Chapter 7

Gapping and the Order of Constituents

'O where are you going?'
Said reader to rider,
'That valley is fatal where furnaces burn,
Yonder's the midden whose odours will madden,
That gap is the grave where the tall return.'

W. H. Auden, *Five Songs, V*

The Dutch pattern of argument cluster coordination in subordinate clauses—briefly introduced in section 6.4.1 and discussed in greater depth in section 7.1—is a case of the more general universal identified by Ross (1970) noted in chapter 2, concerning the tendency of argument cluster coordination to conserve or "project" the directionality of the lexicon across SOV, VSO, and SVO languages and/or constructions (see Koutsoudas 1971; Lehmann 1978; and Mallinson and Blake 1981):

(1) a. SOV: *SOV and SO, SO and SOV
b. VSO: VSO and SO, *SO and VSO
c. SVO: SVO and SO, *SO and SVO

The SOV and VSO cases are essentially symmetrical, and are discussed in sections 7.1 and 7.2. Certain cases of Dutch main-clause argument cluster coordination fall under the VSO heading, as the choice of a VSO category for main-clause verbs predicts (see chapter 6). These are also discussed in section 7.2.

The remainder of the chapter explains gapping in SVO languages like English, as in *Dexter ate bread and Warren, potatoes*, and the related cases of forward gapping in Dutch and other Germanic languages, in terms of the combinatory theory. In particular, the theory predicts Ross's generalization that verb-medial languages and constructions necessarily pattern with the verb-initial ones rather than the verb-final ones in permitting forward, but not backward, gapping, as in (1c) above, and explains why certain SOV languages like Dutch and certain VSO languages like Zapotec show exceptions to the above pattern.[1]

Although the basic SOV and VSO cases reduce to argument cluster coordination, and (as Maling (1972) pointed out in different terms) do not require a

distinctive rule of gapping as such, I will argue that SVO gapping is also, in a sense, argument cluster coordination. For that reason, I will continue to refer informally to this whole collection of phenomena as "gapping."

7.1 Gapping and SOV Word Order

As we saw, type-raising arguments over an SOV verb, composing them, and then conjoining the resulting nonstandard constituents permits the "backward gapping" construction characteristic of coordinate clauses in SOV languages. Thus, in Japanese a subject and an object NP can not only combine with the verb by forward application, but also forward-compose as follows, via forward type-raising over the following SOV verb:

(2) $\text{tazuneta}_{SOV} := (S\backslash NP_{nom})\backslash NP_{acc} : \lambda x.\lambda y.visit' xy$

(3) a. Ken-ga Naomi-o tazuneta.
Ken-NOM Naomi-ACC visit-PAST.CONCL
'Ken visited Naomi.'

b.
$$\dfrac{\dfrac{\text{Ken-ga}}{S/(S\backslash NP_{nom})}\,{>}\mathbf{T} \quad \dfrac{\dfrac{\text{Naomi-o}}{(S\backslash NP_{nom})/((S\backslash NP_{nom})\backslash NP_{acc})}\,{>}\mathbf{T} \quad \dfrac{\text{tazuneta}}{(S\backslash NP_{nom})\backslash NP_{acc}}}{S/((S\backslash NP_{nom})\backslash NP_{acc})}\,{>}\mathbf{B}}{S}\,{>}$$

The resulting nonstandard constituent *Ken-ga Naomi-o* can therefore conjoin:

(4) [Ken-ga Naomi-o], [Erika-ga Sara-o] tazuneta.
$S/((S\backslash NP_{nom})\backslash NP_{acc})$ $S/((S\backslash NP_{nom})\backslash NP_{acc})$ $(S\backslash NP_{nom})\backslash NP_{acc}$
Ken-NOM Naomi-ACC Erika-NOM Sara-ACC visit-PAST.CONCL
'Ken visited Naomi, and Erika, Sara.'

What is more, the Principles of Adjacency, Consistency, and Inheritance, together with the order-preserving constraint on type-raising that is the sine qua non of an order-dependent language, again limit the possible constituent orders. They do not permit any raised categories or rules of composition that would produce a *leftward*-looking function, so that the corresponding "forward gapping" construction is disallowed on the SOV lexicon:[2]

(5) *Ken-ga Naomi-o tazunete, Erika-ga Sara-o
Ken-NOM Naomi-ACC visit-PAST.ADV Erika-NOM Sara-ACC
'Ken visited Naomi, and Erika, Sara.'

As noted earlier, this asymmetry tends to be characteristic of strictly SOV languages. However, a number of important qualifications to the generalization have to be made. Most importantly, like other Germanic languages, Dutch *does* allow coordinations on the above pattern in subordinate-clause conjunctions:

(6) ... dat Maaike aardappels eet en Piet bonen
... that Maaike potatoes eats and Piet beans
'... that Maaike eats potatoes and Piet beans.'

We will see that this exception to the SOV pattern is related to the fact that these languages possess an SVO main-clause constituent order as well, which Japanese lacks.

Japanese also allows OSV word order, as in (7):

(7) Naomi-o Ken-ga tazuneta.
Naomi-ACC Ken-NOM visit-PAST.CONCL
'Ken visited Naomi.'

However, the temptation to allow this by introducing forward crossed composition, so that the raised Japanese subject could compose with the Japanese verb, as in the following derivation, should be resisted:

(8)
$$\dfrac{\dfrac{\text{Naomi-o}}{(S\backslash NP_{nom})/((S\backslash NP_{nom})\backslash NP_{acc})}\,{>}\mathbf{T} \quad \dfrac{\dfrac{\text{Ken-ga}}{S/(S\backslash NP_{nom})}\,{>}\mathbf{T} \quad \dfrac{\text{tazuneta}}{(S\backslash NP_{nom})\backslash NP_{acc}}}{S\backslash NP_{acc}}\,{*>}\mathbf{B}_{\times}}{S}\,{>}$$

Such crossed composition, besides introducing a whole new combinatory rule schema into the grammar, would immediately need to be heavily constrained if it were not (for reasons familiar from chapter 4) to give rise to very free word order indeed, including non-clause-bounded scrambling. Moreover, such an analysis would fail to account for the fact that OS order can also give rise to constituent cluster coordination parallel to (4), as in (9):

(9) [Naomi-o Ken-ga,] [Sara-o Erika-ga] tazuneta
$S/((S\backslash NP_{acc})\backslash NP_{nom})$ $S/((S\backslash NP_{acc})\backslash NP_{nom})$ $(S\backslash NP_{acc})\backslash NP_{nom}$
Naomi-ACC Ken-NOM, Sara-ACC Erika-NOM visit-PAST.CONCL
'Ken visited Naomi and Erika, Sara.'

Although the generalized type-raised categories can compose in OS order to yield such OS argument clusters, the order-preserving nature of the rules concerned forces them to deliver a cluster demanding a different OSV category

for the verb, as the agreement features in (9) reveal, rather than the SOV verb standardly assumed for Japanese.

It is therefore tempting under present assumptions to regard so-called scrambling constituent orders in Japanese as lexically specified, either via multiple verb categories or via explicitly unordered leftward categories, a move that is in keeping with the observation that scrambling (as distinct from true extraction) is clause-bounded. However, I will not pursue the question of OS order in Japanese any further here.[3] Under the analysis presented in the chapter 6, Dutch subordinate clauses are predicted to exhibit the SOV pattern (1a). For example, the entire preverbal argument cluster can coordinate:

(10) a. ... dat [Jan de kinderen en Marie de nijlpaarden] zag zwemmen.
... that [Jan the children and Marie the hippos] saw swim
'... that Jan saw the children swim and Mary saw the hippos swim.'

b. ... dat [Jan Marie en Cecilia Henk] de kinderen zag helpen
... that [Jan Marie and Cecilia Henk] the children saw help
zwemmen.
swim
'... that Jan saw Marie and Cecilia saw Henk help the children swim.'

c. ... dat [Jan Marie de kinderen en Henk Cecilia de nijlpaarden] zag
... that [Jan Marie the children and Henk Cecilia the hippos] saw
helpen zwemmen.
help swim
'... that Jan saw Marie help the children swim and Henk saw Cecilia help the hippos swim.'

