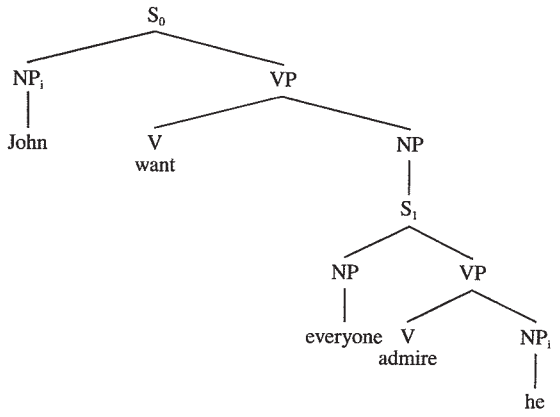
# 6. *Rule Interaction*

## *a. Feeding and Bleeding*

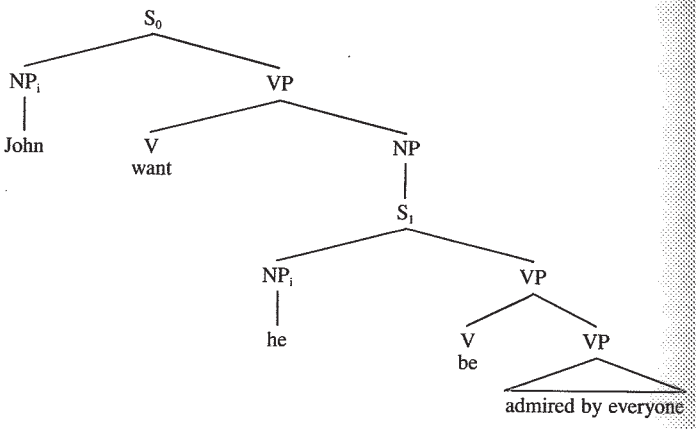
According to the analysis developed in chapter 5, (1a) has a deep structure as in (1b), aside from the complementizer, which is omitted here; in its derivation, Passive applies to $S_1$, yielding the intermediate structure (1c), and Equi-NP-deletion applies to $S_0$, deleting $he_i$;[1]

(1) a. John wants to be admired by everyone.

b.




c.





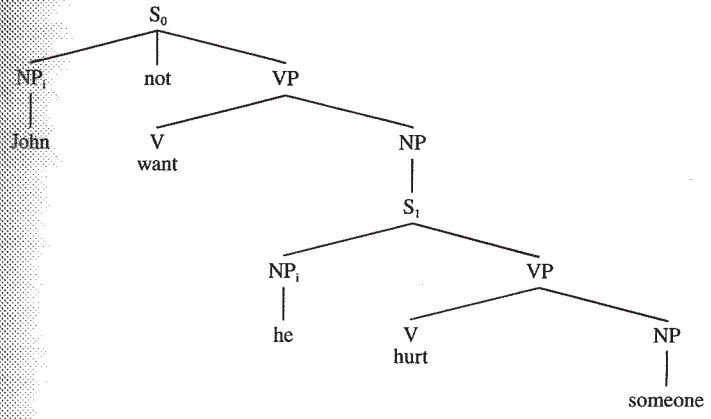



Note that the condition for the application of Equi-NP-deletion to $S_0$ (namely, that the subject of $S_1$ be coreferential with the subject of $S_0$) is not met in (1b) but is met in (1c), that is, the condition for the application of Equi-NP-deletion to $S_0$ comes to be fulfilled through the application of Passive to $S_1$. In such cases we will speak of one rule application feeding another; here, the application of Passive to $S_1$ feeds the application of Equi-NP-deletion to $S_0$, or, for short, $\text{Passive}_1$ feeds $\text{Equi-NP-deletion}_0$.

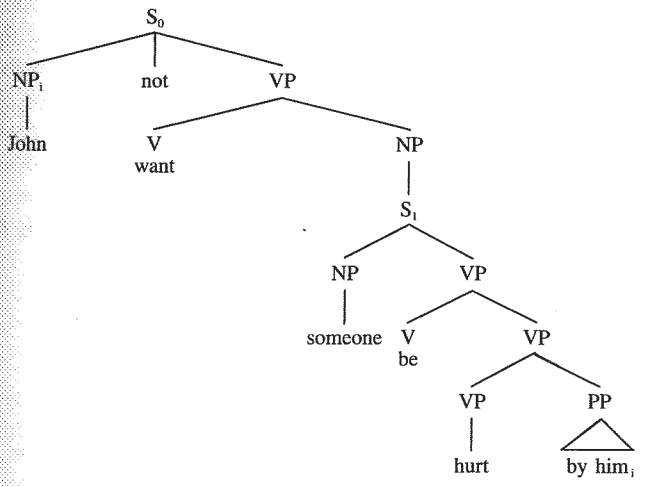
Consider now the derivation of (2a) that conforms to the analysis of chapter 5, in which the deep structure is (2b) and Passive applies to $S_1$ to yield the output (2c):[2]

(2) a. John doesn't want anyone to be hurt by him.

b.




c.




Here, although the condition for the application of Equi-NP-deletion is met in (2b), it is not met in (2c), that is, the application of Passive to $S_1$

destroys the conditions for the application of Equi-NP-deletion to $S_0$, and Equi-NP-deletion indeed cannot be applied to $S_0$ in a derivation that involves the step from (2b) to (2c). In such cases we speak of one rule application **bleeding** another. Here, $\text{Passive}_1$ bleeds $\text{Equi-NP-deletion}_0$.

The terms "feed" and "bleed," introduced in Kiparsky (1965), are part of a system of terminology for relations between rule applications in a given derivation, not only rule applications that actually occur in the derivation but also those that fail to occur, such as the application of Equi-NP-deletion to $S_0$ in (2).[3] These terms will figure prominently in the remainder of this chapter, which is concerned with how transformations may interact in derivations, i.e., with working out the principles that determine what feeding and bleeding relationships are possible in derivations.

Two caveats before delving into some aspects of that question. First, it is important to recognize that feeding and bleeding are relationships not between rules but between rule applications in a given derivation. For example, in the derivation of (3) there are two applications of Passive, and it is only its application to the most deeply embedded S that feeds the application of Equi-NP-deletion:

(3) It is believed by many people that John wants to be admired by everyone.

Linguists do occasionally speak of one rule feeding another, but that usage is derivative from the usage just described: if a linguist says, say, that Passive feeds Agreement, what he probably means is that whenever both rules are potentially applicable to a given domain, the application of the one to that domain feeds the application of the other. Second, I will use the term "feed" not only in cases where the application of one rule makes the other rule applicable at all, but also in cases where the application of the one rule affects what the other rule does in the given case. Thus, in the derivation of (4a) we will say that $\text{Passive}_1$ feeds $\text{Raising}_0$ (calling the embedded S $S_1$ and the main S $S_0$) because, though Raising is applicable to $S_0$ even when Passive is not applied to $S_1$, as in (4b), Passive affects whether it is *Holmes* or *the solution* that $\text{Raising}_0$ makes the derived subject:

(4) a. The solution seems to have been found by Holmes.
b. Holmes seems to have found the solution.

