for several years)

As (507b—e) show, and (513a—d), which lack the to complement, and the unacceptable most-less (514a—b) confirm, mosts and superlatives can also have a characteristic time adjunct.

(513) a. the most charming person in a long while
    b. the most comprehensive report ever
    c. the most accurate so far
    d. the shortest for several years

(514) a. *the charming person in a long while
    b. *the comprehensive report ever

Also, observe how mosts and superlatives license negative polarity items, such as ever.

(515) a. the best film ever (made)
    b. *a/the good film ever (made)

(516) a. the most exciting film ever (made)
    b. *an/the exciting film ever (made)

(517) a. the first film that impressed me at all
    b. *a/the film that impressed me at all

The syntax of (507a—e) is something of a goldmine for future research, but I shall leave it undelved here, since it is irrelevant to our present purpose, which is to show that mosts and superlatives behave alike, arguably because superlatives involve coenunciation. I mentioned above that ordinals and fractions appear to involve coenunciation, and it could be that such coenunciations involve some variety of most. Superlative mosts, fractions and ordinals (and next and last which pattern with ordinals) all have an of complement: the nicest/seventh/next/last of the cakes, a seventh of the cake. Like sup-
eralative 
mosts, ordinals (i) semantically pick out a single individual from a set, and (ii) also allow an xcomp, (518a–e).

(518) a. She met *(the) next (person) to arrive.
   b. She met *(the) last (person) to arrive.
   c. She met *(the) first (person) to arrive.
   d. She met *(the) second (person) to arrive.
   e. She met *(the) five hundred and ninth (person) to arrive.

All fractions, except for half (Eat (a) half (of) the cake, Eat *(a) third *(of) the cake), behave like regular countable common nouns. Ordinals are similar to superlatives, but they differ in certain ways. For example, ordinals can occur in the pattern shown in (519a–e), where the ordinal has an xcomp, and is functioning as an xcomp rather than a complement of a THE. As (520a–e) show, this pattern isn’t available for superlative 
mosts. In this respect ordinals are like the countable common nouns in (521a–d), which can be an xcomp rather than a complement of a determiner when they have the meaning “have the role of X”, where X is a lexically-specified relational concept (e.g. “cobbler to”/“president of”/“sister of”/“xcomp of”).

(519) a. She was next (?person) to arrive.
   b. She was last (?person) to arrive.
   c. She was first (?person) to arrive.
   d. She was second (?person) to arrive.
   e. She was five hundred and ninth (?person) to arrive.

(520) a. * She is best to go.
   b. * She is most charming to have graced this podium in a long while.
   c. * This is most comprehensive ever to be compiled.
   d. * This is most accurate to be published so far.
   e. * This is shortest to receive acclaim for several years.

(521) a. She was cobbler to the king.
   b. She was president of the republic.
   c. She is sister of Edgar.
   d. This is xcomp of that.
Another respect in which superlative *most* differ from ordinal *most* is that superlative *most* precede the adjective they coenounce with while ordinal *most* follow the word they coenounce with. The evidence for this is that uncoenounced superlative *most* precede the adjective, and if coenounced superlative *most* followed the adjective they coenounce with then tangling would result. (522a—b) show the structures differing according to the position of the *most*.

\[
\begin{align*}
(522) \quad \text{a.} & \quad \text{She is the keenest to go.} \\
(522) \quad \text{b.} & \quad \text{She is the keenest to go.}
\end{align*}
\]

We can see that ordinal *most* follow the word they coenounce with because if they preceded it then associations between words and enunciations would cross — that is, if the ordinal *most* is to be adjacent to the word it coenounces with, as it must be, then it must follow it. (523a—b) show this. The structure in (523c) is ruled out by projectivity (the diagram arbitrarily shows the AND as the root of the number — the syntactic structure internal to the number is not an issue here). However, if *five hundred and twenty two* is one word the enunciation of which comprises five vocables, then the structure in (523d) is in principle possible, though the morphological rules generating the phonological form of the enunciation would be excessively complicated.\(^{191}\)

\[^{191}\text{If, contrary to our present assumptions, one word can have more than one enunciation then if the most precedes the number then enunciation associations will cross, as (i—ii) show.}\]

\[
\begin{align*}
(i) & \quad \text{five hundred and twenty second}
\end{align*}
\]
In this respect ordinals resemble the — perhaps coenunciate — words foremost, hindmost, topmost, bottommost, leftmost, rightmost, centremost, utmost, outermost, uppermost, innermost, nethermost, furthermost, northernmost, southernmost, easternmost, westernmost. In each of these bar the last four the first element is either clearly a noun (fore, centre), or at least not clearly an adjective — top, bottom, left and right have noun-like and preposition-like properties as well as adjective-like ones, hind, ut(ter), outer, upper, inner and nether can, unlike adjectives, usually function only as premodifiers of nouns, and further and outer, upper, inner, nether and ut(ter) look, at least morphologically, like comparatives, which, as we’ll see next, are not straightforwardly adjectives. So, in summary, we could make the generalization that a most can coenonuce with a following adjective or a preceding non-adjective.

Given that a complement of a noun is by default an of, and a complement of a determiner is a common noun, a possible analysis is one where more, comparatives, mosts and superlatives are all common nouns. However, a problem for taking them to be common nouns is that unlike common nouns they can’t have an adjective adjunct, as (524a—d) show, so in fact if more, comparatives, mosts and superlatives are all nouns, they belong to a class that subsumes common nouns, such that it is a property of

Note, incidentally, how allowing a word to have more than one enunciation would predict that two vocables can’t both be enunciation of the same two words, since associations would inevitably cross. Plus others unfamiliar to me: hindermost, upmost, inmost, headmost, endmost, undermost, uttermost, hithermost.
this class that they are complements of a determiner, while it is a property only of common nouns that they can have an adjective adjunct.

(524) a. * She read the momentary more pleasant of the books.
b. * She read the convenient shorter of the books.
c. * She read the momentary most pleasant of the books.
d. * She read the convenient shortest of the books.

To complicate matters, Mores, comparatives, Mosts and superlatives can have adverb adjuncts, as in (525a–d), which makes them look like adjectives.

(525) a. She read the momentarily more pleasant of the books.
b. She read the conveniently shorter of the books.
c. She read the momentarily most pleasant of the books.
d. She read the conveniently shortest of the books.

There is no reason in principle why they can’t be simultaneously adjectives and nouns, but in fact it is far from clear that possession of an adverb adjunct diagnoses adjective- hood. Rather, any word functioning predicatively (however we end up defining that concept) can have an adverb adjunct: in (526) a noun functioning predicatively has an adverb adjunct.

(526)
6.2.3 Comparative adverbs

When the dependent of a **ly** is a **good**, their shared enunciation is, irregularly, **well**, (527a—c).

(527) a. She writes well.
    b. She writes badly.
    c. She writes poorly.

In general, coenunciations of a **more**, an adjective and a **ly** are not possible: *darkerly vs. more darkly*. These are ruled out perhaps by a conflict between the rules that build the morphology of **blacker** and the rules that build the morphology of **blackly**. But two of these three-way coenunciations are possible, (528a—b). 193

(528) 

\[
\text{[[[MORE] GOOD] LY}_{\text{ADV}]} = \text{better} \\
\text{[[[MORE] BAD] LY}_{\text{ADV}]} = \text{worse}
\]

a. She writes better than him.
   b. She writes worse than him.

It is arguable that (529—530) involve an adjectival **ly** whose dependent is a **good**, a **bad** or a **poor**.

(529) a. She is well.
    b. She is ill.
    c. She is poorly.
    d. She is sickly.

(530) a. Well people struggle.
    b. Ill people struggle.
    c. Poorly people struggle.
    d. Sickly people struggle.

193 If, as I have suggested, all adverbs involve coenunciation of an adjective with a **ly**, then we also find three-way coenunciations in **faster, sooner, oftener** as in *She'll eat faster/sooner/oftener*. This seems to confirm the suggestion that there is a morphological restriction against **-erly** (cf. Zwicky 1989, 1995). Also, in colloquial speech **quicklier, beautifullier** etc. are possible; morphological constraints tend to be thus violable (cf. **un-ring-up-able**).
When comparative, these behave in two different ways, (531–532), which perhaps can be accounted for on the basis of the different structures shown, the difference being whether the more is dependent of the adjective or the (adjectival) ly. In this case, if the more is dependent of the adjective then, given (528a–b) we expect to find the forms better and worse in (531a–b). And if the more is dependent of the adjectival ly then, given the rules responsible on the one hand for the forms in (529–530), and on the other hand for regular coenunciations of a more with an adjective, we expect to find the forms weller and iller in (532a–b). These predictions seem at least to be on the right track.

(531) [[[MORE] GOOD] LY]  
[[[MORE] BAD] LY]  
  a. She got better/?weller/?more well than she had been.  
     [cf. She got well!/good.]  
  b. She got worse/iller/more ill than she had been.  
     [cf. She got ill!/bad.]

(532) [[[MORE] [GOOD] LY]  
[[MORE] [BAD] LY]  
  a. She is weller/more well/*better than him.  
     [cf. She is well more than him.]  
  b. She is iller/more ill/*worse than him.

In support of this analysis, we have (533–534), where the forms are exactly as would be predicted given the assumptions we've been making.

(533) [[[GOOD] LY] LY]  
[[[BAD] LY] LY]  
  a. She looked at it well-ly.  
  b. She looked at it ill-ly.