Example (10b) is related to the following "backward gapping" example, due to van Oirsouw (1982, 555, example (8b)), apart from the fact that in the latter case the verbs are in the "German" order, as is common in standard Dutch with the auxiliary *hebben*:

(11) Ik geloof dat Jan *Syntactic Structures* en Piet *Aspects* gelezen heeft.
I believe that Jan *Syntactic Structures* and Piet *Aspects* read has
'I believe that Jan has read *Syntactic Structures* and Piet *Aspects*.'

The grammar also allows contiguous preverbal argument sequences that do not include the subject to compose, so the following examples are also allowed, on the assumption that the dative PP *aan Henk* has the category $T/(T\backslash PP_{DAT})$ and that one category of the stem *geef-* 'give' is $(VP\backslash NP)\backslash PP_{DAT}$:

(12) ... dat Jan de kinderen [een treintje aan Piet en een pop aan Henk] zag
... that Jan the children [a train to Piet and a doll to Henk] saw
geven.
give.
'... that Jan saw the children give a train to Piet and a doll to Henk.'

(13) ... dat Jan [de meisjes een treintje aan Piet en de jongens een pop aan
... that Jan [the girls a train to Piet and the boys a doll to
Henk] zag geven.
Henk] saw give
'... that Jan saw the girls give a train to Piet and the boys give a doll to
Henk.'

(14) ... dat Jan [de meisjes een treintje en de jongens een pop] aan Henk
... that Jan [the girls a train and the boys a doll] to Henk
zag geven.
saw give
'... that Jan saw the girls give a train and the boys give doll to Henk.'

Sentence (12) is completely acceptable (cf. Bresnan et al. 1982, 619). The grammaticality of a sentence parallel to the second of these is questioned by by Bresnan et al. in the course of justifying a rather different account of the NP sequence. It seems to be accepted by some informants. The third is not discussed by Bresnan et al., but seems to be also accepted.

Similarly since the complete verb sequence can combine by forward application with preverbal NPs in the crossed order, any subsequence that includes all of the verbs and some rightmost subsequence of the preverbal NP sequence can be a constituent, and may also conjoin. Examples are:

(15) a. ... dat ik Henk [de nijlpaarden zag voeren en de olifanten hoorde
... that I Henk the hippos saw feed and the elephants heard
wassen.]
wash]
'... that I saw Henk feed the hippos and heard him wash the elephants.'

b. ... dat ik [Cecilia de nijlpaarden zag voeren en Henk de olifanten
... that I Cecilia the hippos saw feed and Henk the elephants
hoorde wassen.]
heard wash]
'... that I saw Cecilia feed the hippos and heard Henk wash the elephants.'

Unless such conjuncts include the entire verb group, the combination with the NPs will not be possible with the rules as set out above. It follows that sentences like the following are excluded by the present grammar:

(16) a. ?...dat ik Cecilia [de nijlpaarden zag en de olifanten hoorde]
...that I Cecilia [the hippos saw and the elephants heard]
wassen.
wash
'...that I saw Cecilia wash the hippos and heard her wash the elephants.'

b. ?...dat ik [Cecilia de nijlpaarden zag en Henk de olifanten hoorde]
...that I Cecilia the hippos saw and Henk the elephants heard
wassen.
wash
'...that I saw Cecilia wash the hippos and heard Henk wash the elephants'

Some informants seem prepared to tolerate such examples. Whatever their status, it is striking that so much freedom is allowed in Dutch, and that it can be accounted for within the same degrees of freedom that are required for English grammar.

7.2 Gapping and VSO Word Order

As Dowty (1988) was the first to point out, the position is reversed for verb-initial languages such as Irish. Again, subject and object can raise and compose with each other and with adjuncts in an order-preserving way to yield a single function over a transitive verb like *chonaic* 'saw', this time via leftward type-raising and composition; and again, the nonstandard constituent can coordinate:[4]

(17) $\text{chonaic}_{VSO} := (S/NP)/NP : \lambda x.\lambda y.see'yx$

(18) a. Chonaic Eoghan Siobhán.
saw Eoghan Siobhán
'Eoghan saw Siobhán.'

b.
$$\dfrac{\dfrac{\overline{(S/NP)/NP}\quad \dfrac{\overline{(S/NP)\backslash((S/NP)/NP)}^{<\mathbf{T}}\quad \dfrac{\overline{S\backslash(S/NP)}^{<\mathbf{T}}}{S\backslash((S/NP)/NP)}^{<\mathbf{B}}}{S}^{<}}{}$$
Chonaic Eoghan Siobhán

(19) Chonaic [Eoghan Siobhán] agus [Eoghnaí Ciarán].
$(S/NP)/NP$ $S\backslash((S/NP)/NP)$ $S\backslash((S/NP)/NP)$
saw Eoghan Siobhán and Eoghnaí Ciarán
'Eoghan saw Siobhán, and Eoghnaí, Ciarán.'

Again the three principles exclude the "backward gapping" construction that Ross (1970) held to be generally disallowed in strictly verb-initial languages:

(20) *[Eoghan Siobhán] agus chonaic [Eoghnaí Ciarán].
$S\backslash((S/NP)/NP)$ $(S/NP)/NP$ $S\backslash((S/NP)/NP)$
Eoghan Siobhán and saw Eoghnaí Ciarán

As in the case of Dutch and SOV, there are exceptions to Ross's generalization for VSO languages. Later we will examine the case of Zapotec (Rosenbaum 1977), a VSO language that allows backward gapping and that like Dutch, will turn out to do so because it allows other constituent orders in main clauses.

The derivation and the category for the VSO transitive verb assume that the subject—that is, the NP corresponding to the least oblique element x at predicate-argument structure—is the *first* argument of the VSO verb, not the last, as in the Germanic languages. As a result, the object commands the subject in a purely applicative context-free derivation. This assumption is an instance of the generalization noted at the end of chapter 4 and is forced by the present theory. However, objects may *always* c-command subjects in CCG derivations, even in SVO languages. And the separation of predicate-argument structure from derivations and surface categories (which is crucial to the present analysis of binding in English) allows us to capture the fact that binding phenomena in VSO languages and constructions strongly parallel those in SVO languages and all others with respect to the obliqueness hierarchy (see Keenan 1988).[5]

VOS word orders do not arise in general in Irish. See Baldridge 1999 for an argument parallel to that given earlier for Japanese OSV orders to the effect that VSO languages that do allow VOS orders (such as Tagalog) do so via base generation.

Many coordination phenomena in Dutch main clauses conform to the pattern of VSO languages and are predicted by the fragment presented in chapter 6 and the assumption that main-clause verbs have VSO categories. Such phenomena include examples like the following, parallel to English argument cluster coordination and VSO complement cluster coordination (van Oirsouw 1987, 58):

(21) Wil$_{S_{-SUB}/NP/NP}$ [jij een ijsje en Marietje limonade?]$_{T\backslash(T/NP)/NP}$
want you an ice-cream and Marietje lemonade
'Do you want an ice-cream, and Marietje lemonade?'