## *b. Schemes of Rule Interaction*

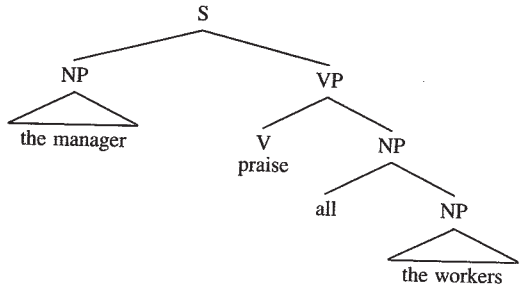
We have already given a rough sketch of a principle that will form a large part of the answer to the question of how rules may interact in a derivation, namely, the principle of the cycle. That principle says (at least in its pristine form—an important qualification will have to be added in §6c) that



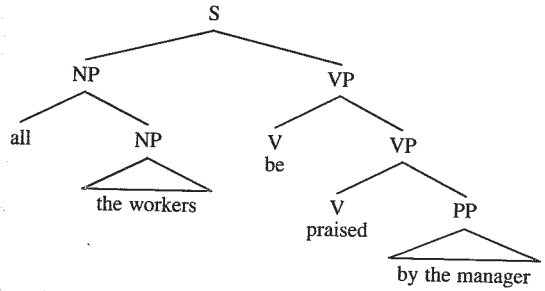
when one domain is contained in another, applications of rules to the former domain precede applications of rules to the latter, so that, e.g., in the examples given in the last section, any application of rules to $S_1$ will precede any application of rules to $S_0$. In the two cases taken up in the last section, that principle implies exactly the interaction between Passive and Equi-NP-deletion that we observe: in both cases, the input to $\text{Equi}_0$ is not the deep structure but rather the structure resulting from application to $S_1$ of whatever rules are going to apply to it (here, only Passive).

The principle of the cycle tells us how an application of a rule to $S_1$ will interact with an application of a rule to $S_0$, but it tells us nothing about how it will interact with an application of another rule to $S_1$. For example, it tells us nothing about the feeding and bleeding that can be observed in cases where there appears to be only one S, as in (1a), with the presumable deep structure (1b), where Passive feeds Quantifier-float, i.e., Passive applies to a structure in which the conditions for Quantifier-float are not met (since *all* can be floated only from a subject NP) to yield an output (1c) to which Quantifier-float can apply:

(1) a. The workers all were praised by the manager.
b.



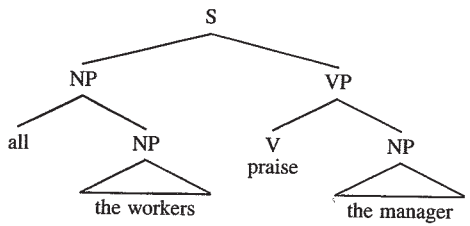

c.



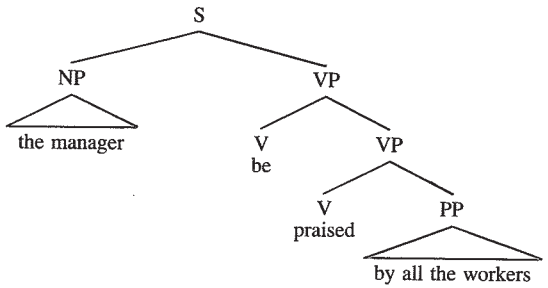

Likewise, in (2a), with the presumable deep structure (2b), Passive bleeds Quantifier-float:

(2) a. The manager was praised by all the workers.
a′. *The manager {all was / was all} praised by the workers.

b.




c.




What is of interest here is not that (2a) is acceptable but that (2a′) is unacceptable, i.e., with a deep structure as in (2b), Passive can apply to yield the structure (2c), to which Quantifier-float is inapplicable, but it must be kept from applying to a structure to which Quantifier-float has already applied, so as to exclude a derivation of (2a′).[4]

Let us list some ways in which transformations applying to a single domain might conceivably interact and see which of them are consistent with the interactions that we observe in (1)–(2) and other examples. A first possibility is that the application of transformations to a given domain is **simultaneous**, i.e., that the given structure is converted in one fell swoop into a structure that differs from the first one in all respects specified by all the transformations whose conditions for application are met in the first structure. This possibility fails miserably in (1)–(2): it implies that there can be no feeding such as is observed in (1), since only rules whose conditions of application were met in (1b) would apply, and it implies that one or other version of (2a′) should be possible, since (2a′) is what would result if Passive and Quantifier-float, whose conditions for application both are met in (2b), both applied.

A second possibility is that the application of transformations to any given domain is **random sequential**, in the sense that the application of transformations to the domain is a sequence of steps, each consisting of the application of **any** transformation whose conditions for application are met in the output of the preceding step. This second possibility provides derivations for (1a) and (2a) but incorrectly also provides a derivation for (2a′).



Random sequential application allows, e.g., for the desired derivation of (1a), since the conditions for application of Passive are met in the deep structure (1b), and the conditions of the application of Quantifier-float are met in the output of Passive. In the case of (2b), since the conditions of application for both Passive and Quantifier-float are met in the deep structure, random sequential application allows the derivation to begin with an application of either transformation. If the derivation begins with an application of Passive, all is well, since the conditions for Quantifier-float are not met in the output, and (2a) results. However, if the derivation begins with an application of Quantifier-float, there is nothing to prevent Passive from then applying to the output of Quantifier-float, in which case the unacceptable (2a′) results.

A third possibility is to posit what has generally been accepted in transformational grammar, namely a fixed **order** of application of the transformations, with each transformation allowed to apply only when its "turn" is reached. Note that since we wish to allow derivations in which Quantifier-float applies to the output of Passive (1a) but exclude derivations in which Passive applies to the output of Quantifier-float (2a′), if there is a fixed order of application in which Passive precedes Quantifier-float, we will get exactly the derivations that we want in these cases. While this third possibility is thus apparently consistent with the facts that we have taken up so far, it also appears to allow for an extraordinary amount of variation in language, namely, variation with regard to the order that is imposed on a given set of rules, and there is little evidence that any such variation actually exists. For example, as it stands, the third proposal for rule interaction seems to imply that there could be a dialect of English that was just like the variety of English under discussion except that Passive and Quantifier-float were in the opposite order, which would imply that (2a′) was acceptable and (2a) unacceptable in that dialect. No such dialect is known to exist, and it is very doubtful that such a dialect is even possible.

The three schemes of rule interaction just sketched of course come nowhere near exhausting the possibilities, since there are for example other schemes in which these possibilities are combined with each other and/or supplemented by additional principles. For example, possibilities two and three are combined in the suggestion of Anderson (1969, 1974) that rules are subject to a partial rather than a total ordering and that their application is sequential but only partially random, in that the partial ordering must be respected. And possibilities one and two are combined in the detailed scheme developed by Koutsoudas, Sanders, and Noll (1974), for whom a derivation consists of a sequence of steps in which the rules whose conditions for application are met in the output of the preceding step are applied simultaneously. Among the principles that have been combined with all

three of these schemes of rule application is the **Elsewhere Principle** (due originally to Panini, ca. 500 B.C., and reintroduced in Anderson 1969, Kiparsky 1973, and Koutsoudas, Sanders, and Noll 1974): when the conditions of application for one rule are a special case of those for another rule, the latter (more general) rule is inapplicable in those cases in which the conditions for the former (more specific) rule are met, i.e., specific rules preempt the application of general rules.

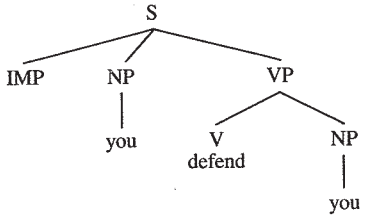
No attempt will be made here to survey the full range of conceivable schemes of rule interaction and their implications. Readers who are interested in pursuing this question in all the detail that it deserves are referred to Pullum (1976), the definitive treatment of this topic. For the remainder of this book, I will adopt a fairly agnostic position on how rules applying to the same domain interact. I will argue, in fact, that there are far fewer cases in which rules apply to the same domain than is generally alleged, so that the question of how rules applying to the same domain interact arises much less frequently than is generally thought.

In a large proportion of the cases that are alleged to require the fixed ordering of transformations of proposal three, a perfectly satisfactory alternative analysis is available in which the domain to which one of the transformations applies is not identical to but is properly contained in the domain to which the other applies, and the principle of the cycle predicts correctly how the two transformations interact. To take one especially popular putative example of a fixed order of application of transformations, consider the interaction observed in examples like (3) between the transformations responsible for the reflexive pronoun ("Reflexivization") and for the absence of the subject ("Imperative subject deletion" = ISD):

(3) Defend yourself/*you.