---

This data may have important implications for an analysis of the grammar of comparative constructions.
a. She looked at it {more well-ly/*better} than him.
b. She looked at it {more ill-ly/*worse} than him.

And finally, it might be thought that in (535–536) there may also be a severely distributionally restricted form of Ly that is perhaps nominal. As these and the previous examples show, GOOD + LY always has irregular morphology, and BAD + non-adverb LY has irregular morphology.

(535) a. This news bodes *(well).
b. This news bodes *(ill).

(536) a. She means (him) well.
b. She means (him) ill.

However, if these have the structure [[GOOD] Ly^noun], then we predict that this would give rise to two sorts of comparative, one where a more modifies the Ly, giving [[MORE] [GOOD] LY], more well/ill (but not weller/iler, because the Ly is not an adjective), and one where the more modifies the adjective, giving [[MORE] GOOD] LY, better/worse. As (537–538) show, it appears that only the former prediction holds good, so we ought therefore to conclude that well and ill in (535–536) are non-coenunciate nouns meaning, very roughly, “goodness” and “badness”.

(537) a. Today’s news bodes {?better/?weller/more well} than yesterday’s.
b. Today’s news bodes {?worse/?iller/more ill} than yesterday’s.

(538) a. She means (him) {*better/*weller/?more well} than she means us.
b. She means (him) {*worse/*iller/?more ill} than she means us.

6.3 Coenunciation and verbs

We begin by returning to Hudson’s (1990) analysis of gerunds, which was presented in §5.4.3. The structures he proposes are given again in (539a–b).
There are a number of pros and cons to this analysis. The pros are as follows.

- The analysis captures the way the gerund as a dependent behaves like a noun, while as a head it behaves as its usual non-nominal self.
- The analysis is a step towards capturing the way verbs as dependents do not constitute a natural class. (I will expand on this point below.)

The cons are as follows.

- But for coenunciation, Hudson’s analysis would rely on the invariably regular morphology of gerunds. He takes the *ing* to be a clitic in the traditional sense, morphologically realizing the gerund, but with its position determined by its host. However, this rests upon the assumption that the single morpho-phonological structure of the enunciation of coenunciate words is necessarily a concatenation of the morpho-phonological structure proper to each of the words it represents. The existence of ‘articled prepositions’, and most of the rest of the coenunciation data presented so far, shows that this assumption is incorrect.
- The dependency structure of (445a) cannot give rise to a projective s-dependency structure. Depending on whether the *ing* precedes or follows the *kiss* we’d get one of the defective structures (540a–b).

(539) a. Him kissing her delighted them.

(540) a. Him kissing her delighted them.
Faced with this problem, we can either revise the dependency structure so that it can yield a projective s-dependency structure, or we can seek an alternative account of gerunds. I will first mention an alternative account, but then I'll go on to take up Hudson's analysis, revise the dependency structure so it is compatible with projectivity, and extend the analysis into other constructions.

Instead of analysing gerunds as coenunciate, we could analyse them as instances of (i) Ing-participle, (ii) Verb, and (iii) Noun. The analysis works because — contrary to standard belief — Noun and Verb are very different categories. Noun is not so exceptional: the generalizations that apply to it are (i) it defines a class of words that can function as certain types of dependent, e.g. as subject of a verb, and (ii) every noun is arguably an instance of singular or of plural. Verb is unusual: the only rules that apply to instances of Verb are firstly that they may be instances of various other categories, i.e. finite, imperative, infinitive, ing-participle, passive or perfect, and secondly the generalization that all verbs belong to some larger class of predicative words that can have subjects and have certain kinds of adjunct, such as already. There are no generalizations specifically about verbs about what dependents they can have or what words they can be dependents of: not all verbs have subjects, and no word subcategorizes specifically for a verb (rather, all finites have subjects, and some words subcategorize for, say, an infinitive). A gerund, then, could be an ing-participle form of verb whose nominality makes it eligible to function as a dependent in the same way as other nouns. The failure of the gerund to be modified by an adjective shows it is not a common noun (just as in Hudson's analysis).

I am not aware of any obvious problems with this analysis, but this is not to say that a refinement of Hudson's, making use of the explanatory possibilities coenunciation affords, will not prove to be a better one. If such a refinement were part of a larger analysis that led to a redefinition of the category Verb, this would be a positive step, since the status of Verb as a natural class is extremely questionable, given that as mentioned above, only one generalization applies to verbs. In other respects, the categories that Verb generalizes across (i.e. infinitive, perfect, etc.) behave dissimilarly, which is another reason to wonder if they don't constitute a natural class. The table in (541) illust-

---

I mean 'Verb' in the traditional sense, generalizing across different inflectional forms.

This analysis was not considered and then rejected by Hudson. This is because he was assuming, without any real debate, that cross-classification can involve only a specified range of categories (i.e. morphosyntactic feature values). This would rule out any word being an instance of both noun and verb.
rates the behaviour of members of each of the classes in the disjunctive set of classes to which every verb must belong (these are the boldface rows in the table) and of various other major word classes, with respect to the following factors:

I. Is it ever okay with a preceding adjunct such as *suddenly* or *already?*

II. Can it function as an xcomp?

III. Can it function as a following adjunct of a noun (in the so-called ‘reduced relative’ construction)?

IV. Can it occur in the ‘absolute’ construction?

V. When it is not an xcomp or absolute, can it have a subject that is also its s-dependent in a non-finite sentence?

VI. When it is not an xcomp or absolute, can it have a subject that is also its s-dependent in a finite sentence?

VII. Can it function as an object?

VIII. Can it function as a subject?

<table>
<thead>
<tr>
<th>(541)</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
<th>VI</th>
<th>VII</th>
<th>VIII</th>
</tr>
</thead>
<tbody>
<tr>
<td>ing-participle</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>passive</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>esmerge FOR-TO</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>preposition</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>adjective</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>noun</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>non-esmerge TO</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>infinitive</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>non-adjectival perfect</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>?</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>finite</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>imperative</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>non-esmerge FOR</td>
<td>no</td>
<td>?</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>adverb</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
</tbody>
</table>

\(\alpha\) As in *Him forgetting her birthday precipitated some acrimony.*

\(\beta\) I.e. a FOR-TO whose extractee and whose subject are merged, as in, say, *a book for her to read from.*

\(\gamma\) Excluding so-called adjectives that can’t be predicative.
As in *Anyone a fan of Mel Gibson should write in*, and also as in *Curtains the colour of mouldy cheese have never appealed to her.*

As in *Someone to mend the pipe has turned up and One to fill this gap is still needed*. There are special semantic restrictions on these adjuncts when they're non-esmerge tos, which don't apply to such adjuncts otherwise.

Apparent non-esmerge to absolutes are possible, (i–ii), but are poor when they have an s-dependent subject, as in (iii).

(i) Ultimately to be commended, the proposal is still open to various criticisms.
(ii) To be continued after lunch, the lecture will be interrupted at this point.
(iii) ? Him to be met at the station soon, the necessary arrangements must be set in motion.

An apparent examples of a non-esmerge to as an absolute is *Soon to be demolished, the building is derelict*, but given the relative badness of *To be demolished, the building is derelict*, it might be better to analyse the example as involving an absolute that is an adjective *soon* obligatorily requiring a non-esmerge to as its xcomp.

Examples of adjectival perfects are *well-read, well-travelled (children), fallen (leaves)* and maybe *broken (windows)*.

But only of a HAVE.

Examples of the yeses in Column V are given in (542). The sorts of context we might find them in are captions to pictures, newspaper headlines, terse note-taking style in diaries, and so on.

In *She longs for him to sleep* there are certain advantages to treating *for* as xcomp of *longs*.

If *for* in *She longs for him to sleep* isn't xcomp of *longs* then it would presumably be object.

Ignoring idiomatic expressions like *Gently does it*.

---

(542) a. Sophy running into the room.

b. Sophy molested by an ostrich.

c. Sophy (for Edgar) to collect from the station.

d. Sophy in trouble.

e. Sophy naked.

f. Sophy a fan of Mel Gibson.

g. Sophy the colour of blood oranges.

h. Sophy to arrive at six o'clock.
Infinitives have an s-dependent subject only in what Akmajian (1984) named ‘Mad Magazine Sentences’,\(^{197}\) (543a). These are a kind of ‘echo exclamation’, though strictly speaking there needn’t be echoing, as (543b) shows.

(543) a. What? Sophy take a holiday?
    b. A: She said she’d give you a wedding present.
       B: What? *Sophy be generous?! I can’t believe it!

It is hard to find any acceptable example, even a Mad Magazine one, in which a non-adjectival perfect has an s-dependent subject.

(544) a. * Sophy been to Paris.
    b. *? What? Sophy seen a ghost?

As explained earlier, our explanation of such phenomena as double depictives and mediopassives with accessible demoted subjects require that every verb is xcomp of a possibly coenunciate aux-like word. On that analysis, (543a) must contain a morphologically invisible, coenunciate non-finite auxiliary\(^{198}\) whose xcomp is the take, and (544a—b) would either be ungrammatical or require us to recognize a morphologically invisible, coenunciate non-finite auxiliary have that is found only in sentences like (544a—b).