This example is simply the mirror-image of the "backward gapping" subordinate clauses discussed in section 7.1. Because of the involvement of the backward crossed composition rule (79) of chapter 6, VP coordination is also captured:

(22) a. Hij [at aardappels en dronk bier.]$_{S/NP_{+ANT}}$
he ate potatoes and drank beer

b. Hij [gaf Marie appels en verkocht Hendrik peren.]$_{S/NP_{+ANT}}$
he gave Mary apples and sold Harry pears

To capture the full variety of constituent coordinations in Dutch main and subordinate clauses, we must apply the same generalization as in English to the relative-pronoun and the topic categories, using the $ convention to schematize over functions of up to four arguments, as in the appendix to chapter 6:

(23) a. $(N\backslash N)\$/(S_{+SUB}\$\backslash NP)$
b. $S'_{-SUB}\$/(S_{-SUB}\$/NP)$

All of the following are then predicted to be grammatical; their derivations are suggested as an exercise:

(24) a. [Hendrik kocht en Wim at]$_{S'_{-SUB}/NP}$ de aardappels.
Hendrik bought and Wim ate the potatoes

b. [Hendrik kan en Karel moet]$_{S'_{-SUB}/VP_{+SUB}}$ aardappels eten.
Hendrik can and Karel must potatoes eat
Hendrik can and Karel must eat potatoes.'

c. aardappels die Hendrik kocht en Wim at
potatoes that Hendrik bought and Wim ate

d. Hij geeft [de politieman een bloem en de leraar appels.]$_{T\backslash((T/NP)/NP)}$
he gave the policeman a flower and the teacher apples

We also need the following version of forward composition (25) (see appendix to chapter 6) to encompass coordinations like (26) (from van Oirsouw 1987, 253).

(25) *Forward composition II* (>**B**)
$X/Y \quad Y/Z \Rightarrow_{\mathbf{B}} X/Z$
where $Y = S_{-SUB}/\$$

(26) De hond voer en de kat aai ik.
the dog feed and the cat stroke I
'I feed the dog and stroke the cat.'

The earlier examples are unaffected by these generalizations, although in many cases alternative derivations are made available.

7.3 Gapping and SVO Word Order

According to the combinatory theory, VSO "forward gapping," SOV "backward gapping," and "right node raising" in all three major word order groups, reduce to simple constituent coordination, as Maling (1972) implies they should when she equates "backward" and "forward" gapping with varieties of node raising, rather than with deletion or copying. But what about sentence-*medial* ellipsis in SVO languages like English and in Dutch/German SVO main clauses? I will begin by reviewing some salient properties of this remarkable construction.[6]

7.3.1 The Natural History of Gapping in English

Gapping in English is unlike all other varieties of constituent coordination in being almost completely insensitive to agreement. In this respect it bears some resemblance to truly elliptical constructions like VP ellipsis. (Ross 1967). Compare the following examples:

(27) a. *I cook beans and eats potatoes.
b. Harry eats beans, and I, potatoes.
c. Harry eats beans, and I do too.

However, as Jackendoff (1971) notes, gapping differs from VP ellipsis in being strictly restricted to root sentences. Hence, (28a,b) are very bad indeed, unlike VP ellipsis, (28c), and the related "pseudogapping" construction, (28d), both of which involve explicit anaphoric verbal elements:[7]

(28) a. *I know that Dexter read *Ulysses* and you say that Warren, *Dr. Zhivago*.
b. *I know that Dexter read *Ulysses* and that Warren, *Dr. Zhivago*.
c. I know that Dexter read *Ulysses* and you say that Warren did too.
d. I know that Dexter read *Ulysses* and you say that Warren did *Dr. Zhivago*.

A large number of apparent further constraints on the gapping construction have been proposed within transformational frameworks by Jackendoff (1971),

Hankamer (1971), Langendoen (1975), Stillings (1975), Hankamer and Sag (1976), and Sag (1976), reviewed by Neijt (1979). Examples like the following have been held by some of these authors to be ungrammatical under the readings indicated by the brackets:

(29) a. Dexter [went to] London, and Warren, Detroit.
b. Dexter [will give] an apple to the teacher, and Warren, a flower to a policeman.

However, Kuno (1976) has pointed out that the acceptability of gapped sentences is highly dependent upon discourse context. Sentences (29a, b) are acceptable when preceded by sentences establishing appropriate themes, presuppositions, and "open propositions" (in the sense of Wilson and Sperber 1979, Prince 1986, and Levin and Prince 1986), such as the following questions, which we will assume are asked in the context of a discussion of Dexter and Warren:

(30) a. Which city did each man go to?
b. Which man will give what to whom?

Even the most basic gapped sentence, like *Warren ate bread, and Dexter, bananas*, is only really felicitous in contexts that support (or can accommodate) the presupposition that the question under discussion is *Who ate what?* Conversely, contexts that establish a *different* open proposition cause gapping to fail. For instance, the following example (from Williams 1978) fails because by the time the putative gap is encountered, the theme is $\lambda x.hit'\ bill'\ x$, rather than $\lambda x.\lambda y.hit'xy$:

(31) Fred hit Bill. *Then Alice did too, and Bert, Tom.

Kuno (1976) shows that many other apparent constraints noted by earlier authors supposedly prohibiting gapping of strings that are in present terms surface constituents, as evidenced by the fact that they can be coordinated, are equally sensitive to context and to the inclusion of materials that are compatible with the discourse functions associated with gapping:[8]

(32) a. Twenty percent of the population [wants the president] to raise taxes, and eighty percent, to lower them.
b. Twenty percent of the population [keeps coal] in the cellar, and eighty percent, in the bath.
c. Twenty percent of the population [believes that the country is run] by madmen, and eighty percent, by crooks.

Kuno suggests on the basis of related examples and their sensitivity to island conditions that there is a close relation between gaps and residues of relativization (Kuno 1976, 317, n.29). Since in present terms the residue of relativization is a constituent, and island constraints arise when it is impossible to build such a constituent, this observation suggests that—unlike some other elliptical constructions—gapping is closely related to surface syntax.

However, not only does the "residue" of the gapping process in the rightmost conjunct appear to correspond to a discontinuous part of the leftmost conjunct—the gapped part of the leftmost conjunct may also be discontiguous:

(33) Dexter wants Watford to win, and Warren, Ipswich.

Discontinuous gapping of this kind is even more widespread in German and Dutch main-clause coordinations, like the following, because of the "V2" requirement:

(34) Jacob *heeft* appels *gegeten*, en Hendrik, peren.
Jacob *has* apples *eaten*, and Hendrik, pears
'Jacob has eaten apples and Hendrik, pears.'

Gapping therefore appears to involve something more than surface grammar, which under present assumptions is subject to the Principle of Adjacency.

Nevertheless, CCG already affords almost everything we need to account for the above phenomena. For example, the residues and the gapped element itself in each of the well-known family of gapped sentences mentioned in chapter 2 are all constituents under one or another of the possible analyses of *you want to try to begin to write a play*:

(35) I want to try to begin to write a novel, and ...
a. you, to try to begin to write a play.
b. you, to begin to write a play.
c. you, to write a play.
d. you, a play.

Conversely, when failure of coordination suggests that a substring cannot be a constituent, even in this extended sense, it cannot be a gap either. Compare (36a) (see (17) of chapter 1) with (36b) under the intended reading:

(36) a. *Three cats in twenty like velvet and in ten prefer corduroy.
b. *Three cats in twenty like velvet, and two dogs, corduroy.