According to a widely accepted analysis, (3) has a deep structure like (4), involving a symbol that marks the sentence as imperative, and the derivation involves applications of Reflexivization, which converts into a reflexive pronoun a NP that is coreferential with the subject of its clause,[5] and of ISD, which deletes a second person subject in the presence of the imperative marker:

(4)

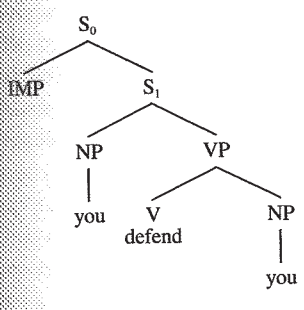





Note that with the deep structure (4) and the assumed transformations, the rules cannot apply in random sequential fashion. The conditions for the application of both Reflexivization and ISD are met in (4), and of the two derivations that random sequential application would allow, one yields the desired *Defend yourself* (namely, the derivation in which Reflexivization applies to (4) and then ISD applies to the output of Reflexivization), but the other yields the unacceptable **Defend you* (namely, the derivation in which ISD applies to (4), yielding an output to which Reflexivization is no longer applicable, since the subject NP to which the second *you* is coreferential has been deleted). This result argues against random sequential application of the assumed transformations but is consistent with a fixed order of application in which Reflexivization precedes ISD.[6]

An extremely minor change in the proposed deep structure changes this situation drastically, however. Suppose that the "IMP" appears not as a constituent of but as a sister of the S in question:

(5)




With respect to almost everything, (5) and (4) are no more than minor typographical variants of one another. However, the principle of the cycle yields many implications when applied to (5), though none when applied to (4). In particular, the principle of the cycle, in conjunction with deep structures like (5), implies that ISD (which is sensitive to the presence of "IMP" and whose conditions for application thus could not be met in $S_1$ but only in $S_0$) can apply to structures derived by any of the transformations that affect declarative sentences (i.e., any transformations that would apply to the domain $S_1$ in (5)), and thus that it is possible to form imperatives in which a subject *you* is deleted from a S that has undergone Passive or *Tough*-movement (or Reflexivization). As (6) shows, this implication is correct:

(6) a. Don't be surprised by what I'm going to say.
b. Please be easy for them to deal with.

Where the approach that posits deep structures like (4) must specify an ordering of transformations in which Reflexivization, Passive, and *Tough*-

movement are ordered before ISD, an approach with deep structures like (5) is able to predict their interactions without assuming any fixed ordering of the transformations, i.e., the approach with (4) must make do with a description of how these transformations interact, while the approach with (5) provides an explanation of their interaction.

In the following chapters, I will pursue a research program of seeking analyses that exploit the principle of the cycle to predict the interactions of the transformations that are posited. I will regard the program as successful if the analyses that it leads to are on other grounds desirable, for example, if the rules can be stated in their pristine forms and the deep structures posited accurately represent the meanings of the sentences. (On the last criterion, (5) is preferable to (4): semantically, *Defend yourself* is not a combination of IMP with two separate elements *you* and *defend you* but is rather an order or request that you defend yourself, i.e., it is a combination of the meaning of "IMP" with the meaning of the posited $S_1$.) I will regard the program as a failure if it forces one to accept objectionable analyses, e.g., an analysis in which the statement of the conditions under which some rule applies requires complications that are avoided under alternative analyses, or one that misrepresents the meanings of the relevant sentences. An analysis will be presented in §18a, incidentally, that will allow the interaction between Passive and Quantifier-float noted in (1)–(2) to be predicted from the principle of the cycle, namely, an analysis in which quantified NPs in deep structures are outside the Ss that they appear in in surface structure, so that the principle of the cycle will predict that applications of transformations to the "matrices" (such as *the manager praised x*) with which a quantified NP is combined will precede applications of transformations to structures that contain the quantified NP. Until we work out that analysis, however, we must continue to regard (1)–(2) as a serious problem that must be solved by any approach to rule interaction that does without a fixed order of application of transformations.

### c. *The Cycle*

By all rights, this section should consist mainly of arguments for accepting the principle of the cycle. The section will in fact be mostly devoted to discussion of what principle of the cycle should be adopted, and indeed only one real attempt at an argument for accepting such a principle will be made.

The reason for the absence of such arguments in this section is that to give a solid argument for something, one must be clear about what it is to be an argument against, and there happens to be no alternative position on the relationship between rule applications on different domains that has

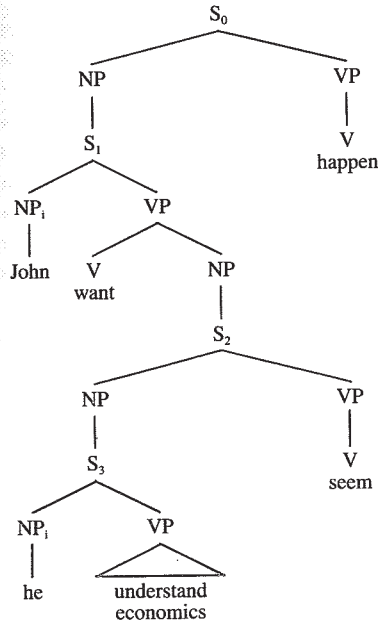


achieved even the status of a respectable straw man, let alone that of a serious contender to the cyclic principle. Accordingly, the best arguments that can be offered for the cyclic principle must to some extent be arguments from ignorance: demonstrations that the principle successfully accounts for the observed rule interactions, especially those that are fairly intricate, combined with statements that no alternative principle is known to account satisfactorily for how rules apply to structures that contain multiple domains of rule application. Rather than presenting such arguments here, perhaps the most useful thing I can do is simply urge the reader to consider some of the more involved derivations given in this book and ponder how the various transformations could be made to interact correctly without reliance on the principle of the cycle.

There are many supposed arguments for the cyclic principle that rely on the fact that certain sentences demand derivations in which an application of one transformation is "sandwiched" between applications of another transformation in a way that the cyclic principle allows. For example, in the derivation relating (1a) to the deep structure (1b), $\text{Raising}_2$ feeds $\text{Equi-NP-deletion}_1$, which in turns feeds $\text{Raising}_0$, and in the derivation relating (2a) to the deep structure (2b), $\text{Passive}_1$ feeds $\text{Raising-to-object}_0$, which in turn feeds $\text{Passive}_0$:

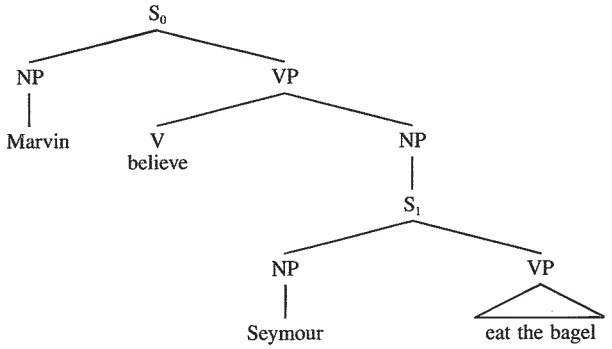
(1) a. John happened to want to seem to understand economics.

b.



(2) a. The bagel is believed by Marvin to have been eaten by Seymour.

b.




However, the necessity of such derivations provides an argument for the cyclic principle only to the extent that one can exclude other schemes for rule interaction that also allow such derivations. There is one obvious logical possibility for a scheme of rule interactions in which such derivations are allowed, namely, the possibility that the sequence of steps in a derivation is unconstrained both with regard to what transformation applies in any step and what domain it applies to in that step. Any serious argument for the cyclic principle will have to rest on facts relating not only to what derivations it allows but also to what derivations it excludes.