The verb classes in (541) hardly cluster together in any obvious way. But other patterns do emerge. With respect to factors I—VI, ing-participles, passives, esmerge for-tos, prepositions, adjectives and nouns, and perhaps non-esmerge tos, behave alike: they can all be xcomps, they can all be adjuncts of nouns, every absolute belongs to one of these classes, they can all be modified by a preceding adjunct like already or suddenly, and they can all have s-dependent subjects in non-finite sentences (and not necessarily just Mad Magazine ones), and (except for ing-participles) not in finite sentences.\(^{199}\) Infinitives and perfects are very similar to each other but not really to the others. So are

---

\(^{197}\) See also Aarts (1992: 38), Radford (1988: 330).

\(^{198}\) These are ‘cryptic supines’, which are discussed below.

\(^{199}\) I am pointing out this pattern simply to show that the patterns that exist in the data are not at all captured by the class Verb (as traditionally understood). The explanation for the pattern itself is not really relevant to the present discussion, but clearly we are seeing evidence of a category of words (‘predicators’) characterized by the property of ‘predicativity’. 
finites and imperatives. And so are non-esmerge for and adverbs, though only in that they lack all the properties that define the other clusters of similarity.

All things considered, some sort of reanalysis of the categorial status of verbs could lead to a somewhat simpler and more explanatory grammar. I propose now to pursue such a reanalysis, employing the grammar of coenunciations. The necessary exposition is rather tricky and will of necessity be gradual. I’ll first map out the general outline of the analysis and then go back over it in more detail. I begin by defining a couple of very general intercategorial relationships: Subclass and Subtype. Both these define a subset relationship, but Subtype additionally requires that every member of a category that has subtypes is a member of one of the subtypes of that category. An example of a category that has subtypes is Mammal: one cannot just be a mammal; one must also belong to some subtype of Mammal, such as Dog or Human.

(545) a. Subclass

\[ \forall a, \forall b, a \text{ is subclass of } b \text{ iff } [\forall c, c \text{ is instance of } a \land c \text{ is instance of } b] \]

“A is a subclass of B iff every instance of A is an instance of B.”

b. Subtype

\[ \forall a [\exists b, b \text{ is subtype of } a] : \forall c, c \text{ is instance of } a : \exists d, d \text{ is subtype of } a, c \text{ is instance of } d \]

“If A has a subtype, then every instance of A is an instance of a subtype of A.”

We also need the principle, (546), that by default all classes can intersect. If any instances of two intersecting categories would inherit conflicting properties from them, and if the conflict is not resolved by rule, then the intersection will remain empty. For example, a tiger can just about be a pillow, but can’t be a telephone; it is most implausible to suppose that either of these facts are stored in the encyclopaedic entries of Tiger, Pillow or Telephone. Only unexpectedly impossible intersections need be explicitly prohibited.

200 The relationships Subclass, Subtype and Instance (and its synonym ‘Member’) are all covered by Hudson’s (1990) ‘isa’ relationship. The chief innovation here is in making a distinction between subset relationships and membership relationships.

201 I can’t think of any examples.
Default cross-classification

∀a, ∀b, ∀c, a is subclass of c, b is subclass of c: ∃d, d is subclass of a, d is subclass of b

"If A is a subclass of C and if B is a subclass of C then there is a class that is a subclass of both A and B."

I'll notate the intersection of class X and class Y by ‘X&Y’ — e.g. ‘Tiger&Pillow’.

Our story starts with two categories, Verb and Aux. Roughly speaking, Verb contains infinitives and perfects, while Aux contains auxiliaries, infinitival to, and coenunciate items involved in finite verbs, passives, and ing-participles. Some important rules pertaining to verbs are (547a–c).

(547) a. Perfect is subclass of Verb
b. Every verb is xcomp of an aux.
c. Every perfect is xcomp of a HAVE.

The effect of (547a–b) combined is to require the HAVE that the perfect is xcomp of to also be an aux. More on HAVEs below. An attraction of (547b–c) is that it accounts for the very restricted and extremely similar distribution of infinitives and perfects, as illustrated in the table in (541).

Some important rules pertaining to auxes are (548a–e). (548a–b) state that auxes are subject-raising. (548d–e) will be explained shortly.

(548) a. Aux is subclass of Osmerge [i.e. an aux’s object and subject merge]
b. Aux is subclass of X-licence [i.e. an aux has an xcomp]
c. If X is xcomp of an aux then X is a verb.
d. A sense of an aux is also a sense of the xcomp of the aux.
e. An aux can be an ‘elliptic’.

(548d) says that auxes have the same sense as their xcomp. The referent of finite and imperative auxiliaries is exceptionally not an instance of their senses. For finites, the referent is a proposition — something like “it being the case that ∃x, x instance of
[sense of the finite]. Imperatives are commands for the addressee to make it the case that \( \exists x, x \) instance of [sense of the imperative]. Weak evidence for sense sharing is that with some negated auxiliaries, e.g. Don't go, She mustn't go, the negation element has scope over the aux syntactically, but semantically it negates not the proposition (the referent) but the situation (the sense of the aux and its xcomp).

(548e) says that auxes can be 'elliptics'. By this is meant that (i) the aux's xcomp can be omitted, in which case (ii) the aux optionally has an extractee. Coenunciate auxes can't be elliptics, because they are required to have the same enunciation as their xcomp, which couldn't happen if they had no xcomp. (549a—b) are examples where the aux (italicized) lacks an xcomp but has no extractee. Note that xcomps of verbs cannot always be omitted in the same way; cf. (549c).

(549) a. She may.
   b. She intends to.
   c. * She seems.

(550a—b) are examples where the aux has an extractee but isn't an elliptic.

(550) a. He wants to read every book she wants to write.

---

204 The explanation is that the object of a verb must either receive a semantic role from W (i.e. be a 'ject' of W) or be subject of an xcomp of W. Thus the objects in (i—iii) (where the verbs in (i—ii) are osmerges) are okay even if the xcomp is omitted, because they're jects. But in *She seems, the object (i.e. She) is neither a ject nor a subject of an xcomp.

   (i) She tried (to rise).
   (ii) She began (to dance).
   (iii) She persuaded him (to dance).

This constraint on objects of verbs does not apply to objects of auxes, since the object of an elliptic aux is neither ject nor subject of an xcomp.

205 The diagrams show coenunciate finite auxiliaries, for a full discussion of which sufficient space is lacking here.
b. He’s read every book she has written.

These are normal extractions in that, to take (550b) as an example, the book can be an extractee of the write because it is a complement (object and inject) of the write. Because the book is an extractee of the write and the write is a complement of the have, the book can be an extractee of the have. In (551a–b), however, the extraction of the book is not thus sanctioned by it being a complement. In these examples to and has are elliptics: they therefore lack an xcomp and are licensed to have an extractee.

(551) a. He wants to read every book she wants to.

[“He wants to read every book she wants to read.”]

b. He’s read every book she has.

[“He’s read every book she’s read.”]

This analysis, in which elliptic auxes optionally select for an extractee solves what would otherwise be a problem for a monostratal model of syntax. A polystratal framework (and in this respect even traditional models like that of Quirk et al. (1985) are polystratal) can offer an analysis wherein the extraction originates in the usual way, with its original extraction site being subsequently deleted (by ellipsis).

---

206 There are, though, remaining problems not solved by the analysis. Dick Hudson (p.c.) has pointed out to me data like (i–ii), where the extractee, on is apparently subcategorized for by the absent, ellipted rely.

(i) I rely on books on which she does
(ii) I wonder on which books she does

One must conclude either that ellipsis in fact works in a way such that the rely is in some wise present (cf. how when pointing to a pair of trousers or some bathroom scales one says Look at those rather than Look at that), or that on can be understood with the meaning (or meaninglessness) that it has when complement of a rely, even when no rely is syntactically present.
6.4 Supines

In this section we'll deal with other sorts of aux. I remain vague about the details of which classes besides aux they may belong to, but for convenience of description I'll call them 'supines'. The supines comprise (i) infinitival to, (ii) infinitive xcomps of verbs, (iii) some element coenounced in ing-participles, (iv) some element coenounced in passives. I'll discuss each of these in turn.

- **Infinitival to.** This is the intersection of Aux, For-to, and the lexeme To. Like all auxes, infinitival to (i) is subject-raising, (ii) has a verb as its xcomp, (iii) can be elliptic, and (iv) shares its sense with its xcomp. For-to is a category (originating in Hudson (1990)) each of whose instances is either a For or an aux and a To. Instances of For and of To are instances either of For-to, as in (552a–b), or of Preposition, as in (553a–b).

(552) a. It's time for her to go.
    b. It's time to go.

(553) a. She has a craving for chocolate.
    b. She turned to her neighbour.

There are two chief motivations for recognizing For-to as a category. Firstly, both infinitival to and infinitival for can be esmerges — that is, their extractee and their subject merge. (This doesn't quite define For-to as a natural class, since there is some evidence that imperatives and various auxes can be esmerges too; see Rosta (in prep.).) The other motivation is that wherever an infinitival for is subcategorized for or otherwise allowed, so is an infinitival to; cf. (555–560).

(555) a. She waited for him to yodel.
    b. She waited to yodel.

(556) a. She longed for him to yodel.
    b. She longed to yodel.