In all of the earlier examples the coordination of argument sequences was brought under the general mechanism of constituent coordination by type-

raising the arguments and composing to yield a function over verbal and sentential functors—as in the English argument cluster coordination example (40) of chapter 3, repeated here:

(37) Give$_{(VP/NP)/NP}$ [a teacher an apple]$_{VP\backslash((VP/NP)/NP)}$ and [a policeman a flower]$_{VP\backslash((VP/NP)/NP)}$

It is therefore tempting to believe that the sequence of arguments that is left behind by gapping is also a constituent assembled by order-preserving type-raising and composition, that coordinates with an adjacent category of the same type. Gapping would be an instance of constituent coordination under the extended sense of the term implicated in combinatory grammar. Such a constituent would semantically be a function over a tensed verb, so its syntactic category would have to follow suit, as in (38):

(38) (A teacher likes an apple, and) [a policeman, a flower.]$_{S\backslash((S\backslash NP)/NP)}$

The theory of SVO gapping presented below has two components. First, I will show that gapped right conjuncts like *Warren, potatoes* also have the status of constituents under the present theory, just like argument sequences in SOV and SVO languages, and just as in argument cluster coordination. In particular, like argument cluster conjuncts, the gapped constituent has an interpretation that enables it to combine with the missing verbal component to yield a correct interpretation for the whole. Moreover, no rule that will produce a backward-gapped rightward-looking function from the English type-raised argument categories is permitted by the universal Principles of Adjacency, Consistency, and Inheritance that were in earlier chapters claimed to constrain combinatory rules, together with the "order-preserving" constraint on the type-raised categories that are allowed in an SVO language.

More controversially, I will propose, following Steedman 1990, that the "gapped" conjunct is coordinated with a "virtual" constituent of the same type in the ungapped left conjunct. The second part of the argument suggests a way to recover this virtual constituent, together with another virtual constituent corresponding to the gap, even though neither may be a derivational constituent—or even a contiguous substring—of the left conjunct. The possibility arises because *associativity* of functional composition induces semantic equivalence over certain classes of derivations. Furthermore, the *parametric neutrality* of combinatory rules like composition and application allows the recovery in a restricted sense of certain constituents under one derivation from the result of another.

7.3.2 The Category of the Right Conjunct[9]

I have assumed so far that leftward type-raising in English yields the following quite general category:

(39) $T\backslash(T/NP)$

On the assumption that English auxiliaries bear the category of VSX verbs as in (40a) and that coordinations like (40b) are allowed, it must be the case that subjects, including explicitly nominative pronominal ones, also bear this category, schematizing over functions over S/NP, $(S/NP)/NP$, $(S/PP)/NP$, and so forth, as they do in the German and Dutch main clauses discussed in chapter 6.

(40) a. are := $(S_{inv}/NP)/NP_{2s}$

b. Are you now, or have you ever been, a member of the Friends of the Legume Film Society?

It follows that we already have in the form of the backward composition rule (41) of chapter 3 a rule that will make a nonstandard constituent out of the gapped right conjunct, complete with an impeccable semantic interpretation, as in the derivation in (41) below.

(41) Dexter eats beans, $\dfrac{\text{and}}{CONJ}$ $\dfrac{\dfrac{\text{Warren,}}{T\backslash(T/NP_1)}<\mathbf{T}\quad \dfrac{\text{potatoes}}{T\backslash(T/NP_2)}<\mathbf{T}}{T\backslash((T/NP_2)/NP_1)}<\mathbf{B}$

The resultant argument cluster category is specified as needing to combine with a VSO verb, which do not in general exist in English. As a consequence, the category will be unable to take any further part in any normal derivation. This much is desirable, as it correctly prevents the following from being accepted in English, unlike the related examples in Dutch and German:

(42) $\dfrac{\dfrac{*\text{Eats}}{(S\backslash NP)/NP}\quad \dfrac{\dfrac{\text{Warren}}{T\backslash(T/NP_1)}<\mathbf{T}\quad \dfrac{\text{potatoes}}{T\backslash(T/NP_2)}<\mathbf{T}}{T\backslash((T/NP_2)/NP_1)}<\mathbf{B}}{}*$

Nevertheless, we have at least found a way to make the gapped conjunct a constituent, which is the first step toward making grammatical rules apply to it under the Constituent Condition on Rules.

It is striking that related derivations of illegal gapped clauses like (28a–d) are ruled out for the same reason as the Fixed-Subject Condition violations

discussed in chapter 4, because of the continued exclusion from the grammar of English of the Forward Crossed Composition rule:

(43) *Warren eats beans, and I believe that Dexter, potatoes

$$\dfrac{\dfrac{\dfrac{}{CONJ}\ \text{and} \quad \dfrac{}{S/S}\ \text{I believe that} \quad \dfrac{}{\mathrm{T}\backslash((\mathrm{T}/NP)/NP)}\ \text{Dexter, potatoes}}{}}{}\ *$$

(44) *Warren eats beans, and I believe that Dexter, potatoes

$$\dfrac{\dfrac{\dfrac{}{CONJ}\ \text{and} \quad \dfrac{}{S/(S\backslash NP) : \lambda p.believe'(p\ dexter')me'}\ \text{I believe that Dexter,} \quad \dfrac{}{\mathrm{T}\backslash(\mathrm{T}/NP) : \lambda p.p\ potatoes'}\ \text{potatoes}}{}}{}\ *$$

For the same reason, gaps within the scope of a complementizer are blocked:

(45) *I believe that Warren eats beans, and that Dexter, potatoes

$$\dfrac{\dfrac{\dfrac{}{S}\ \text{I believe that Warren eats beans,} \quad \dfrac{}{CONJ}\ \text{and} \quad \dfrac{}{S'/S}\ \text{that} \quad \dfrac{}{\mathrm{T}\backslash((\mathrm{T}/NP)/NP)}\ \text{Dexter, potatoes}}{}}{}\ *$$

(46) I believe that Warren eats beans, and Dexter, potatoes

$$\dfrac{\dfrac{\dfrac{}{S}\ \text{I believe that Warren eats beans,} \quad \dfrac{}{CONJ}\ \text{and} \quad \dfrac{}{\mathrm{T}\backslash((\mathrm{T}/NP)/NP)}\ \text{Dexter, potatoes}}{}}{}$$

Thus, we correctly capture the restriction of gapping to clauses immediately dominated by the conjunct.[10]

On the assumption that adjuncts bear the category $(S\backslash NP)\backslash(S\backslash NP)$, the following derivation will block, because backward composition will deliver a fragment that is incompatible with any analysis of the left conjunct:[11]

(47) Dexter ran quickly, and Warren, slowly

$$\dfrac{\dfrac{\dfrac{}{S}\ \text{Dexter ran quickly,} \quad \dfrac{}{CONJ}\ \text{and} \quad \dfrac{\dfrac{}{\mathrm{T}\backslash(\mathrm{T}/NP_1)}\ \text{Warren,} \quad \dfrac{}{(S\backslash NP_2)\backslash(S\backslash NP_2)}\ \text{slowly}}{S\backslash((S\backslash NP_2)/NP_1)}\ <\mathbf{B}}{}$$

This is in fact a desirable result, because if *Warren slowly* were to come in general to bear a gap-permitting category, it would threaten to allow examples like the following to mean something like *Dexter ran and Warren ran quickly* (see discussion in Sag 1976 and Wood 1988):

(48) *Dexter ran, and Warren, quickly

$$\dfrac{\dfrac{\dfrac{}{S}\ \text{Dexter ran,} \quad \dfrac{}{CONJ}\ \text{and} \quad \dfrac{\dfrac{}{\mathrm{T}\backslash(\mathrm{T}/NP_1)}\ \text{Warren,} \quad \dfrac{}{(S\backslash NP_2)\backslash(S\backslash NP_2)}\ \text{quickly}}{S\backslash((S\backslash NP_2)/NP_1)}\ <\mathbf{B}}{}$$

We must instead assume that examples like (47) arise from the tendency noted in section 4.3.2 for verbs to behave as if they subcategorized for certain types of adverbs, to allow derivations like the following for sentences like *Dexter ran*

quickly, and Warren, slowly, while continuing to exclude examples like **Ran Warren slowly*:

(49) Dexter ran quickly, and Warren, slowly

$$\dfrac{\dfrac{\dfrac{}{CONJ}\ \dfrac{}{T\backslash(T/NP)}\ \dfrac{}{T\backslash(T/ADV)}}{T\backslash((T/ADV)/NP)}{<}\mathbf{B}}$$

Although we have yet to see how a gapped verb *ran* of category $(S/ADV)/NP$ can be accessed in such examples, it is clear that such a verb cannot be involved in examples like (48).