Let us then, before turning to the questions with which this section will be mainly concerned, give one argument for the cyclic principle, an argument showing that the principle correctly excludes certain derivations that an alternative scheme (namely, that of totally random rule application) allows. The cyclic principle (but not the straw man alternative with which it is being compared) excludes derivations in which application of a rule to a domain feeds or bleeds application of a rule to a domain properly contained within the latter domain. The possibility of such a derivation arises when we consider possible interactions between Reflexivization and Raising-to-object (RO). In the oversimplified sketch of Reflexivization given in the preceding section, it was stated that the antecedent of a reflexive pronoun must be the subject of the S to which Reflexivization applies. This statement of the conditions for Reflexivization is in fact overly restrictive, since a reflexive pronoun in English can have an object NP as its antecedent (3a), but in any event Reflexivization must be restricted so that a NP in a VP cannot be the antecedent of a reflexive pronoun in a lower VP (3b) (cf. (3b′), where the antecedent is presumably not the object of *force* but the underlying subject of the infinitive):

(3) a. Mary asked $\text{John}_i$ about $\text{himself}_i$.



b. Mary promised $\text{John}_i$ to defend $\text{him}_i$/*$\text{himself}_i$.

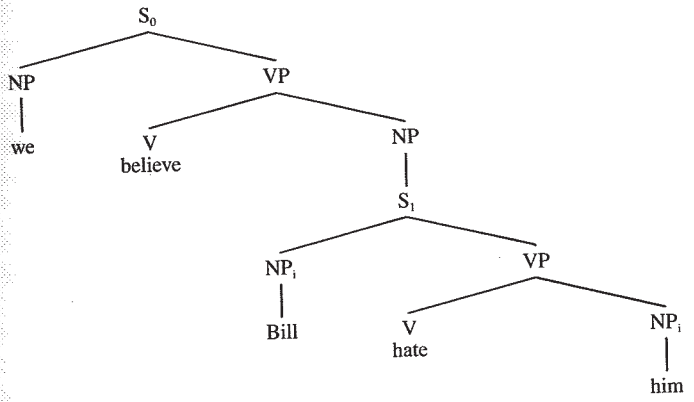
b′. Mary forced $\text{John}_i$ to defend $\text{himself}_i$.

Now consider the possibilities for RO and Reflexivization (4a–a′) in the structure (4b):

(4) a. We believe $\text{Bill}_i$ to hate $\text{himself}_i$.

a′. *We believe $\text{Bill}_i$ to hate $\text{him}_i$.

b.




If the application of transformations were totally random, there would be nothing to rule out a derivation in which $\text{RO}_0$ bleeds $\text{Reflexivization}_1$: note that application of RO to $S_0$ would move *Bill* into such a position that, as in (3b), it could not serve as antecedent in an application of Reflexivization. The cyclic principle, however, rules out such a derivation: Reflexivization would apply to $S_1$ before RO had its chance to apply to $S_0$. Thus, random application of transformations incorrectly allows derivations for both (4a) and (4a′), whereas the cyclic principle correctly allows only a derivation for (4a).[7]

Having given one argument for "the" cyclic principle, let us turn to the question of what cyclic principle we should adopt. In what has been said so far in this book about the principle of the cycle, we have sidestepped the two important issues of what exactly the domains are to which transformations apply, and of whether the cyclic principle governs the application of all transformations or only of certain kinds of transformations.

With regard to the first issue, we have spoken so far as if the domains to which transformations applied were precisely the Ss that figure in the underlying structure of a sentence. It is far from obvious that only Ss can be domains of application of transformations. For example, note that in the application of Equi-NP-deletion in the derivations of sentences like (5), nothing outside the VP of the main clause plays any role in deter-

mining whether the transformation is applicable, or on the effect of the transformation:

(5) John forced Mary$_i$ [(she$_i$ → ∅) to empty the garbage].

There is thus nothing in principle to prevent one from taking the domain to which "object-controlled Equi-NP-deletion" applies to be not a S of the shape (6a) but a VP of the shape (6b):

(6) a. [$_S$ NP [$_{VP}$ V NP$_i$ VP]]]
b. [$_{VP}$ V NP$_i$ [$_S$ NP$_i$ VP]]

Likewise, there is nothing in principle to prevent one from taking the Dative-movement transformation that is often posited in the derivations of sentences like (7a) (cf. (7a′), in which the transformation has not applied)* to be a VP of the shape (7c) rather than a S of the shape (7b):

(7) a. Mary sold John the candlestick.
a′. Mary sold the candlestick to John.
b. [$_S$ NP [$_{VP}$ V NP [$_{PP}$ to NP]]]
c. [$_{VP}$ V NP [$_{PP}$ to NP]]

There are also several transformations that could be taken to have a NP rather than a S as domain of application, e.g., the optional replacement of *the* by a dependent NP that will be proposed in chapter 12 as part of the account of NPs like those in (8):

(8) a. the denunciation of Smith by Jackson
b. Jackson's denunciation of Smith

To make the cyclic principle precise, we must thus specify what exactly the possible domains of application of transformations are (only Ss? only Ss and NPs? VPs as well? constituents of all categories?) and clarify whether we should retain our tacit assumption that all domains of application are "cyclic domains," e.g., whether we should interpret the cyclic principle as implying that the application of the NP movement transformation of (8) must precede the application of any transformations to the S in which the NP is contained.

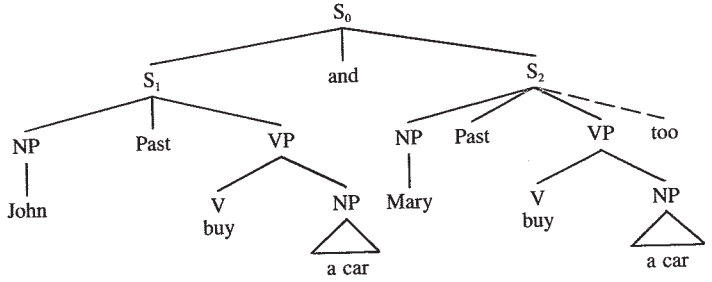
Surprisingly, very little attention has been devoted to these questions, with virtually all published discussion of domains and the cycle restricted to the question of whether NPs as well as Ss are cyclic domains. Only in Williams (1974) is anything approaching the proposition that all constituents are cyclic domains taken up seriously, though as I pointed out in McCawley 1977, when viewed from a certain perspective, Montague grammar can also be interpreted as regarding all constituents as cyclic domains.



Since the existing literature thus provides little on which one can base an answer to these questions, I will, provisionally, simply adopt the answer that best fits my research program of making the cyclic principle do the greatest amount of work, i.e., I will provisionally assume (i) that every syntactic constituent is a cyclic domain (in the sense that the cyclic principle will be taken as implying that applications of transformations to any constituent must precede applications of transformations to anything that the given constituent is contained in, irrespective of the categories of the two constituents) and (ii) that the domain to which any transformation applies will be the smallest constituent containing all material relevant to its application, so that e.g., the domain to which Dative-movement or Object-controlled Equi-NP-deletion applies will be a VP and the domain to which the NP movement of (8) applies will be a NP. Note that these two answers do in fact maximize the amount of work done by the cyclic principle: they maximize the class of cases in which the cyclic principle will say which of two applications of transformations must precede the other.

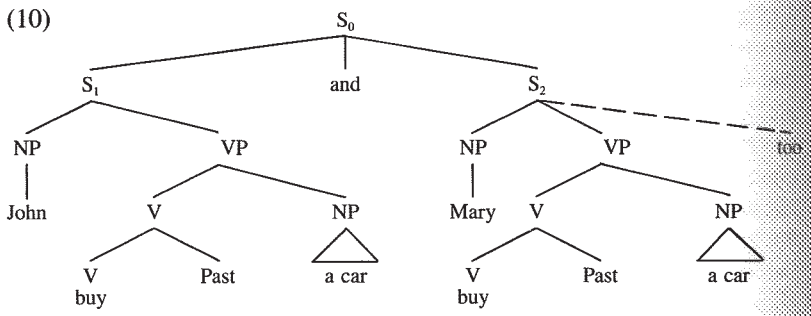
Let us turn now to the question of whether all transformations are subject to the principle of the cycle. It turns out that some transformations must be exempted from the cyclic principle, as can be seen easily from the following example. Suppose that all transformations applied according to the cyclic principle. Then there would be no derivation relating sentences like (9a) to the presumable deep structure (9b):

(9) a. John bought a car, and Mary did too.
b.