---

207 According to the O.E.D. there is some precedent for borrowing the term 'supine' from Latin grammar to denote a class of non-finite verb forms including to-infinitives.
The rules generating these examples would permit FOR-TOS in general to function as xcomps and adjuncts of their heads. Unfortunately we fail to predict that nothing subcategorizes specifically for an infinitival for. The present analysis allows subcategorization for either (i) a FOR-TO, (ii) a FOR-TO and TO, or (iii) a FOR-TO and FOR. But this third option never occurs. It may therefore ultimately prove better to alter the analysis so that it allows subcategorization only for either (i) a FOR-TO, or (ii) a FOR-TO and TO. This can be achieved by making FOR-TO a subtype of FOR, (561a-b). Every FOR is a preposition or a FOR-TO, and every FOR-TO is a FOR.

The rules for To would then be as in (562a-b). Every TO is either (i) a preposition or (ii) an aux and a FOR-TO. The phonological properties such a TO would inherit by virtue of being a FOR (i.e. /fo:/) would be overridden, either by stipulation or by Elsewhere principles, by the phonological properties it would inherit by virtue of being a TO (i.e. /tu/).208

---

208 There is a conflict between the default xcomp for a FOR-TO, which is a TO (as in for her to) and the xcomp of a FOR-TO that is also a TO, which is a verb (as in to go). It is interesting to ask whether this conflict needs to be resolved by stipulation, since it is a question of how WG’s mechanism of default inheritance works. According to my present assumptions, if a rule actually states that the default xcomp of a FOR-TO is a TO then the conflict will have to be resolved by stipulation, since the class of infinitival TOs is the intersection of (i) FOR-TO (whence the requirement that the xcomp be a TO), (ii) Aux (whence the requirement that the xcomp be a verb), and (iii) TO. The requirements of intersecting classes (i) and (ii) conflict. However, if (as seems reasonable) the default xcomp of words in general is a TO, unless otherwise specified, then the class of infinitival TOs is the intersection of (i) FOR-TO, which would be unspecified for what their xcomp is, (ii) Aux, which is specified for the requirement that the xcomp be a verb, and (iii) TO, which is also unspecified for an xcomp. In this case, the requirements of the intersecting classes don’t conflict, and the xcomp requirement from Aux should win out without need for stipulation.
(562) a. \( \forall a, a \) is subtype of Preposition, \( a \) is subtype of To
b. \( \forall a, a \) is subtype of Aux, \( a \) is subclass of For-to, \( a \) is subtype of To

So, rules for the examples in (555—560) would just select for a For-to. By default this comes out as for, but it is free to come out as to instead.

- **Infinitive xcomps of verbs.** In order to maintain the generalization that every verb is xcomp of an aux, and, more importantly, in order to allow the double depictive solution presented earlier, it is necessary to analyse apparent infinitive xcomps of verbs, as in (563a—h), as in fact consisting of an aux coenounced with its xcomp. That is, the structure is along the lines of (564), which is essentially the same as (463b) which concluded our earlier discussion of double depictives. I’ll call this aux ‘Cryptic’, since it shows no morphological manifestation.

(565) a. She helped him yodel.
b. She let him yodel.
c. She had him yodel.
d. She made him yodel.
e. She bade him yodel.
f. She has never known anyone yodel so well.
g. Outdoors you will find Wren create new green dimensions with sensitive landscaping that creates a community and not just a row of houses.  

h. She saw/felt/heard/noticed him yodel.

---

This solution to the double depictive problem requires every verb to be xcomp of an os-merge whose sense is also the sense of its xcomp. The 'object-related' depictive will be depictive of the verb, and the 'subject-related' depictive will be depictive of the aux. The aux and the verb xcomp must have the same sense, because both depictives semantically modify the same sense: in our examples what happens both during the meat's rawness and during his nakedness is him eating the meat.

- **Ing-participles and supine haves.** We return now to the problem of the nonprojectivity of Hudson's analysis of gerunds (illustrated in (540a–b) earlier). Two possible solutions are (567a–b), in which, depending on the order of the ing and the infinitive, the object of the infinitive is extraposed or the subject is extracted.
Both analyses are ad hoc: they both require rules stipulating the extraction or extraposition, and neither resembles other more familiar constructions. But there is a far better analysis.

We've already seen part of the motivation for analysing gerunds as involving co-enunciation: as dependents gerunds behave as nouns and as heads they behave as verbs. So we follow Hudson in analysing gerunds as involving a nominal ING whose complement is a verb. We also have our general rule that every verb is xcomp of an aux. The natural analysis is therefore one in which INGS are auxes and sometimes or always also nouns. Their syntax is illustrated in (568); Him kissing her has exactly the same dependency structure as Him to kiss her.

I propose that all ing-participles have this structure. The ING is an aux and, at least sometimes, a noun, as we can tell from its ability to be the subject of a verb, as we see in (568), and from its ability to be the complement of a preposition, as in She talked about him cavorting. At present I've no analysis to offer for adjectival ing-participles, such as (very) interesting (books), gum-chewing colonels. Ing-participle-like common nouns, as in the cleaning of the windows, should, I suggest, be analysed as non-coenunciate common nouns which are syntactically equivalent to other nominalizations like inspection, and whose enunciation's morphology is composed in the same way as the morphology of the enunciation of INGS.

Progressives are grammatically different in some way from other ing-participles, as can be seen from the acceptability and grammaticality contrast between (569a—b) on the one hand, and (570a—b) on the other.

(569) a. ! She is owning a car.
     b. * She is having swum.
Owning a car, she finds it easy to travel to work.

Having swum a mile, she felt entitled to lounge about for the rest of the day.

In some ways, the be plus progressive construction is like the have plus perfect construction. The variety of hases that have a perfect as their xcomp can occur in the extremely limited range of contexts exemplified in (571a–e). They can be finite, (571a), they can be xcomp of a modal, (571b), they can be xcomp of an infinitival to, (571c), they can be xcomp of an ing, (571d), they can be xcomp of an imperative auxiliary do, (571e), and they can be imperative, (571f).

Has she swum?

She could have swum.

She yearns to have swum.

Having swum a mile, she paused for a rest.

Do please have swum a mile by the time I get back.

Please have swum a mile by the time I get back.

But, they cannot be a progressive, (572a), they cannot themselves be a perfect xcomp of a have, (572b), they cannot be xcomp of a finite auxiliary do, (572c), and they cannot be xcomp of a cryptic supine, (572d–e).

She is having swum.

She has had swum.

She did have swum.

She helped/let/had/made/bade/saw/noticed/heard/felt him have triumphed.

She made him have been happy.

BES that have a progressive xcomp occur in almost the same grammatical contexts as haves that have a perfect xcomp, (573a–h). But, unlike a have, a be can be a perfect, (573i). And progressive xcomps of cryptic supines are possible, albeit not normal, as (573j–k) show.
(573)  a.  Is she swimming?
     b.  She could be swimming.
     c.  She yearns to be swimming.
     d.  Being swimming at the time, she was unable to attend.
     e.  Do please be swimming when I get back.
     f.  Please be swimming when I get back.
     g.  * She is being swimming.
     h.  * She did be swimming.
     i.  She has been swimming.
     j.  She helped/let/had/made/bade/saw/noticed/heard/felt him be having fun.
     k.  She made him be listening.

It might appear that progressives can be xcomp of a wider range of verbs than can perfects, since, say, (574a—b) but not (575a—b) are possible. But since (576a—b) are also possible, but (577a—b) aren’t, I conclude that (574a—b) and (576a—b) don’t contain progressives.

(574)  a.  She kept swimming.
     b.  She likes swimming.

(575)  a.  * She kept swum.
     b.  * She likes swum.

(576)  a.  She kept having been sick.
     b.  She likes having been successful.

(577)  a.  * She is having been sick.
     b.  * She is having been successful.

The constraints on the syntactic environment of HAVEs, I won’t investigate here. Ideally we should bring out the parallel between (572a) and (572c), and (573g—h).

   The competing analyses of progressives are (i) one in which progressives are verbs and are always complements of a BE, as in (578a), and (ii) one in which there is a progressive form of ING, as in (578b).
(578) a. She is swimming.

b. She is swimming.

The former position, (i), may ultimately be preferable, but I’ll stick with the latter, (ii), for the time being since it’s simpler; it doesn’t entail the sort of complexities we’ll now launch into in a consideration of HAVEs with perfect xcomps.

Since these HAVEs have a verb as their xcomp, and since we have the rule that verbs are always xcomps of auxes, we conclude that these HAVEs are auxes. Moreover, since aux HAVEs can be xcomp of an aux, as in (571b–e) and arguably also (571f), we conclude that in these constructions the aux HAVEs are also verbs. But finite HAVEs can’t be xcomps of an aux (*She wants to has enjoyed herself), so it is only supine HAVEs and not finite HAVEs that are verbs as well as auxes.210 Thus, in She has gone, has is an aux but not a verb, while in She longs to have gone, have is both a verb and an aux.

In some, mainly British, dialects, possessive HAVEs are auxes, as (579a–c) show. In these lects, when the HAVE is an aux the structure is probably that shown in (579d).

(579) a. Has she any money?
    b. I’ve no money.
    c. I haven’t a clue.
    d. She has money.

210 For the sake of a quiet life I desist from debating the categorial status of the HAVE in Has she any money?
This construction is exceptional firstly in that the xcomp is a noun, and secondly in that the semantic relationship between the noun and its subject is one of possession.\(^{211}\) Thus it is not surprising that the construction is falling out of the language.

- **Passives.** The case for passives involving coenunciation has already been made in §5.4.2: there are good reasons for thinking passives involve a coenunciation of a verb and some other supine element similar to infinitival to. Since the morphology of passives is always the same as the morphology of perfects, it would be fairly explanatory to suppose that this morphological sameness is due to the passive supine elements taking a perfect as its xcomp. Since perfects are xcomps of have, we should therefore conclude that the supine element is an aux have. I will not now try to decide whether this passive aux have is also a verb.