The schema $T\backslash(T/NP)$ merely abbreviates categories raised over the categories of English grammar that are permitted under the Principle of Categorial Type Transparency, as informally defined in chapter 3. It follows that the present mechanism for constructing gapped right conjuncts by composing leftward type-raised arguments will permit such conjuncts only when they preserve the linear order of subject, object, and more oblique arguments.

The rule allows several more complex types of gapped right conjunct. For example, it allows (50), adapted from Aoun et al. 1987:[12]

(50) Dexter gave a teacher an apple, and Warren, a policeman a flower

$$\dfrac{\dfrac{}{S}\ \dfrac{}{CONJ}\ \dfrac{\dfrac{}{T\backslash(T/NP_1)}\ \dfrac{}{T\backslash((T/NP_3)/NP_2)}}{T\backslash(((T/NP_3)/NP_2)/NP_1)}{<}\mathbf{B}}$$

On the assumption that subject phrases like *which woman* also bear the leftward type-raised category, as they must when in situ, as in *What book did which woman buy?*, the following example (also of a kind discussed in Aoun et al. 1987) is accepted:

(51) I wonder which man met Dexter, and which woman, Warren

$$\dfrac{\dfrac{}{S/S_{iq}}\ \dfrac{}{S_{iq}}\ \dfrac{}{CONJ}\ \dfrac{\dfrac{}{T\backslash(T/NP_1)}\ \dfrac{}{T\backslash(T/NP_2)}}{T\backslash((T/NP_2)/NP_1)}{<}\mathbf{B}}$$

Parallel derivations are allowed for *I wonder which teacher you gave an apple, and which policeman, a flower*, and *Which apple did you give to the teacher, and which flower to the policeman?*

However, the rule correctly excludes all of the following, because the arguments in the gapped conjunct are not in canonical order:[13]

(52) a. *Which man did Dexter invite, and which woman, Warren?
b. *Which man did Dexter introduce to Warren, and [which woman,]$_{(S/PP)\backslash((S/PP)/NP)}$ [Gilbert,]$_{((S/PP)/NP)\backslash(((S/PP)/NP)/NP)}$ [to George?]$_{S\backslash(S/PP)}$
c. *Freeman wondered what Hardy gave to Willis, and [what]$_{(S/PP)/((S/PP)/NP)}$ [Gilbert,]$_{((S/PP)/NP)\backslash(((S/PP)/NP)/NP)}$ [to George.]$_{S\backslash(S/PP)}$

Examples (52b) and (52c) are also of a kind considered by Aoun et al. 1987, (12-13).

Finally, the theory correctly predicts that the gap may be discontinuous, as in (33) (repeated here), since the second NP has the same category as any NP complement.

(53) Dexter wants Watford to win, and Warren$_{T\backslash(T/NP)}$, Ipswich$_{T\backslash(T/NP)}$.

These possibilities for gapping are a direct corollary of the way in which the permitted combinatory rules project directionality from the English SVO lexicon. It is also a corollary of the Principles of Adjacency, Consistency, and Inheritance, and the fact that English verb are confined to SVX and VSX patterns, that no forward combining gapped constituent over tensed verbs can be constructed. In particular, the Principle of Inheritance ensures that the composite function will be *backward* looking, just as in the case of a VSO language (see (17) and (19)).[14]

Now if only (41) had the following analysis, we would have an answer to the question of why SVO languages pattern with the VSO alternative in gapping on the right:

(54) *eats$_{((S/NP)/NP)}$ [Dexter beans]$_{S\backslash((S/NP)/NP)}$ and [Warren potatoes]$_{S\backslash((S/NP)/NP)}$

The nonstandard constituent is leftward-looking, so it must occur to the right of the verb. That fact would enable the coordination rule to apply to yield the effect of a gap on the right. A gap on the left would be impossible with this category, just as it is in VSO languages.

Of course, (54) is *not* a possible surface analysis of (41), and we still need to say how both the gap and the appropriate nonstandard gapped constituent can be recovered, in the face of the fact that both the putative conjunct and the gap itself may correspond to discontinuous substrings of the sentence. But the directionality result, coupled with capturing several empirical constraints that

set gapping apart from other elliptical constructions, is a strong argument. It suggests that we should resist any solution to this problem that extends the calculus by compromising the Principle of Adjacency. The next sections propose one possible alternative. It will be convenient to refer to the two putative virtual constituents as the "virtual gap" and the "virtual left conjunct," respectively.

7.3.3 A Hypothesis Concerning the Left Conjunct

It is important that any proposal for revealing "virtual" adjacent nonstandard constituents in the left conjunct should conform to the Principles of Adjacency, Consistency, and Inheritance, if it is not to compromise the claims of earlier sections concerning Universal Grammar. Interestingly, there is a way of using the rules of the grammar itself to yield the virtual constituent, so that the grammar as a whole continues to respect the basic constituent order specified in the lexicon in the way it has up to this point, even though the subject and the object are not contiguous in the string.

The device in question depends on a property of the combinatory rules that was first identified by Pareschi (1986, see Pareschi and Steedman 1987) as providing a possible basis for a technique for parsing in the face of so-called spurious ambiguity, a topic to which I will return in part III. I will call this property "parametric neutrality." It can be described as follows:

(55) *Parametric neutrality*
Specifying the syntactic type of any two categories that are related by a given combinatory rule determines the syntactic type of the third.

For example, we normally think of a rule like application as taking a function of type X/Y or $X\backslash Y$ and an argument of type Y as input parameters, and combining them to yield the result X. But because any two categories between them specify all the information that is required to determine the type of the third, we can consider *any two* of the three categories that such a rule relates as the input parameters and use the rule to determine the type of the third. For example, we can define the argument type Y and the result type X of application to determine the type of the third category X/Y.

The observation holds for all of the combinatory syntactic rules, as may be verified by inspecting the three familiar rule types exemplified here.

(56)

Application	X/Y	Y	$\Rightarrow$	X
Composition	X/Y	Y/Z	$\Rightarrow_{\mathbf{B}}$	X/Z
Substitution	Y/Z	$(X\backslash Y)/Z$	$\Rightarrow_{\mathbf{S}}$	X/Z

This observation concerning syntactic types does not extend without qualification to their interpretations. In Pareschi and Steedman 1987 we assumed that categories were matched to rules via first order unification of the kind familiar from the programming language Prolog in which λ-terms have to be simulated, in a manner discussed at length by Pereira and Shieber (1987). In this framework the related nonstandard invocations yield only "dummy" constant functions. For example, consider the gapped sentence *Dexter eats apples, and Warren, pears*, and consider instantiating the backward application rule with Y as the "virtual" VSO verb category $(S/NP)/NP : v$ and X as the left conjunct $S : eats'apples'dexter$. First-order unification would yield the following constant function as the value of $X\backslash Y : f$:

(57) $S\backslash((S/NP)/NP) : \lambda v.eats'apples'dexter$

Crucially, although this virtual category can only reduce to yield the same proposition we started with, it can first coordinate with the semantically nonvacuous right conjunct, since it has the same type. If the gapped verb is somehow made available, the whole coordinate fragment can then combine to yield an S with the following interpretation, as desired:

(58) $S : and'(eats'warren'pears')(eats'apples'dexter')$

Crucially, the verbal argument must be to the left.