In $S_1$ and $S_2$ the conditions for Tense-hopping are met.[9] Thus, the application of Tense-hopping to $S_1$ and $S_2$ would precede any application of a transformation to $S_0$. Since $S_0$ is the smallest constituent containing the two occurrences of *buy a car,* that means that Tense-hopping would attach Past to the two occurrences of *buy* before VP-deletion had a chance to apply. But then VP-deletion could not delete *buy a car,* the stage of the derivation to which it might apply would no longer contain two occurrences of *buy a car* but only two occurrences of *buy-Past a car:*

(10)




Thus, if VP-deletion is simply the deletion of a repeated VP, and (9a) has (9b) as its deep structure, the application of Tense-hopping to $S_1$ and $S_2$ will have to follow the stage of the derivation at which VP-deletion gets its chance to apply to $S_0$.

This is an anomaly from the point of view of the cyclic principle in its pristine splendor, and either of the two transformations involved here could in principle be held responsible for the anomaly: one could say that the derivation from pristine cyclicity is either that VP-deletion must apply earlier than it "ought to" or that Tense-hopping must apply later than it ought to. More precisely, one could propose either that VP-deletion is **precyclic** (i.e., that it applies prior to the application of all transformations that conform to the cyclic principle) or that Tense-hopping is **postcyclic** (i.e., that it applies after the application of all transformations that conform to the cyclic principle). I will reject the former of these possibilities and adopt the latter, not because of any really clearcut arguments against the one or in favor of the other, but simply on the grounds that taking VP-deletion to be precyclic would subvert the entire program of exploiting the cyclic principle. If VP-deletion were precyclic, then so also would be the vast bulk of the transformations taken up above or in the following chapters, and the cyclic principle would tell us nothing about the interactions of the transformations that we are most interested in. To see this, note that all transformations that feed a precyclic transformation must themselves be precyclic, and that all transformations that change subjects (e.g., Passive, *Tough*-movement, and *There*-insertion) feed VP-deletion:

(11) a. Tom was arrested by the FBI, and Dick was too.
b. Bill is hard for us to get along with, but Agnes isn't.
c. We said that there was a leak in the boat, and in fact there was.

By contrast, taking Tense-hopping to be postcyclic does not appear to force one to take any other transformations to be postcyclic, and thus under that alternative the cyclic principle continues to have rich implications for the



interaction of transformations. I accordingly rate the second horn of our dilemma the less objectionable one and will henceforth adopt the conclusion that we must weaken the cyclic principle to the extent of admitting a set of postcyclic transformations, with Tense-hopping among them.

The argument just given is in fact valid only for the case of Tense-hopping that combines tenses with main verbs;[10] it is easy to see that it does not carry over to the combining of tenses with auxiliary verbs, in view of the lack of parallelism between (9a), where a tense is "stranded" by the deletion and consequently appears attached to *do*, and (12), where the tense cannot be stranded but must appear in combination with the auxiliary verb:

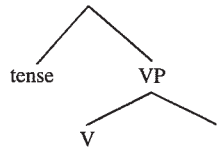
(12) a. Tom was arrested by the FBI, and Dick was/*did ∅ too. (= (11a))
b. The FBI has bugged my office, and the KGB has/*does ∅ too.

It is thus necessary to divide what we have so far lumped together under the name "Tense-hopping" into two transformations: a postcyclic one that combines a tense with a main verb (and for which I will retain the name **Tense-hopping**), and a cyclic one that combines a tense with an auxiliary verb (henceforce to be called **Attraction-to-tense**).

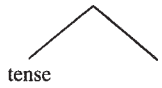
Besides Tense-hopping (in the narrower sense in which we are now using that term), the transformations that are known to be postcyclic are *Do*-support, the deletion of prepositions before complement Ss, the deletion of *for* before *to*, and the irregular deletion of *for* after *want*. Pullum (1976) makes some tentative suggestions as to what kinds of transformations can in principle be postcyclic and makes one proposal that I find particularly attractive, namely that a postcyclic transformation must be **local** in the sense (Emonds 1976) that it can involve only two constituents, which must be adjacent with regard both to linear order (nothing intervenes between them) and to constituent structure (they either are sisters or are aunt and niece).[11] The four transformations listed above meet this description.

Pullum's proposal heavily restricts the possibilities for interactions among postcyclic transformations: the smaller the parts of a syntactic structure that can be involved in the application of two transformations, the less opportunity there is for them to interact in any way. There is indeed only one case that I know of where the possibility of interactions among postcyclic transformations even arises, namely that of Tense-hopping and *Do*-support. *Do* must be adjoined to precisely those tenses that cannot undergo Tense-hopping. But that interaction is predictable if one assumes the Elsewhere Principle (§6b). Tense-hopping applies to the configuration (13a) and *Do*-support to the more general configuration (13b),[12] and thus by the Elsewhere Principle the application of *Do*-support is excluded in those cases where Tense-hopping applies:

(13) a.

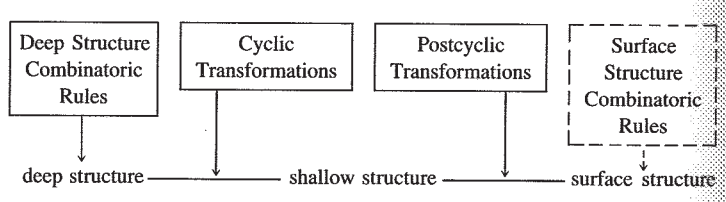



b.




The level of structure that is the output of the cyclic transformations and the input to the postcyclic transformations is often called **shallow structure**. We will thus henceforth assume the following picture of levels and sets of transformations:[13]

(13)




In this diagram, each box represents a set of rules, and the arrow coming out of each box indicates what that set of rules constrains. The arrow may point either to a structural "level," in which case the rules constrain what structures can occur at that level, or to a line connecting two levels, in which they constrain the relationship between the two levels, e.g., the cyclic transformations specify what deep structures may correspond to what shallow structures.

### *d. Root Transformations*

This section is devoted to a notion that, as a result of some quirks of history, is often confused with the notion of postcyclic transformation. It is taken up at this point largely because I believe that the reader will be less likely to confuse the two notions if he sees them taken up together in a context that makes clear that they not only are distinct but are indeed unrelated.

Of the transformations that we have taken up so far, exactly one, namely Inversion, has been subject to the restriction that it applies only in main clauses (i.e., independent questions), not in subordinate clauses (i.e., dependent questions):

(1) a. What did George talk about?
b. Ann asked what George talked about.
b′. *Ann asked what did George talk about.



By contrast, the other transformations that we have taken up so far have applied without regard to whether the domain to which they were potentially applicable made up the whole sentence or was embedded in a larger structure. For example, the applicability of Passive in such a sentence as (2a) is not affected when the structure underlying that sentence is embedded in a larger sentence like (2b):

(2) a. The bomb was planted by terrorists.
b. The newspapers report that the police are convinced that the bomb was planted by terrorists.

The term **root transformation** has been introduced by Emonds (1970, 1976) to refer to transformations whose application is restricted, as in the case of Inversion, to main clauses.[14]

The two postcyclic transformations that we have taken up so far, namely Tense-hopping and *Do*-support, clearly are not root transformations:

(3) a. The State Department refused to confirm the rumor that an ambassador had told a reporter from *Le Monde* that George Washington owned slaves.
b. It is widely recognized that George Washington owned slaves, but the State Department refuses to confirm the rumor that an ambassador had told a reporter from *Le Monde* that Thomas Jefferson did ∅ too.