**6.5 Rethinking Verb and Perfect**

So far, Verb has been treated as the WG equivalent of a primitive,\(^{212}\) and Perfect has been treated as a subclass of Verb. Both ideas may be incorrect.

It may prove possible to dispense with 'Verb' altogether, and we want to do so, as a way of explaining the exceptionlessness of the generalization that verbs are xcomps of auxes. One solution is to equate the class Verb with the class of xcomps of auxes. There are three problems. First, the xcomp of bes and possessional aux have is not a verb. But if, as is probably the case, the xcomps of bes and possessional aux have are xcomps of a type, such as 'statal', distinct from the type of the xcomp of other auxes, then rules that would otherwise have applied to 'Verb' could now apply to the class of words that are xcomp of an aux but not statal of an aux.\(^{213}\) A second, unsolved problem is oughts and useds, which are auxes whose xcomp is a to: generalizations ap-

---

\(^{211}\) Though I don't know of any other examples where the referent of the subject of a noun refers to a possessor of the referent of the noun, there are strong connections between the possessor relationship on the one hand, and 'er' and 'ee' relationships on the other, as (i–iv) show. In (i) Sophy is the looker, in (ii) Sophy is the kissee, in (iii) Sophy is the inspector, and in (iv) Sophy is the inspectee.

(i) Give Sophy a look.
(ii) Give Sophy a kiss.
(iii) Sophy's inspection (of the troops).
(iv) Sophy's inspection (by the troops).

I have no explanatory account for this, though Gisborne (in prep.) has given it some thought.

\(^{212}\) 'Primitive' is a descriptive rather than formal label. A primitive class has two criterial properties. First, it should bring together a cluster of properties rather than be defined by a single property. Second, it should be expected to have exceptional members.

\(^{213}\) With the redefinition of 'xcomp' as merely 'complement of W whose subject is W's object', the reformulation would in fact be that 'Verb' is the class of words (i) that are not statal of an aux, (ii) that are complement of an aux, (iii) whose subject merges with the object of the aux.
plying to non-statal xcomps of auxes would apply to these tos, yet they have the properties not of verbs but of all other tos. The third problem is that we can’t state the generalization (which has exceptions) that xcomps of auxes are verbs. What, then, will prevent a preposition from being an xcomp of an aux, as in She’ll in the room?

The problem with Perfect being subclass of Verb is that not all perfects are xcomps of auxes (e.g. in fallen leaves). When they’re not xcomps of auxes they look rather like adjectives. So we maintain the generalization that every verb is an xcomp of an aux by treating Perfect not as a subclass of Verb but as a class that intersects with Verb. But this leaves us with a problem: if not all perfects are xcomps of auxes, or, even, if not all perfects are verbs, however verbs are defined, how, for numerous lexemes, e.g. Tolerate, can we avoid the disjunctive statement that each instance of the lexeme is either (i) a non-statal xcomp of an aux (i.e. a verb) or (ii) a perfect? That disjunction would have to be repeated for any lexeme some of whose instances are, in the traditional sense, verbs.

The solution is to make all perfects xcomps of auxes. So far, the perfects in (580a–b) are xcomps of an aux, but not that in (580c). In (580d), though, which is a construction we might expect to be similar to (580c), there is a verb that is xcomp of an aux. So we’ll revise the analysis of (580c) to (581a). In this structure there is a coenunci­ate supine have that differs from the one in (580b) and (581b) in that it’s not an unac­tive.

(580) a. aux perfect
    TREE HAVE FALL
    trees have fallen

    unactive
    aux aux perfect
    TREE BE HAVE FELL
    trees were felled

b. perfect
    EVERY TREE FALL
    every tree fallen

c. aux
    EVERY TREE ING FALL
    every tree falling

d. of d of of
   s of of
   . of x of
   of of
   aux perfect
   EVERY TREE HAVE FALL

(581) a. every tree fallen

[309]
A consequence of this revised analysis is that the class Perfect can also be dispensed with. Perfect is now equivalent to the class of xcomps of \texttt{HAVE} auxes, the only problem being the noun xcomp of possessional aux \texttt{HAVES}. 
Postscript: Implications for syntactic theory

This postscript addresses first the implications for syntactic theory of allowing coenunciation in general (§7.1), and in particular coenunciation as it features in the analysis of node-raising (§7.2), the focus being on whether and to what extent the theory overall is being taken towards a direction of less constrainedness. Second, in §7.3 it addresses the relationship between the version of WG I’ve been developing and two hypothetical transformational versions of WG.

7.1 The possibility of coenunciation

Allowing coenunciation, rather than requiring that nothing be enunciation of more than one word, makes the grammar considerably less restrictive (in the sense that the mere quantity of morphophonological words (i.e. vocables) of an utterance does not set an upper limit on the number of syntactic words present), and accordingly in many ways makes the syntactician’s job harder, in that the range of possible analyses for a sentence is that much the greater when coenunciation is allowed as a possibility. However, a strong case has nonetheless been made for the possibility of coenunciation: it has been well-motivated, and put to good use, showing that some attractive analyses are made possible with coenunciation.

I have not proposed any constraints on coenunciation apart from No Crossing (the vitally important requirement that no word be separated from a word it coenounces with by a word that does not coenounce with them both). This is because no obvious constraint is suggested by the data, except maybe that all words permitted or required to coenounce are ‘function’/‘closed class’ words; but I’ve not formulated this apparently
viable and plausible constraint, because WG so far has no formal distinction between 'function'/'closed class' and 'content'/'open class'. One can easily think up additional constraints, e.g. that a specified dependency must hold between coenounced words, but such constraints conflict with the data — in this case, that poss-S coenounces with the preceding word whether or not there is any dependency between them. In such cases we have a choice between going with the data or going with the constraint, and we should go with the data, since the only factor in support of the constraint is precisely that it restricts the grammar. Essentially the same point applies to an appeal to constrainedness as an argument in favour of coenunciationlessness: the data is against it (as are purely conceptual considerations discussed in Chapter 5).

7.2 The relaxation of the No Crossing constraint for node-raising

It is hard to imagine how to proceed if No Crossing were dropped, but syntactic words were nonetheless sequenced. How on earth could one discern their sequence? Fortunately, in general the data supports No Crossing. But in the case of node-raising, the data does call for a — very tightly restricted — relaxation of No Crossing. (381), the statement of the relaxation and the restrictions upon it is repeated here for convenience as (582).

(582) If (i) A, B, C and D are sequences of words in the sequence A < B < C < D or D < C < B < A, and (ii) every word in C either (a) is s-dependent of another word in C or (b) is s-dependent of a word in D or (c) is unsubordinate of every word in C, then (α) the vocables corresponding to D can also correspond to B, and (β) associations between words in B and their enunciations are exceptionally allowed to cross associations between vocables and words not in B.

\[312\]

\[314\] It could be that poss-S must coenounce with an unsuperordinate of its inject, but it is not possible to test the validity of a constraint on poss-S, since the inject of a poss-S is tail of the poss-S, and thus every word between the poss-S and its inject must be subordinate of the poss-S. But the option remains to treat the apparent requirement that poss-S’s inject be its tail as a mere accidental consequence of a constraint that poss-S must coenounce with an unsuperordinate of its inject.
An alternative analysis without the need for such a relaxation certainly seems preferable, but none seems available; the analysis I have offered is thus a last resort. Even though it seems No Crossing must be relaxed, note the following.

- The restrictions on the relaxation are tight in (582).
- (582) is wholly unlike other rules of grammar. To formalize (582) in more orthodox WG style (that is, in the style of the other rules I give elsewhere) would necessitate the introduction of new machinery (e.g. word sequences).
- Node-raising is syntactically sui generis; there are no similar phenomena.
- Node-raising is stylistically marked. In speech it requires exaggerated and unusual intonation, though if node-raising occurs in speech one would suppose it to be scripted, for node-raising is almost exclusively found in writing, particularly in relatively formal and pedantic writing.

All this leads one to conclude that node-raising is a cuckoo in the grammatical nest: the nature of the written medium makes node-raising more processable, and creates a functional pressure for it — the avoidance of repetition, mainly, and secondarily the avoidance of anaphora and its attendant ambiguity. So convention (582) is adopted, not so much as a rule of grammar proper, but as a conventional and limited licence to be ungrammatical (a bit like en passant and castling in chess). Node-raising is the result of relatively conscious language engineering; I am suggesting that if the functional pressure for some grammatical phenomenon is strong enough, and apparent enough in relatively conscious use of language (as happens especially in written language), then it can override grammatical constraints. Node-raising is not the only example of such a phenomenon: other examples are the use of parenthetical bibliographical references in acad-

---

215 I am not aware of any analyses of node-raising that deal with the full range of data that I present in §5.3. Only RNR is generally dealt with, as a variety of coordination construction; that is only a subportion of what the phenomenon involves. I myself have been unable to think of an alternative, and less drastic, analysis of it.

216 The avoidance of ambiguity can be seen from comparing (i—iii). (i), which has the syntactic structure of (ii), is pragmatically less open to misassignment of referents.

(i) They sold coffee to _ and bought cars from the government.
(ii) They sold coffee to the government and bought cars from the government.
(iii) They sold coffee to the government and bought cars from it.