However, mere first-order unification will not yield a verb that can combine in this way. It will again yield only a dummy category, which again will semantically be a constant function that either will refuse to combine with the right conjunct, blocking the derivation, or will yield an incorrect meaning. Accordingly, the analysis in Steedman 1990 used unification only to specify the syntactic type of the verb. The claim was that its interpretation was obtained from an extra sentential discourse context including representations of elements such as given information or background, via anaphora rather than via the unification process itself.

Although we have already noted an interesting relation between possibilities for gapping and the state of the discourse context, reflected to some extent in intonation, the sententially bound nature of gapping suggests that this information must in fact be obtained from the left conjunct rather than from extra sentential context.

One mechanism for the kind of abstraction or matching that appears to be involved that has recently received attention in the literature is higher-order unification. Pareschi (1989) has proposed a number of applications for higher-

order unification in natural language grammar, and Dalrymple, Shieber and Pereira (1991) and Shieber, Pereira and Dalrymple (1996) show how a generalization of the technique described above could be applied more generally to a wide variety of elliptical and anaphoric constructions, via higher-order unification over *typed* λ-terms (Hindley and Seldin 1986) using Huet's algorithm (Huet 1975; Huet and Lang 1978).

This observation is interesting for present purposes, since interpretations are coupled in CCG with a syntactic category that is simply their semantic type plus some directional information. CCG categories are themselves in effect typed λ-terms. Higher-order unification and other implementations of matching or abstraction therefore remain interesting possibilities to explore as a basis for gap retrieval. In particular, the fact that gaps exhibit the same ambiguities between "strict" and "sloppy" anaphora as VP ellipsis suggests a common mechanism:

(59) a. Dexter fed his cat chicken, and Warren did too.
b. Dexter fed his cat chicken, and Warren, tunafish.

Nevertheless, higher-order unification is equivalent to quite general abstraction over typed terms and will deliver spurious interpretations including vacuous abstractions and interpretations that violate the generalization that rightward arguments in English must "wrap," discussed in chapter 4. Dalrymple, Shieber and Pereira (1991) and Shieber, Pereira and Dalrymple (1996) filter these spurious Logical Forms via structural criteria of "primary occurrence," domination, and identification of "source parallel elements." Many of these structural criteria replicate parts of the grammar itself, and for the present purpose we must seek a more purely grammatical mechanism.

These observations provide a strong motivation for trying to exploit the property of parametric neutrality within the competence grammar itself to subsume gapping to ordinary constituent coordination.

One way to do this is to assume that the gap is interpreted as the theme of the left conjunct, made available from the interpretation of the left conjunct via a discourse mechanism related to Kuno's (1976, 310) Functional Sentence Perspective Principle.

Mats Rooth (personal communication) has suggested that the following kind of exchange presents problems for the idea that the gap is a theme, and for information-structure based proposals for gap recovery in general:

(60) Q: Do Sid and Nancy like Dexter and Warren?
A: No! Sid LOATHES Dexter, and NANCY, WARREN.

However, the example is at least compatible with an L+H* pitch accent on *loathes* and a mid-sentential LL% boundary, and it is therefore possible to argue in the terms of chapter 5 that it involves the accommodation of a marked theme $\lambda x.\lambda y.loathe'xy$, standing in contrast to the alternative theme $\lambda x.\lambda y.like'xy$.

The theme of the left conjunct can be made available to the right conjunct by defining the term $\theta''left$ as an anaphor that picks out the theme that the left conjunct has established along lines described in chapter 5. This term is introduced grammatically via the following production, related to backward application:[15]

(61) *Virtual conjunct–revealing rule (<dcomp)*

$X : left \implies Y : \theta''left \quad X\backslash Y : \lambda y.left$

where $Y = S/\$$

The interpretation of the revealed function $X\backslash Y$ is defined as a constant function, rather than as applying y to the discontinuous rheme of $left'$, since the latter would eventually yield $left'$ anyway.

Although the rule as written is nondeterministic, in a parser we would arrange to constrain the category $X\backslash Y$ to be that of the right conjunct.[16] It will be convenient to refer to such a use of combinatory rules in the grammar as "category decomposition." The application of this rule in derivations will be indicated by a *dotted* underline and an index $<dcomp$, identifying the combinatory rule involved as backward application.

For example, rule (61) can be used to deliver the following categories from a nonstandard invocation of backward application in which the result category X is set to $S : eats'bread'dexter'$, the function type $X\backslash Y$ among the inputs is specified as the same type as the right conjunct $S\backslash((S/NP)/NP)$, and the decomposition forces the revealed verb to be a "virtual" VSO verb:

(62)

$$\dfrac{\dfrac{\dfrac{\text{Dexter eats bread,}}{\substack{S \\ : eats'bread'dexter'}} \quad \underline{\text{and}}_{CONJ} \quad \dfrac{\text{Warren, potatoes}}{\substack{S\backslash((S/NP)/NP) \\ : \lambda f.f\ potatoes'warren'}}}{\dfrac{\substack{((S/NP)/NP) \\ : \theta''(eats'bread'dexter')} \quad \dfrac{\substack{S\backslash((S/NP)/NP) \\ : \lambda y.eats'bread'dexter'}}{\substack{S\backslash((S/NP)/NP) \\ : \lambda f.and'(f\ potatoes'warren')(eats'bread'dexter')}}\ <\Phi>}{\substack{S : and'(\theta''(eats'bread'dexter')potatoes'warren')(eats'bread'dexter') \\ = S : and'(eats'potatoes'warren')(eats'bread'dexter')}}\ <$$

The attraction of category decomposition is that it exploits exactly the same rules as the original grammar and therefore preserves the projection of lexical directionality from the lexicon under coordination.

The prediction that gapping in English and every other SVO language is *forward* gapping (see section 7.3.2) remains in force when category decomposition is included. Even though an English subject and object can raise and compose on the left of a conjunct, backward gapping on the SOV pattern as in (63) is excluded by universal principles. The recovery of the virtual conjunct would require a rule of decomposition that violated the Principle of Consistency, and it would yield a result requiring a similarly illegal rule to recombine.

(63) $$\dfrac{\dfrac{*\text{Warren, potatoes}}{S\backslash((S/NP)/NP)}\quad \dfrac{\text{and}}{CONJ}\quad \dfrac{\text{Dexter bought bread}}{\dfrac{S}{S\backslash((S/NP)/NP)\quad (S/NP)/NP}\ *dcomp}}{}$$

Nor could this example be permitted in a language like English by composing forward type-raised categories and including a forward virtual conjunct-revealing rule, as in the following illegal derivation:

(64) $$\dfrac{\dfrac{\dfrac{*\text{Warren,}}{T/(T\backslash NP)}\quad \dfrac{\text{potatoes}}{T/(T\backslash NP)}}{S/((S\backslash NP)\backslash NP)}\ *>\mathbf{B}\quad \dfrac{\text{and}}{CONJ}\quad \dfrac{\text{Dexter bought bread}}{\dfrac{S}{S/((S\backslash NP)\backslash NP)\quad (S\backslash NP)\backslash NP}\ >dcomp}}{S/((S\backslash NP)\backslash NP)}\ <\Phi>$$

The crucial composition of the two forward–type-raised subjects is ruled out for English (though not for German and Dutch) by the restrictions on type-raising to raising to parametrically licensed categories permitted by Type Transparency, discussed in earlier chapters, as shown in example (42) of chapter 3.