In (3a), a tense marker gets attached to the following verb just as well in a deeply embedded subordinate clause as it does in a main clause, and in (3b) we see that deletion of the verb to which a tense would otherwise be attached can affect a deeply embedded clause, and that clause can be the domain to which *Do*-support applies, inserting a *do* to serve as bearer of the tense.

Since Tense-hopping provides such a clear example of a postcyclic transformation that is not a root transformation, it is hard to see how the notions could be confused. Such confusion does however exist, as a result of terminological practices in which different linguists in the late sixties and early seventies distinguished "cyclic" transformations from some other kind of transformations, not appreciating that the "other kind" differed from one group of linguists to another. The "other kind" was postcyclic transformations for one group, root transformations for a second group, and something called "last-cyclic transformations" (which is of at most historical interest and will not be taken up in this book) for a third group. Confusion among these three terms was engendered by linguists'

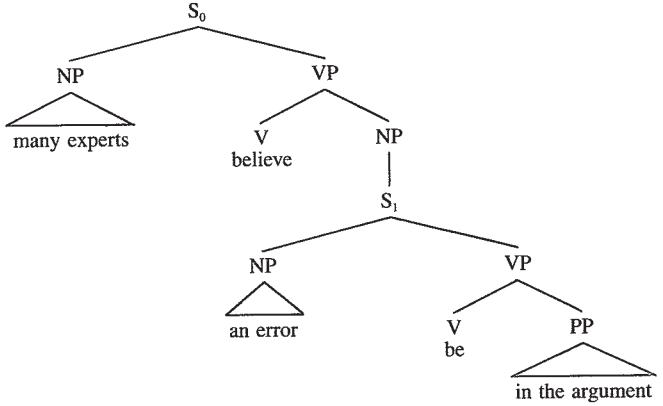
failure to make clear which of (at least) three distinctions they had in mind when they argued about whether such-and-such transformation is "cyclic."

Both the notion of postcyclic transformation and that of root transformation will play a role in the chapters that follow, and each notion will be given the appropriate name wherever it arises.

### *e. Obligatory and Optional Transformations*

In the preceding sections, a distinction has implicitly been drawn between **optional** and **obligatory** transformations. Passive provides a fairly clear example of an optional transformation, in the sense that when a derivation contains a structure that meets the conditions for the application of Passive, the derivation is normally allowed but not required to have an application of Passive, e.g., both (1a) and (1a′) have the deep structure (1b), and their derivations (involving *There*-insertion$_1$ and RO$_0$) are identical up to the point where Passive applies in the one derivation but not in the other:

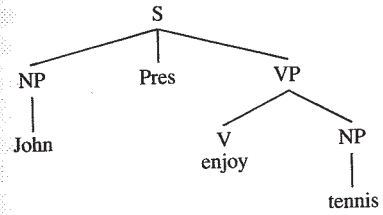
(1) a. There is believed by many experts to be an error in the argument.
a′. Many experts believe there to be an error in the argument.
b.




Tense-hopping provides an equally clear example of an obligatory transformation, in the sense that when a derivation contains a structure that meets the conditions for Tense-hopping, the derivation is required to have an application of that rule. For example, the deep structure (2b) can figure in a derivation of (2a), involving an application of Tense-hopping, but not in a derivation of (2a′) or (2a″), to give the only apparent things that one might get without an application of Tense-hopping:[15]



(2) a. John enjoys tennis.
a′. *John does enjoy tennis.
a″. *John -s enjoy tennis.
b.




The characterizations just given of "optional" and "obligatory" are far too imprecise to be of much use. Let us attempt to say more precisely what the terms mean. Note that the characterization of "optional" just given is in conflict with the common practice of calling a rule optional even when there are instances in which some factor extraneous to the conditions of application of the rule happens to exclude derivations in which there is no application of the rule. For example, corresponding to (3a) and (4a), whose derivations involve applications of Passive, there are no sentences like (3b) and (4b), whose derivations are like those of (3a) and (4a) except for the nonapplication of Passive:

(3) a. John forced Mary to be examined by Dr. Grimshaw.
b. *John forced Mary for Dr. Grimshaw to examine her.
(4) a. Mike's being bothered by this is ridiculous.
b. *This's bothering Mike is ridiculous.

Here, failure to apply Passive results in violation of one or other constraint on derivations: in (3), violation of the constraint that the complement of *force* must undergo Equi-NP-deletion, and in (4), violation of the constraint that in surface structure all words must be in forms that their morphology allows (*this* has no genitive form).

We will speak of Passive as being optional notwithstanding the existence of cases where one "has to apply Passive," in that what is wrong with the derivations that yield (3b) and (4b) is the violation of the constraints just mentioned, not the nonapplication of Passive that gave rise to those violations. If application of Passive is the only way to avoid these violations in derivations with the given deep structures (and one must make a systematic search of the syntactic devices of English before one is in a position to say that it is), that is true only by accident. There are many cases in which there are two or more optional transformations, the application of any of which could suffice to avoid violation of a particular con-

straint. For example, the oddity of (5a–a′) (which reflects the fact that a conjoined NP with a personal pronoun as one conjunct and a nonpronoun as the other conjunct has no genitive form) could be avoided not only by passivizing the embedded S (5b) but also by "dislocating" the offending NP, as in (5c):

(5) a. *You and Nancy's annoying Tom is ridiculous.
a′. *Your and Nancy's annoying Tom is ridiculous.
b. Tom's being annoyed by you and Nancy is ridiculous.
c. You and Nancy, your annoying Tom is ridiculous.

To call a transformation optional is simply to say that the rule in itself allows both for derivations in which its conditions for application are met and it is applied, and for derivations in which its conditions for application are met but it is not applied. This is completely neutral with regard to whether there is anything else that in a given case might rule out one or both of these alternatives. This may be a good point at which to remind the reader of something that has largely been lost sight of in this chapter, namely, that a grammar is **not** a procedure for constructing sentences by assembling a deep structure and converting that deep structure into a surface structure, and the notion of derivation that is assumed in this book does not commit one to the latter conception of a grammar.

In the conception of a grammar assumed in this book, neither deep structure nor surface structure, nor any other linguistic level, has any "priority" in relation to other levels. A grammar specifies what derivations a language allows, by means of rules of which some (the "combinatoric" rules) specify what structures are allowed on particular levels, and others (the "transformations") specify how the stages in a derivation may or must differ from one another. In speaking, or in understanding speech, one mentally constructs a derivation,[16] assembling structures on all levels simultaneously, in such a way that (if all is successful) the structures and the relations among them conform to the rules of the grammar. An optional transformation does not direct the language user to flip a mental coin in choosing between two alternative structures; rather it simply makes available to him two possibilities that he is free to use as he puts together syntactic and semantic structures, and it carries no guarantee that even one of those two possibilities, let alone both of them, will be of any use to him in a given case.