The avoidance of repetition can be seen from comparing the relative length of (iv—v). It's clearly a lot less effort to utter (v).

(iv) They sold coffee to the government of each of the provinces with a local GDP above the national average and bought cars from the government of each of the provinces with a local GDP above the national average.
(v) They sold coffee to _ and bought cars from the government of each of the provinces with a local GDP above the national average.
emic writing, and the use of parentheses around structurally integral words as in (583a–d).

(583) a. a (painting by) Picasso
    b. reactions to (allegations about) the genocide
    c. reactions to the (allegations about the) genocide
    d. reactions to (the allegations about) the genocide

The phenomenon illustrated here is, like node-raising, pragmatically-driven but syntactically constrained, in that the utterance must be syntactically well-formed both with and without the inclusion of the parenthesized material.

7.3 A comparison with transformational syntax
I'll here discuss (i) whether and how it is possible to develop a transformational counterpart of the version of WG I've been developing so far, and (ii) whether such a transformational version is superior or inferior to its nontransformational equivalent. I'll call the two WG versions 'TWG' ('T' for 'Transformational' and 'Trace theory') and 'SWG' ('S' for 'S-dependency').

I will take it that the main difference between TWG and SWG is that in TWG there are no dependencies that aren't s-dependencies, so the notions of s-dependency and dependency can be merged, with the resulting relation type having properties of both dependency and s-dependency. The questions we'll be asking are whether it is possible to develop a TWG that does this, and what such a TWG would be like, and whether it is worthwhile and as good as or better than SWG.

I'll be discussing two different hypothetical versions of TWG. Bear in mind that these are not TG recast in WG-like terms. Rather, they are versions of WG, differing from SWG only in the respects explicitly stated; I am in no way implicitly assuming that any of the other apparatus of Transformational Grammar is part of the TWG mod-

---

217 If I have neglected other possible versions of TWG, such as ones whose principal difference is something other than there being no dependencies that aren't s-dependencies, this may be because I haven't thought of them, or it may be because they are too inadequate to merit discussion here. Any version of TWG in which anything is reducible to purely configurational (i.e. constituency) factors is undiscussed for, I think, the latter reason.

218 It should be emphasized that in the TWG versions, not every dependent/s-dependent of W is branch-dependent (i.e. niece) of W. The s-dependency loops found with legates (in such constructions as coordination and pied piping — see Chapters 2–4) still remain.
My conclusions will be that it is certainly possible to develop a TWG, but that there is no advantage in doing so, since SWG is marginally simpler. In other words, the main body of the thesis motivates SWG principally in opposition to Hudson's version of WG, while this section shows its superiority to a transformational alternative. But it should be borne in mind that I consider the differences between SWG and Hudson's WG to be major and the differences between SWG and TWG to be minor, on the grounds that SWG and TWG can handle data that Hudson's WG, without any mechanism comparable to s-dependency, can't.

It appropriate to consider this matter here, firstly because the average linguist is more familiar with a transformational approach, and secondly because it might seem superficially as though, compared to classical WG (e.g. Hudson 1970), SWG already is moving in the direction of TWG, what with its introduction of constituency and morphologyless terminal nodes (i.e. some coenounced words), and its introduction of s-dependency to deal in part with phenomena (such as extraction, passive, extraposition, raising) that TG uses traces and other phonologically null items for. The question may arise as to whether SWG is already tantamount to TWG, and if it isn't, why we should not go the whole hog and espouse TWG. The following discussion explains what the whole hog might look like, and why SWG is preferable to it.

7.3.1 Anaphor-and-antecedent-and-agreement TWG

Though SWG doesn't allow enunciationless syntactic words, it does allow words that don't contribute any of their own morphology to their enunciation, so in effect phonologically null words are allowed. This then allows in principle structures like (584a–g), which are examples of what I'll call an anaphor-antecedent-agreement version of TWG (AAA TWG).

![Diagram](584)

(584) a. [[[[-1] Smiling] students[-1] waved].

---

219 A comparison of WG with TG might be of value as a method of introducing people schooled only in TG to WG, but it would be of dubious value as a way of trying to establish which is 'the better theory', since, due to their foundational assumptions — their different stances with respect to universalism — they amount to theories of different things. In principle one could for the purposes of comparison treat TG as a theory only of English, but this would be placing it at an undeserved disadvantage, since one would be asking it to do a job for which it was not tailored.

220 I mean what I here call 'anaphors' to be understood as anaphors, not as traces. Traces come into the picture in the TWG model discussed in §6.3.2.
b. [Anyone₁ [[-₁] smiling]] will be reprimanded.

c. [[It₁] keeps [[-₁] raining]].

d. [Make [it₁] [[-₁] rain]].

e. [[Doors₁] were [marked [on [the [top [of [the [left [of [-₁]]]]]]]]]].

f. [[Sordello₁], [he] will [[-₁] forget [[-₁] [she] tried [[-₁] to [[-₁] read [-₁]]]]]

The coindexation here at minimum indicates coreferentiality (in a loose sense of ‘semantically but not syntactically identical’), which pretty well by definition amounts to an associacy. So the question would be: Is this associacy a dependency? If the answer were No, then we would indeed have dispensed with the dependency/s-dependency distinction.

But there’s no need to address this question, since it is not sufficient for the coindexation to indicate mere coreferentiality. Consider (585–586). (586b) is ungrammatical because the xcomp of *become* is a preposition, while (585b) is grammatical because the xcomp of *become* is an adjective. Since (585a) is grammatical and (586a) isn’t, we are
led to conclude that the null anaphor in (585a) is an adjective and the null anaphor in (586a) is a preposition.

(585) a. How ill did she become \( _{-1} \)?
   b. She became ill.
(586) a. * In how poor a state of health did she become \( _{-1} \)?
   b. * She became in a poor state of health.

But in this case it must be that the null anaphor shares not only its meaning with its antecedent, but also its word class. In fact, it’s not just the word class that must be shared, but any property that can be selected for, such as a word’s lexeme, as (587) shows.

(587) On whom does she rely \( _{-1} \)?

Here, the null anaphor must belong to the lexeme ON.

Similarly, the Italian examples (588—589)\(^{221}\) show that null anaphors must agree with their antecedent in lexeme-specific, semantically unpredictable gender.

\[
\begin{array}{c}
\text{s} \\
\text{fill} \\
\text{x}
\end{array}
\]

(588) a. La pera\(_{1}\) è \( _{-1} \) caduta.
   ["The pear [fem.] is fallen [fem.]."]
   
\[
\begin{array}{c}
\text{s} \\
\text{fill} \\
\text{x}
\end{array}
\]

b. Il figo\(_{1}\) è \( _{-1} \) caduto.
   ["The fig [masc.] is fallen [masc.]."]

\[
\begin{array}{c}
\text{x} \\
\text{fette} \\
\text{s} \\
\text{x}
\end{array}
\]

(589) a. Ho fatto la pera\(_{1}\) \( _{-1} \) purpurea.
   ["I made the pear [fem.] purple [fem.]."]

\(^{221}\) The article—noun sequences are treated as single words to avoid a decision about which is subordinate to which.
In summary, a null anaphor must agree with its antecedent not only in meaning but also in gender and more generally in any property that is selectable for. There are two objections to this. First, it is a strength of the version of WG that I’ve been using that the grammatical mechanism of agreement is unnecessary for English. I mean ‘agreement’ here not as a name for a certain construction (as in “subject-verb agreement”) but as denoting the device whereby for a given feature (e.g. Number) two words have the same value (e.g. Singular). A feature in this sense is a relationship between a word and a category it belongs to, just as the relation Species links living beings to the species they belong to; for example, the relation Number relates singular words to the category Singular, and the relation Species relates me to the category Homo Sapiens. To repeat, then, it is possible to do without relations such as these, and thus to do without the rules such relations make possible. An agreementless grammar is more restrictive than one allowing agreement.

The second objection to agreement between null anaphor and antecedent is that for cases such as passive, raising and tough-movement, where a word must satisfy selectional requirements as a subject of one word and as a complement of another, we must set up a feature ‘Property-that-is-selectable-for’, a feature for which a word can have sever­al values. Admittedly, WG still lacks even a way of stipulating which properties are and aren’t selectable for, but one can hardly find in such barefaced stipulation a satisfying solution.

Notwithstanding the above objections to a form of TWG employing anaphora and agreement, it could still work. But, seductive though it is by virtue of such factors as easy diagrammability, it on balance yields a more complex and less restricted grammar than SWG. Here’s a comparison between SWG and AAA TWG.

---

222 I also wonder whether it is necessary for other languages, or whether what is called agreement simply amounts to the morphophonological blueprint of the enunciation of word X being constrained by properties of another word that X is in a certain kind of dependency relationship with. I have not been able to think of a hypothetical example for which this reanalysis would necessarily be unworkable. (It should be borne in mind that, as can be seen from the analysis of pied piping in Chapter 2, the Proxy associacy does the job of what in other frameworks is done by feature percolation.)
• There is one respect in which AAA TWG is simpler, in the sense of requiring fewer rules: since AAA TWG does not require s-dependency, only SWG needs to stipulate that all s-dependents of W are dependents of W.

• A small minority of rules, such as (590a–b), are simpler in AAA TWG.

(590) a. SWG: If X is s-dependent of Y and X is subject of Y then X is precessor of Y.
b. AAA TWG: If X is subject of Y then X is precessor of Y.