Nor does the inclusion of category decomposition in the theory permit "anti-gapping"—that is, overgenerations of the following kind, in which the *leftmost* product of decomposition is made available for coordination on the pattern of a VSO language, rather than the rightmost, because virtual VSO verbs bear a different category from real SVO verbs and inverting auxiliaries like (40a):[17]

(65) $$\dfrac{\dfrac{*\text{Cooks,}}{(S\backslash NP)/NP}\quad \dfrac{\text{and}}{CONJ}\quad \dfrac{\text{Dexter eats beans}}{\dfrac{S}{(S/NP)/NP\quad S\backslash((S/NP)/NP)}\ <dcomp}}{}\ *<\Phi>$$

The same technique will reveal nonconstituent virtual gaps as well. Consider example (41), repeated here:

(66) Dexter will buy bread, and Warren, potatoes.

The gapped right conjunct can be assembled in the usual way into a leftward-looking function over transitive verbs, of the following category:

(67) $S\backslash((S/NP)/NP) : \lambda p.p\ potatoes'\ warren'$

By assumption, the left-conjunct-revealing rule (61) can decompose the S on the left into a category with the type of a transitive verb and a category of the same type as the right conjunct. The entire sentence can then be syntactically derived as follows (semantics is omitted, since the real work is done by the anaphoric θ'' anaphor):

(68)
$$\dfrac{\dfrac{\dfrac{\text{Dexter will buy bread,}}{S} \quad \dfrac{\text{and}}{CONJ} \quad \dfrac{\text{Warren, potatoes}}{S\backslash((S/NP)/NP)}}{(S/NP)/NP \quad S\backslash((S/NP)/NP)}\ {<}dcomp}{S\backslash((S/NP)/NP)}\ {<}\Phi{>}}{S}\ <$$

On the assumption that the anaphoric operator θ'' can reveal discontinuous themes of the kind discussed in section 5.3.3, examples like (69) will also be accepted with appropriate interpretations:[18]

(69)
$$\dfrac{\dfrac{\dfrac{\text{Dexter wants Ipswich to win}}{S} \quad \dfrac{\text{and}}{CONJ} \quad \dfrac{\text{Warren, Watford}}{S\backslash((S/NP)/NP)}}{(S/NP)/NP \quad S\backslash((S/NP)/NP)}\ {<}dcomp}{S\backslash((S/NP)/NP)}\ {<}\Phi{>}}{S}\ <$$

For the same reason, discontinuous gapping is not subject to the same Fixed-Subject Condition as extraction. For example, related sentences with non-nominative second NPs in the gapped conjunct seem much better than those with nominatives like (53b). For example, consider the following, uttered in a context in which the question at issue is which team each of the men claim to have won the Cup:

(70) Dexter said (that) Watford won, and Warren, Ipswich.

This exemption from the Fixed-Subject Condition constitutes a major objec-

tion to any purely syntactic account of gapping, as it does for pronoun binding and quantifier scope.[19]

As noted earlier, multiple or discontinuous gapping of the kind illustrated in (69) is even more common in German and Dutch main clauses, so it is time to return to the earlier analysis of Dutch and see how the main-clause grammar emerges from the same generalization.

7.3.4 Verb Gapping in Dutch Main Clauses

Gapping in Dutch and German is even simpler than gapping in English, because the category of main-clause verbs is VSO. It follows that the simple backward composition rule, (21) of chapter 6, repeated here as (72), can apply to the order-preserving backward type-raised categories to deliver gapped right conjuncts like *Hendrik peren* in sentences like the following:

(71) Jacob heeft appels gegeten en Hendrik peren.
Jacob has apples eaten and Hendrik pears
'Jacob ate apples, and Hendrik, pears.'

(72) *Dutch/German backward composition* ($<\mathbf{B}^n$)
$(Y\backslash Z)\$ \quad X\backslash Y \quad \Rightarrow_{\mathbf{B}^n} \quad (X\backslash Z)\$$

It follows that if we further assume that Dutch has the same virtual conjunct–revealing rule (61) that English has, limited to Dutch VSO main verbs as in (73), then identical main-clause gapping is allowed, as in (74):

(73) *Dutch virtual conjunct–revealing rule I (<dcomp)*
$X : left \implies Y : \theta'' left \quad X\backslash Y : \lambda y.left$
where $Y = S/\$$

(74)

Jacob heeft appels gegeten	en	Hendrik	peren
S'_{-SUB}	$CONJ$	$T\backslash(T/NP_1)$	$T\backslash(T/NP_2)$
		$T\backslash((T/NP_2)/NP_1)_{+ANT}$ (<B)	
$((S_{-SUB}/NP_2)/NP_1)_{+ANT}$ $S'_{-SUB}\backslash((S'_{-SUB}/NP_2)/NP_1)_{+ANT}$ (<dcomp)			
$S'_{-SUB}\backslash((S'_{-SUB}/NP)/NP)_{+ANT}$ (<Φ>)			
S'_{-SUB} (<)			

As in the case of English gaps like (33), the gapped material is discontinuous, a possibility that again stems from the assumption that the virtual verb translation is recovered via Information Structure as a discontinuous theme.

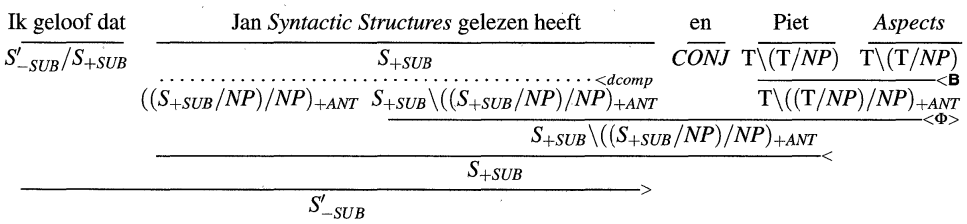
The variable T in the Dutch backward type-raised category $T\backslash(T/NP)$ is free to match the topicalized main-clause category $S'_{\text{-SUB}}$ in the above derivation, despite the fact that there are no "real" verbs of category $(S'_{\text{-SUB}}/NP)/NP$, on

the assumption that such categories are consistent with the directional parameterization of the language. Since such a category is distinct from that of a real main verb, the following potential overgeneration is blocked:

(75) *Kocht$_{(S/NP)/NP}$ en [Hendrik at appels.]$_{S'}$

However, by the same token, T can match S_{+SUB}, the subordinate-clause category, to precipitate a similar decomposition in a subordinate left conjunct. This correctly allows rightward gapping in Dutch embedded clauses, as in the following sentence from van Oirsouw (1982, 555, (8c)), despite the involvement of SOV order in the left conjunct and the absence from the Dutch lexicon of "real" verbs of category $(S_{+SUB}/NP)/NP$:

(76)



This analysis implies the following claim. According to this theory, the SOV+SO pattern of gapping is allowed in Dutch only because the directional parametric specification of legal categories for the Dutch lexicon includes VSO main verbs, which in turn allow arguments to be raised over them to give backward raised categories. If Dutch were a "pure" SOV language, like Japanese, it would not license T\(T/*NP*) categories, and would be predicted not to allow forward gapping on the SOV+SO pattern, as is indeed the pattern with more strictly SOV languages like Japanese.[20]

Of course, predictions cut both ways. If we predict that when an SOV language like Dutch allows VSO/SVO main-clause order, it may allow SOV+SO gapping as well as the usual SO+SOV, then we necessarily also predict that a VSO/SVO language that allows SOV as a main-clause order may allow SO+VSO/SVO as well as the standard VSO/SVO+SO.

Zapotec (Rosenbaum 1977) appears to be exactly such a language, the mirror image in this respect of Dutch. It is clearly a VSO language, since indirect questions and other subordinate clauses require that order. However, it also allows SVO, OVS, and SOV as intonationally and pragmatically marked main-clause orders. It allows SO+VSO/SVO gapping, as well as the VSO/SVO+SO order that Ross's (1970) generalization would predict. Tojolabal (Furbee 1974) appears to be a similar case.