The notion of "obligatory transformation" is fairly unproblematic in orthodox transformational grammar, where a fixed order of application of transformations is assumed: if the conditions for the application of transformations are met at the point in the derivation where the rule gets its chance to apply to the given domain, it must apply, in the sense that the derivation is excluded if the rule is not applied. The notion requires some further explication, however, when it is combined, as it is here, with a picture of rule interaction that does not assume a fixed order of application of transformations. Under either conception of rule interaction, the notion of obligatoriness is constrained by the notion of "getting a chance to apply" that is in force. For example, in view of the principle of the cycle, what is directly relevant to the application of an obligatory transformation to a given domain is not whether in deep structure that domain meets the conditions for application but whether it meets them in a stage of the derivation that follows the application of any cyclic transformations to domains contained with it. However, in the approach that is followed here, we have to contend with the possibility that more than one transformation might "get a chance to apply" at a given stage of a derivation. Nothing that has been said so far excludes, for example, the possibility that a single structure might meet the conditions for application of two or more obligatory transformations. Perhaps (i) "obligatory" should be interpreted according to a literal understanding of "must apply whenever its conditions for application are met," which would imply that all the obligatory transformations whose conditions for application are met by a given structure apply simultaneously. Alternatively, (ii) one might require that rules apply one at a time and allow any one of the obligatory transformations whose conditions for application are met at a given stage of a derivation to be the rule that applies at that stage. Or (iii) one might merely require that at the end of the application of the cycle to a given domain there be no obligatory transformation whose conditions for application are met by that domain. Each of these three possibilities restricts the class of admissible derivations less than does the preceding one; for example, (ii) is less restrictive than (i) because it allows for the possibility that one obligatory transformation might bleed another one, and (iii) is less restrictive than (ii) because it allows for the possibility that an optional transformation might bleed an obligatory one. I know of few instances in which a structure meets the conditions for application of more than one obligatory cyclic transformation (S, on p. 224, to which both Raising and Tense Replacement apply, is one such case) and none in which the choice among (i)–(iii) matters. I will accordingly leave this question unresolved until cases arise where its answer makes a difference

It is worth noting that whether a given transformation is obligatory or optional can depend on what other rules the grammar is to contain and on how exactly one formulates its conditions of application. A transformation that is **de facto** obligatory, in the sense that no derivations are well formed in which its conditions for application are met but it is not applied, need not

be classed as obligatory in the technical sense if the derivations in which it is not applied can be held to be ill formed for some other reason such as, say, that the surface structures violate some combinatoric constraint on surface structure. For example, while *Do*-support is de facto obligatory, the ill-formedness of derivations in which it is applicable but is not applied could be attributed to violation of the morphological requirement that tenses in surface structure be attached to verbs. If *Do*-support were not applied in the derivation of (6a), a surface structure would result in which a tense was in a configuration that its morphology did not allow:

(6) a. Sarah doesn't like Proust.
b. *Sarah -sn't like Proust.

There is thus no apparent obstacle to treating *Do*-support as technically optional, notwithstanding the fact that it indeed "must apply" when it can.

One can even seriously entertain the possibility that no transformations need be regarded as obligatory: perhaps in all cases where nonapplication of a transformation is excluded, nonapplication results in violation of some independently necessary constraint. With the analyses adopted here, it seems to be impossible to maintain this last position, since, for example, Tense-hopping apparently must be taken to be obligatory in order to avoid spurious derivations of such nonsentences as (2a′). Likewise, the treatment of Reflexivization in §6b requires that it be obligatory. If it were optional, there would be a spurious derivation for (7: *him*), since it could fail to apply on the lower S, and RO could move the potential antecedent to a position where it no longer met the conditions for Reflexivization.

(7) We believe $\text{John}_i$ to hate $\text{himself}_i/*\text{him}_i$.

Note that there is no general constraint against surface structures in which a personal pronoun and its antecedent are in the configuration of (7):

(8) We promised $\text{John}_i$ to help $\text{him}_i$.

Reflexivization also illustrates how whether a transformation is optional or obligatory can depend on details of what are taken as its conditions for application. In the analyses developed in this book, "indices" (such as the subscript in (7)–(8)) indicate the purported reference of NPs, and when the applicability of a transformation is contingent on two constituents being identical, a difference in the indices of otherwise identical NPs renders it inapplicable. In some versions of transformational grammar, however, syntactic structures are taken to be unspecified for purported reference, and transformations that here are regarded as obligatory and contingent on identity of indices are regarded as optional. Thus, Lees and Klima (1963) took (9a) and (9b) to have identical deep structures and to differ in whether the option of applying a certain optional transformation was taken or not:



(9) a. John hates himself.
b. John hates him.

In the approach adopted here, (9a–b) have deep structures that differ with regard to whether the two NPs have identical or distinct indices, with the deep structure of (9a) meeting and that of (9b) failing to meet the conditions of application of an obligatory transformation of Reflexivization.

### *f. The Cycle and Language Use*

It is far from obvious how the cyclic principle relates to the mental events that are involved in a language user's production and understanding of speech. This section will be devoted to a rough sketch of an approach to the understanding of language that is a fairly plausible hypothesis as to the sort of computation that goes on in people's heads and which turns out to embody a recognizable variant of the cyclic principle.

The approach in question, due to Lakoff and Thompson (1975a, 1975b), has the following characteristics: (i) Syntactic analysis is done "in real time," that is, words and phrases are assigned roles in a syntactic structure as soon as the hearer hears them—the hearer does not wait until he has heard the whole sentence before he starts doing a syntactic analysis. (ii) The syntactic analyses are **relational** in nature, that is, they consist of specifications of the grammatical relations (such as "is the subject of") that hold between the parts of a sentence. (iii) The analysis proceeds by steps of provisional assignment of grammatical relations and of revision of already assigned grammatical relations. (iv) The analysis yields a structure that is closely analogous to a deep structure rather than to a surface structure.

More specifically, the analysis works as follows. Consider the sentence (1):

(1) The hostages are believed by the police to have been killed.

Let us assume that besides being able to carry out the scheme of analysis about to be described, hearers can identify all the words in question and can identify NPs as being NPs. (The procedure to be sketched takes NPs as given and fits them into a syntactic structure.) The analysis begins by positing a sentence $S_0$. The first NP that is encountered (here, *the hostages*) is tentatively assigned the role of subject of $S_0$, and the first verb that is encountered after it (here, *are*) is tentatively assigned the role of predicate of

$S_0$. The next word encountered is *believed,* which is the past participle of a verb. The combination of a form of *be* and a following past participle triggers a reassignment of grammatical relations: the verb of which the past participle is formed takes over from *be* the role of predicate of $S_0$, and the NP currently assigned the role of subject of *be* is reassigned the role of object of the "main" verb (here, *believe*). The procedure so far can be summarized in the following table, in which entries in lower lines indicate revisions made on the entries in the upper lines:

(2)

| the hostages | are | believed |
|---|---|---|
| Subj of $S_0$ | P of $S_0$ | ——— |
| | ——— | believe |
| Obj of $S_0$ | | P of $S_0$ |

Next comes *the police*. If the S currently under analysis (here, $S_0$) has been identified as passive (i.e., if a NP previously identified as its subject has been reassigned the role of object), the NP of a *by* + *NP* combination may be assigned the role of subject of the S, and I assume that the option of so analysing it is taken here. (There is also the possibility of assigning it other roles, such as a locative or a time specification, as in *He was shot by the lakefront* or . . . *by 10:00*. For expository ease, I ignore such alternative possibilities here.)

We now encounter *to have*. The predicate of the S under analysis is at this point *believe,* and *believe* belongs to a class of verbs that, when followed by *to* and a verb in the infinitive form, trigger the following reassignment of grammatical relations: (i) a new S (here called $S_1$) is posited and takes over the status of "the S under analysis"; (ii) the object of *believe* is reassigned the role of subject of $S_1$; and (iii) $S_1$ is assigned the role of object of $S_0$. The analysis so far is as follows:

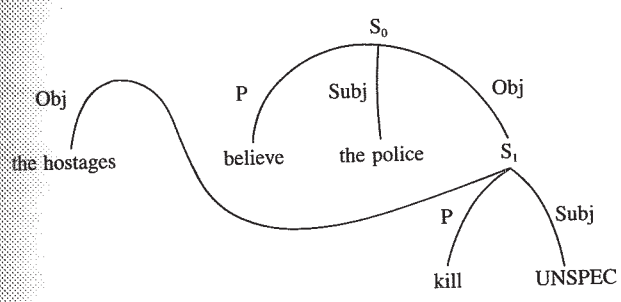
(3)

| $S_0$ | the hostages | are | believed | by | the police | to | $S_1$ |
|---|---|---|---|---|---|---|---|
| | Subj of $S_0$ | P of $S_0$ | ——— | | | | |
| | Obj of $S_0$ | ——— | believe | | | | |
| | | | P of $S_0$ | | | | |
| | | | | | Subj of $S_0$ | | |
| | Subj of $S_1$ | | | | | | Obj of $S_0$ |

Leaving aside *have,* which can be worked into the analysis in a fairly straightforward way, we reach *been killed* and do the same reanalysis that we did in the case of *are believed: be* is first assigned the role of P of $S_1$ but is then supplanted in this role by *kill,* and the role of *the hostages* is reassigned from subject of $S_1$ to object of $S_1$. If we retain only our final assignments of grammatical relations, we then have the following structure, which is a close approximation to the deep structure that would be set up under the analyses developed in chapters 3–5:



(4)




In this diagram, "Subj," "Obj," and "P" are labels not of nodes but of branches, and the curved lines are an informal way of indicating that left-to-right order should not be regarded as significant. The content of (4) is that $S_0$ is made up of *believe* as its predicate, *the police* as its subject, and $S_1$ as its object, with $S_1$ made up of *kill* as its predicate, *the hostages* as its object, and UNSPEC as its subject (the step that would assign UNSPEC that role was not given in the procedure sketched above, but could easily be filled in).