• A minority of rules come out simpler in SWG, e.g.:

(591) a. SWG: “The subject of the xcomp of W is the object of W”.
b. AAA TWG: “The subject of the xcomp of W is a null-anaphor whose antecedent is object of W”.

• Most rules are about equally complex in the two models:

(592) a. SWG: By default, if X is s-dependent of Y then X is successor of Y.
b. AAA TWG: By default, if X is dependent of Y then X is successor of Y.

(593) a. SWG: By default, if X is dependent of Y then X is s-dependent of Y.
b. AAA TWG: By default, every word is not a null-anaphor.

(594) a. SWG: “X must be s-dependent of Y if …”.
b. AAA TWG: “X can’t be a null-anaphor if X is dependent of Y and …”

(595) a. SWG: “X may be s-dependent of Y if …”.
b. AAA TWG: “X may be a null-anaphor if X is dependent of Y and …”

(596) a. SWG: “X can’t be s-dependent of Y if …”.
b. AAA TWG: “X must be a null-anaphor if X is dependent of Y and …”

• A lot of extra rules are required for AAA TWG:

• AAA TWG must stipulate that null anaphors can’t have dependents. (This is because ‘movement’ cannot leave behind stranded dependents of the moved word; rather, the head of a word is a word that has reached its final resting place.)
AAA TWG must stipulate that anaphors agree with their antecedent with respect to meaning and their values for the feature Property-selectable-for.

AAA TWG must stipulate that null-anaphors must have an antecedent in the same sentence. (This is because having a null anaphor without an antecedent would amount to mere deletion, which, if it is possible at all, is not generally possible for most elements that are nonetheless movable.)

AAA TWG might hope to make up for its disadvantages by incorporating the account of anaphor-antecedent associacies into a larger account of binding in general, and thus score points for its greater generality. But that seems the wrong way to go, since the way these null-anaphors link to their antecedents is both better understood, clearer in the data, and less subject to pragmatic factors than other anaphora is. Thus if anything, to the extent that binding is syntactic, it ought to be done in the same way as transformational phenomena such as raising and extraction, i.e. by chains of individually sanctioned local links, so that all long-distance anaphor-antecedent associacies are mediated by intermediate null-anaphors.\(^{223}\)

### 7.3.2 Chain-and-segment TWG

There is an alternative TWG to AAA TWG that makes no use of coenunciation and instead divides the single primitive 'syntactic word' into two: this I'll call 'chain-and-segment TWG' (CS TWG). Instead of having words, we have chains and their segments.\(^{224}\) Some properties of words properly belong to chains, and others to segments:\(^{225}\)

- Dependencies hold between segments.\(^{226}\)
- Segments but not chains can be daughters of phrases.

---

\(^{223}\) Remember that in AAA TWG null anaphors are doing the job of traces in TG. Hence 'intermediate null anaphors' are analogous to traces intermediate in a chain.

\(^{224}\) 'Segment' is meant here merely as 'link in a chain' and involves no allusion to any other technical sense in which the term is used. 'Chain' is meant in the Transformational Grammar sense.

\(^{225}\) I have considered alternatives to this pattern of allocation of properties, and the one I give here presented the fewest problems. To fully justify my tentative conclusions in §7.3.3 that SWG is mildly superior to the TWG versions, I should really first explain why this particular version of CS TWG is the best, but this would turn a brief postscript into a far longer disquisition on the relative merits of versions of WG none of which are actually advocated in the preceding chapters.

\(^{226}\) If it was chains rather than segments that were heads and dependencies, then there would be no motivation for segments. We would have a model in which words can last a long time (e.g. beginning at the start of a sentence and ending at the end (e.g. in the case of extraction from one end of the sentence to another), and we'd then be faced with the problem of stating the position of the chain's enunciation relative to other enunciations, and with the problem of accounting for the manifold evidence (some considered in Chapter 5) that words are sequenced.
• Segments but not chains can have enunciations.227
• Word classes classify chains. Segments select for segments of a given type of chain.
• Chains rather than segments have meanings.
• Chains (or chains’ meanings) rather than segments bear semantic roles. For example, if segment X is object of a kill, then the chain (or its meaning) that X is a segment of is the killee.

The diagrams in (584a–g) above will also serve as representations of CS TWG. Each terminal node is a segment. Coindexed segments belong to the same chain. Uncoindexed segments belong to single-segment chains. Segments without enunciations are represented by ‘_’, and those with enunciations are represented orthographically.

The CS TWG counterpart of SWG’s word that is an s-dependent is a word that has an enunciation. It takes a few rules to stipulate which segment in a chain gets enounced:

(597) a. Exactly one segment in each chain has an enunciation.
   b. If X is dependent of Y, and Z is subordinate of Y, and Z is not dependent of Y, and X and Z are segments of the same chain, then Z has no enunciation.
   c. If X and Z are dependents of Y, and X and Z are segments of the same chain, and Z follows Y, then Z has no enunciation.

Given (597a), (597b) accounts for which segment in the chain coindexed 1 in (598a) is enounced, and similarly for the subject-to-object raising in (598c–d), and (597c) accounts for the object-to-subject promotion in (598b–c) and the extraction in (598c–d).

(598) a. put it_here

---

227 If it’s chains that have enunciations then we have the problem of determining at which point the enunciation happens. Also, the segment that is simultaneous with the enunciation is the only one that can have dependents (since dependents can’t be left stranded without an overt s-head). In the CS TWG it’s the property of having an enunciation that is the equivalent of SWG’s property of being an s-dependent.
b. \[ I_1 \ am \ _1 \]

c. \[ \_1 \ Here_2 \ I_1 \ am \ _1 \_2 \]

d. \[ \_1 \ Here_2 \ am \ I_1 \_2 \]

(597c) can be dispensed with if (598b–d) instead have the structure of (599a–c).

(599) a. \[ I \ am. \]

b. \[ \_1 \ Here \ I_1 \ am. \]

c. \[ \_1 \ Here \ am \ I_1. \]

I have no suggestions for any explanation for (597b–c) (or just (597b)) or for why they are bound up with the syntax–morphophonology interface. Since I don’t advocate the model under discussion, my lack of suggestions should not be found at all troubling.

As with the comparison of SWG with AAA TWG, a comparison of SWG with CS TWG shows SWG to be simpler. As before, the main respect in which SWG is more

---

228 Since I don’t advocate the model under discussion, my lack of suggestions should not be found at all troubling.

229 In SWG the matter of what a word is s-dependent of (e.g. whether it can extract, or where it must come to rest) is a purely syntactic one. In CS TWG, this is translated into the matter of which word has an enunciation. In Minimalism, overt movement is presented as motivated by a need to shed features that are not phonologically interpretable. Hence there is in the architecture of the model an inherent connexion between overt movement and enunciation.
complex is that it requires a distinction between dependency and s-dependency. The additional rules CS TWG needs are (597a–c) and (600). (600) is the CS TWG counterpart of AAA TWG’s stipulation that null anaphors can’t have dependents.

(600) If a segment has a dependent then it must have an enunciation.

As with AAA TWG, a small minority of rules are simpler in CS TWG, e.g.:

(601) a. SWG: If X is s-dependent of Y and X is subject of Y then X is precessor of Y.
    b. CS TWG: If X is subject of Y then X is precessor of Y

A minority of rules come out simpler in SWG, e.g.:

(602) a. SWG: “The subject of the xcomp of W is the object of W”.
    b. CS TWG: “The subject of the xcomp of W is a segment of the same chain that the object of W is a segment of”.

Most rules are about equally complex in the two models:

(603) a. SWG: By default, if X is s-dependent of Y then X is successor of Y.
    b. CS TWG: By default, if X is dependent of Y then X is successor of Y.

(604) a. SWG: By default, if X is dependent of Y then X is s-dependent of Y.
    b. CS TWG: By default, every segment has an enunciation.
    or: By default, every chain has exactly one segment.

(605) a. SWG: “X must be s-dependent of Y if …”.
    b. CS TWG: “X must have an enunciation if X is dependent of Y and …”

(606) a. SWG: “X may be s-dependent of Y if …”.
    b. CS TWG: “X may have an enunciation if X is dependent of Y and …”

(607) a. SWG: “X can’t be s-dependent of Y if …”.
    b. CS TWG: “X can’t have an enunciation if X is dependent of Y and …”
### 7.3.3 SWG, AAA TWG and CS TWG: a provisional final reckoning

I will begin with a summary of the essential differences between these three models. (Item (I) on the list is not a difference, but a reminder of a non-difference.)

<table>
<thead>
<tr>
<th></th>
<th>SWG</th>
<th>AAA TWG</th>
<th>CS TWG</th>
</tr>
</thead>
<tbody>
<tr>
<td>(I)</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>(II)</td>
<td>no</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>(III)</td>
<td>yes</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>(IV)</td>
<td>no</td>
<td>yes$^{230}$</td>
<td>n.c</td>
</tr>
<tr>
<td>(V)</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
</tbody>
</table>

We see from rows (III—V) that each model requires the same number of novel devices, so at the most important level of comparison the three models come out tied. Let us consider further some other relative merits and defects of the models.

On the grounds of ontological parsimony (§1.5), CS TWG is to be dispreferred. SWG is preferable, for the novel device it introduces is not a new sort: it just another relation among many others. But CS TWG requires the recognition of sets of words (i.e. sets of segments), which is a sort of object not required in the other models. As for AAA TWG, we would require some kind of novel sort to handle agreement,$^{231}$ but I return below to the issue of whether a TWG based on coenouncing null anaphors necessarily involves agreement.