Zapotec and Tojolabal have often been cited, following Rosenbaum and Furbee themselves, as counterexamples to Ross's basic generalization. However, once it is understood how true discontinuous gapping actually works, it is clear that these languages only appear to violate the generalization because they are not purely verb-initial, and include verb-final lexical entries for verbs, which in turn determine the categories of type-raised arguments. If Ross's generalization is restated in the following form, then they are entirely consistent with it:

(77) *Ross's generalization (revised)*
The possibility of "rightward gapping" in a given language depends on the availability to its lexicon of rightward-combining verbs (and hence the possibility of rightward categories raised over them), and the possibility of "Leftward Gapping" depends on the availability to its lexicon of leftward-combining verbs (and the associated raised categories).

7.4 Other Elliptical Phenomena

The assumption that gapping arises in languages like English via the availability of a backward type-raised category for subjects as well as other arguments, and of the revealing rule (61), predicts that a subject or any other argument alone should be able to coordinate with a sentence in the same way, giving rise to the construction that Ross called "Stripping," illustrated in the following examples:[21]

(78) a. Dexter ran away, and Warren (too).
b. Dexter ran away, but not Warren.

(79) a. Dexter gave a flower to a policeman, and chocolates (too).
b. Dexter sent a flower to a policeman, but no chocolates.

(80) a. Dexter gave a policeman a flower, and a judge (too).
b. Dexter gave every policeman a flower, but no judge.

For example:[22]

(81)
$$\frac{\dfrac{\text{Dexter}\quad \dfrac{\text{ran away,}}{S}}{S/NP}\quad \dfrac{\text{and}}{CONJ}\quad \dfrac{\text{Warren}}{S\backslash(S/NP)}}{\dfrac{S/NP \quad S\backslash(S/NP)}{S\backslash(S/NP)}\ <\Phi>}\ <dcomp$$

(82) Dexter gave a flower to a policeman and chocolates

$$\frac{\dfrac{\text{Dexter gave a flower to a policeman}}{S} \quad \overline{CONJ} \quad \overline{S\backslash(S/NP)}}{\dfrac{S/NP \qquad S\backslash(S/NP)}{S\backslash(S/NP)}\ <dcomp} <\Phi>$$

As would be expected under the assumption that Information Structure is involved, acceptability is helped by appropriate contextual questions, such as "Which boys ran away?" and "What presents did Dexter give to a policeman?," as well as by discourse-linking particles like *too* and contrastive conjunctions like *but not*.

The implication that gapping and stripping are related seems to be empirically correct. As Morgan (1973) points out, they both share the curious restriction to root subjects, which was explained in the case of gapping by the fact that fragments like *I believe that Dexter* do not bear the category of a VSO subject: this explanation generalizes correctly to stripping:

(83) a. *Dexter likes apples and I think that Warren pears.
b. *Dexter likes apples and I think that Warren (too).

Chao (1987) offers extensive crosslinguistic evidence that languages like English and French that allow gapping, also allow stripping, while those that do not, like Chinese (which otherwise seems qualified to allow rightward gapping), do not allow stripping either.

Sluicing (84a) and VP ellipsis (84b) do not appear to be amenable to analysis in the same syntactic terms as stripping and gapping, since the requisite categories are not otherwise present in the grammar and/or the requisite rules would violate the Principle of Consistency, even on the optimistic assumptions about the categories of the elided conjuncts embodied here:

(84) a. [Dexter did something with the beans,]$_S$ but [I don't know what.]$_{S/(S/NP)}$
b. [Somebody has to do the job,]$_S$ but [I know that I won't.]$_{S/VP}$

Under present assumptions such constructions must be regarded as being mediated by a quite separate, presumably purely anaphoric mechanism, as their freedom to occur outside the context of coordination suggests, rather than as being syntactically mediated, as gapping and stripping are according to the present theory.[23] Their constituent categories are presumably the following, which trivially conform to the assumption made here that only like types can coordinate:

(85) a. [Dexter did something with the beans,]$_S$ but [I don't know what.]$_S$
b. [Somebody has to do the job,]$_S$ but [I know that I won't.]$_S$

The suggestion that they are mediated by anaphoric processes rather than syntactic ones is borne out by the absence of word order-dependent constraints parallel to those noted by Ross for gapping. (That is, we do not find "backward VP anaphora" and "backward sluicing" predominating in verb-final languages, but rather the same "forward" varieties that predominate in English.)

The conclusion that gapping is syntactically unrelated to sluicing and VP-ellipsis is contrary to Hankamer and Sag's (1976) and Sag and Hankamer's (1984) claim that all three fall into their "surface anaphoric" or "elliptical" class of constructions, as opposed to their other class of elliptical constructions mediated by the "deep" or "model-interpretive" anaphora that is characteristic of pronouns. However, see Williams 1977, Schachter 1977, and Chao 1987, 112–127, for further arguments that support the present proposal, according to which VP ellipsis and sluicing are mediated by model-interpretive anaphora, like pronouns, and their more restricted character arises from the special nature of their antecedents. Gapping and Stripping, by contrast, are claimed here to be under the control of syntax, pragmatically specialized though they are.

7.5 A Cautious Conclusion

The introduction of rules of decomposition is a radical departure. It involves an appeal to discourse context to determine Logical Form and potentially threatens the constrained nature of the core CCG. Such rules should not be invoked lightly, and they are not needed in the chapters that remain. If something more purely grammatical will do the job, then it should be welcomed. Nevertheless, the widespread involvement of noncontinuous gapping, particularly in Germanic languages, makes it seem certain that something more than pure combinatory rules will be needed, at least in semantic terms.

If we adopt the hypothesis that rules of decomposition are to be allowed in syntax, then medially gapped sentences arise from the coordination of two nonstandard constituents—in descriptive terms, two gapped sentences—and their combination with a third constituent—the virtual gap. In this respect the present proposal is akin to theories in which gapping arises from the restoration of the gapped conjunct to the status of a standard clause, the gapped material being accessed via processes of anaphora or structure copying. The advantage of the present approach is that an analysis, including an interpretation, can be achieved by combining elements that are strictly adjacent by strictly syn-

tactic type-driven operations. The theory thus explains why constituent order under coordination exhibits Ross's (1970) universals, as presently revised. According to the present theory, as with related categorial analyses that similarly extend the notion of constituent (e.g. Stump 1978; van der Zee 1982; Cremers 1983; Oehrle 1987; Dowty 1988; Moortgat 1988a; Morrill 1988; Wood 1988; Hepple 1990; Solias 1992; Solias Aris 1996; Morrill and Solias 1993; Houtman 1994; Hendriks 1995; and Versmissen 1996), everything that can coordinate, including medially "gapped" conjuncts, is a constituent under the generalized definition of that notion that is afforded by categorial grammars.

Within the framework of CCG the twin principles of Consistency and Inheritance to which combinatory rules are subject predict the observations of Ross and Maling concerning the dependency of forward and backward "gapping" in coordinate structures upon the lexical specification of clause constituent orders in any given language, including the observation that SVO patterns with VSO in forbidding the backward variety. These principles further predict that languages like Dutch/German and Zapotec that are not "purely" verb-final or verb-initial may allow both forward *and* backward gapping, a fact that has hitherto been supposed to controvert Ross's generalization. On the contrary, such languages merely underline the fact that Ross's generalization should be thought of as applying to parametrically specified *lexicons* specifying verbs as having one or more orders such as VSO, SOV etc., rather than as applying to languages via a single "underlying" word order, a notion the present theory entirely avoids.

PART III

Computation and Performance