If we compare the three models on the basis of the complexity of rules required for them, SWG comes out as superior, at least in those respects in which the three models handle the phenomena SWG handles. SWG needs fewer rules and those that it does need are on average simpler. In comparison with CS TWG, SWG does not need to explain (597a—c) (or just (597a—b)), since such rules are entirely absent from SWG.

---

$^{230}$ I said in §7.3.1 that AAA TWG requires agreement, but see below for a reassessment.

$^{231}$ For example, agreement could involve recognizing classes of classes of words. So, for instance, Singular and Plural would be classes of words, and Number would be the class containing Singular and Plural. Classes of classes of words are not otherwise required.
A further reason for preferring SWG is that it does not have to explain a curious state of affairs arising only in TWG, whereby the rule that multiple non-adjunct dependents of the same type must be coordinated applies only to enounced segments (CS TWG) or words that aren't null anaphors (AAA TWG), as (608) shows.

(608) bring either him₁ or he₂ here

A reason for favouring CS TWG is that it provides a slightly less stipulative basis for distinguishing properties that are selectable for from properties that aren't. I'll sketch how it might go. As stated above, word classes classify chains, and segments select for segments of a given type of chain. Only segments have dependents. No segment is ever selected for on the basis of what dependents it has. Therefore we might conjecture that properties that are selectable belong to chains, and properties that aren't selectable for belong to segments. For example, no word/segment is selected for on the basic of which open-class lexeme it belongs to, so it could be that open-class lexemes classify segments rather than chains.

Before we proceed further with the comparison, let me reconsider AAA TWG. A defect of AAA TWG is that it is apparently coincidental that all null-anaphors co-enounce leaving no morphological trace (except in the case of resumptive pronouns?), while in CS TWG enunciationlessness is a criterial rather than incidental property in distinguishing the two types of segments. This objection could be remedied by instead stating that null anaphors have no enunciation at all. This would introduce into the theory a new distinction between words that do have enunciations and words that don't, but at least this move would allow us to define a null anaphor as any word without an enunciation.

I said in §7.3.1 that AAA TWG requires a mechanism for agreement, which is undesirable both because it introduces an otherwise unnecessary formal device, and because that device involves an ontologically novel sort. However, there is in fact a way to revise AAA TWG so that agreement is got rid of in favour of an associacy 'Antecedent', such that only null anaphors have antecedents. This can be done by means of the rules that state whether a word’s dependent can, can’t or must be its s-dependent. Con-
sider some hypothetical rules of that kind, presented first in SWG terms, and then reformulated in an agreementless AAA TWG, which I will now start referring to as 'AA TWG' (i.e. Anaphor-and-Antecedent instead of Anaphor-and-Antecedent-and-Agreement). The AA TWG reformulations assume that null anaphors belong to no word class or lexeme.

If X is complement of Y and Y is instance of Z [a hypothetical word class or lexeme] then:

(i) SWG: X is s-dependent of Y and X is instance of Noun.
   AA TWG: X is instance of Noun.

(ii) SWG: X is not s-dependent of Y and X is instance of Noun.
    AA TWG: the antecedent of X is instance of Noun.

(iii) SWG: X is instance of Noun (and X may be, but needn't be, be s-dependent of Y)
    AA TWG: X or the antecedent of X is instance of Noun.

Comparing, in the tabulation below, AA TWG with CS TWG and SWG using the same criteria as before plus some others brought up in the discussion above, leads me to the conclusion that while CS TWG is inferior to the other two, there's not a lot to choose between SWG and AA TWG.

---

232 The resulting structures are reminiscent of the way Lexicase handles 'unbounded dependency constructions' by means of what it calls 'traces'.

[326]
| (I)  | Does the model allow coenunciation? | yes | yes | yes |
| (II) | Does coenunciation play a role in accounting for 'movement constructions'? | no | yes | no |
| (III) | Is there a distinction between dependency and s-dependency? | yes | no | no |
| (IV)  | Is there a need for agreement (i.e. sharing of attribute values)? | no | no | no |
| (V)   | Is there a distinction between chain and segment? | no | no | yes |
| (VI)  | Is there a distinction between words with and without enunciations? | no | yes/no | yes\(^{233}\) |
| (VII) | Are ontologically novel sorts required? | no | no | yes |

Nonetheless, §7.3 has shown that the formal differences between SWG and TWG are slight, with each translating into the other unambiguously. The important lesson is that we need either SWG or TWG instead of the WG version of Hudson (1990), and the choice between SWG and TWG — or at least between SWG and AA TWG — is a relatively minor one, little more than a matter of taste.

\(^{233}\) There is a distinction between segments with and without enunciations.


Bresnan, J. (1974). 'The position of certain clause particles in phrase structure.'
Cambridge, Massachusetts: MIT Press.
Bresnan, J., & Grimshaw, J. (1978). 'The syntax of free relatives in English.'


[331]
   reply to Emonds’ Natural Language and Linguistic Theory 11: 279–312.
   perspective.’ ms.
   Dordrecht: Reidel.
Jaworska, E. (1986). ‘Prepositional phrases as subjects and objects.’ Journal of
   Copenhagen: Munksgaard.
   Blackwell.
   Chicago: Chicago University Press.


McCawley, J.D. (1993). *Everything that linguists have always wanted to know about logic (but were ashamed to ask)*. Second edition. Chicago: University of Chicago Press.


Rosta, A. (in prep.) ‘English complementation patterns.’ “ms”


Glossary

Advener of. X is advener of Y iff X is legate of Z and Z is s-dependent of Y.

Associacy. A relationship where one word is associate of another.

Associate of. Any word in a grammatically significant syntagmatic relationship with another.

Aunt of. Converse of Niece.

Aux [word class]. Finite and non-finite auxiliaries.

Branch-depency. A relationship where one word is branch-dependent of another.

Branch-dependent of. X is branch-dependent of Y iff X is s-dependent of Y and X is niece of Y.

Branch-head of. Converse of Branch-dependent.

Clausal of. A variety of dependent, and generally a variety of complement.

Conjunct-mate of. A reciprocal relationship holding between roots of the same conjunct (in complex coordination).

Coordinatee of. Roots of conjuncts are coordinatee of the conjunction.

Dependency. A relationship where one word is dependent of another.

Dependent. See §2.1 for a definition.

Elliptic [word class]. An aux with no xcomp.

Esmerge [word class]. A word extractee object is the same word as its subject.

Exclusive. A dialectal/idiiolectal kind of gapping conjunction that does not allow the initial conjunct to be gapless (formally: every coordinatee must be a substitute or a conjunct-mate of a substitute).

Extractee of. A relation involved in extraction; equivalent to Hudson (1990)'s Visitor.


Gapping conjunction [word class]. The kind of conjunction you get in gapping.

Head of. Converse of Dependent.

Inclusive. A dialectal/idiiolectal kind of gapping conjunction that allows the initial conjunct to be gapless (formally: the initial coordinatee can have a legate rather than being a substitute or a conjunct-mate of a substitute).
Inject of. Indirect object.

Legate of. A legate of W is a proxy of a conjunction (in which case W is coordinatee of the conjunction) or a wh-pronoun (in which case W is the wh-pronoun itself). Only legates can be adveners and have wards.

Niece of. X is niece of Y in constituency structures [[X] Y] and [Y [X]].

Osmerge [word class]. A word whose object is the same word as its subject.

Parenthetical of. A variety of adjunct.

Precessor of. X is precessor of Y iff X is s-dependent of Y and X precedes Y.

Promotee of. A relation used in some promotions, e.g. prepositional passives.

Proxy of. See §2.6.

Quantitative of. A kind of dependent, sometimes a complement, that expresses quantities and amounts.

S-dependency. A relationship where one word is s-dependent of another.

S-dependent of. See §2.1–4. The main characteristic is that if X is s-dependent of Y then X takes its position from Y.

S-head of. Converse of s-dependent.

Subclass of. X is subclass of Y iff every member of X is member of Y.

Subjunction [word class]. A supplementary conjunction whose coordinatee is not a substitute; i.e. a conjunction that does no coordinating.

Subordinate of. X is subordinate of Y iff X is s-dependent of Y or X is s-dependent of Z and Z is subordinate of Y.

Substitute of. If X is substitute of Y then X and Y are in the same kind of dependency relationships with Z.

Subtype of. X is subtype of Y iff X is subclass of Y and every member of Y is a member of a subtype of Y.

Successor of. X is successor of Y iff X is s-dependent of Y and X follows Y.

Superordinate of. X is superordinate of Y iff Y is s-dependent of X or Z is s-dependent of X and Z is superordinate of Y.

Supervener of. = Legate.

Supine [word class]. Non-finite auxiliaries.

Supplementary conjunction [word class]. A unary conjunction with only one coordinatee.
Symmetric conjunction [word class]. Normal conjunction with no funny business going on.

Tail of. X is tail of Y if X is s-dependent of Y and every word between X and Y is a superordinate or a subordinate of X.

Unactive [word class]. Passive and mediopassive.

Unary conjunction [word class]. A gapping conjunction that can have only one conjunct.

Vocable. Morphological word.

Ward of. X is ward of Y iff Y is legate of Z and X is s-dependent of Z.

Xcomp of. Predicative complement. The subject of the xcomp of W is object of W.