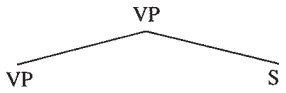
The various steps of reassignment of grammatical relations amount to the steps in a derivation with (4) as deep structure. Moreover, the interaction of these steps is precisely in accord with the cyclic principle: the initial assignments of grammatical relations correspond roughly to the surface structure, and in each reassignment of grammatical relations, either a more deeply underlying structure for the same S is recovered or a derived structure for an embedded S is posited and the task is changed to that of recovering an underlying structure for the latter S. Note in particular that in the application of Lakoff and Thompson's procedure to (1), *the hostages* is successively assigned the status of subject of $S_0$, of object of $S_0$, of subject of $S_1$, and of object of $S_1$, via steps that correspond to Passive on $S_0$, to RO on $S_0$, and to Passive on $S_1$. Moreover, the structures that are directly relevant to the various reassignments of grammatical relations are in most cases neither surface structures nor deep structures but intermediate structures in derivations, e.g., in the step that assigns to *the hostages* the role of subject of $S_1$, it is completely immaterial that *the hostages* is the surface subject of $S_0$: what is relevant is only that at that stage of the procedure it is the object of $S_0$.

## EXERCISES

1. Give derivations for the following sentences, ignoring tenses but giving full attention to complementizers:
   a. The boys all seem to be hated by their teachers.
   b. Your brothers happen to both be admired by Lucy.
2. a. Construct a derivation in which Raising-to-subject on a particular S feeds Q-float on that S.
   b. Construct one in which Tough-movement on a particular S feeds Equi on the next higher S.
   c. Construct one in which Passive bleeds Inversion.
   d. Determine whether Quantifier-float can feed VP-deletion.
3. Because Extraposition is optional, there is sometimes more than one possibility as to which domain it applies to. Show that the sentence

   John believes it to be obvious that God is dead.

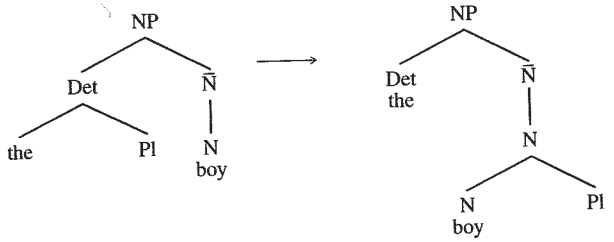
allows two different derivations (with slightly different surface constituent structures), one where Extraposition applies on the lowest possible domain and one where it applies on a higher domain. Assume that Extraposition of object complements, like that of subject complements, yields a derived structure




Give the two surface structures in full detail.

4. a. Provide an argument that the transformation deleting prepositions before *that*-clauses and *for-to* complements must be postcyclic.
   b. Do the same for *Do*-support.
   c. Do the same for the transformation that deletes *for* before *to*.
   d. Explain the cryptic remark in chap. 5, n. 6 about attachment of *-'s* to subjects of *'s -ing* complements.
5. Suppose that (contrary to what you are likely ever to have been told) in deep structure the plural morpheme is part of the Det and that it is moved into its surface position by a transformation something like Affix-hopping. Construct an argument, directly analogous to the one (on pp. 165–66 of this chapter) that Tense-hopping is postcyclic, showing that this transformation must also be postcyclic. In the role played by VP-deletion in the earlier argument, use the transformation that replaces a repeated $\bar{N}$ by *one*.






6. Decide whether each of the following is a root transformation, and give appropriate examples to justify your conclusion.
   a. Preposing of VP in such sentences as
      They said that John would win the prize, and win it he has.
   b. VP-deletion.
   c. Optional inversion accompanying preposed direct quotations, as in
      "You're insane," shouted John in a menacing tone.
   d. Extraposition of relative clauses.
   e. Inversion after preposed negatives, as in:
      Never before had he seen such stupidity.
      Under no circumstances will I talk to them.
   f. Deletion of *for* after *want*.
   g. Any other transformation whose status as a root transformation has not been taken up.

## NOTES

1. In most of the structures presented in this chapter, tenses, auxiliary verbs, and complementizers will be omitted for simplicity's sake. The omissions do not affect the points being made.
2. *Someone* appears in place of *anyone* in (2b), in accordance with the analysis to be adopted in §17a, according to which the *any* of sentences like (2a) is the form that *some* takes when it is in the scope of a negation. The position of *not* in (2b) is a makeshift and should not be taken seriously. A serious proposal regarding the deep structure location of *not* is made in §8b and elaborated in chapter 17.
3. This terminology is further systematized in Koutsoudas, Sanders, and Noll (1974).
4. Which version of (2a′) would result would depend on how exactly Passive would affect a structure that had a VP-modifier, which is a possibility that we have not considered in our discussion of Passive so far. I reject as a cop-out the suggestion that (2a′) be excluded by gerrymandering the details of Passive so that it would be simply inapplicable to such a structure.

5. For a more accurate statement of the conditions of application of Reflexivization, see §11f.

6. These observations are of course perfectly consistent with other schemes of rule application, e.g., simultaneous application or the combined simultaneous/random scheme proposed by Koutsoudas, Sanders, and Noll.

7. This argument is weak not only because it argues only against a straw-man alternative but also because even that straw man can be rescued by combining it with an idea that has considerable independent interest and much popularity, namely, the idea that in the output of movement transformations, the position previously occupied by the moved item is filled by a **trace**, i.e., a phonologically null constituent that functions syntactically as a pronoun whose antecedent is the moved item (Wasow 1979, Chomsky 1981). Under that conception of movement transformations, the above argument collapses, since the antecedent of the reflexive could then be the trace rather than *Bill,* and RO would not bleed Reflexive no matter what sequence the rules applied in.

8. To simplify exposition, an arbitrary choice has been made regarding which of (7a) and (7a′) is closer to deep structure. The few arguments that I know of that allegedly choose between an analysis that derives a structure like (7a) from one like (7a′) and an analysis that derives a structure like (7a′) from one like (7a) have shaky foundations.

9. The dotted line in (9b) indicates that no serious claim about the place of *too* in the constituent structure is intended. I in fact argue in §17e that in examples like (9b), *too* is Chomsky-adjoined to the VP.

10. The argument that Tense-hopping is postcyclic is due to Akmajian and Wasow (1975). The transformation that combines *-ing* and the past participle morpheme with their host verbs is, like Attraction-to-tense, cyclic. A possible case of precyclic transformations will come up in chapter 18, where some reasons will be given for taking "Negative incorporation" (which, e.g., combines *not* and *anyone* into *no one*) to be precyclic.

11. Emonds's definition of "local" required in addition that one of the two constituents involved be of a "lexical" rather than a "phrasal" category.

12. The basis of this claim will be presented in chap. 8.

13. The inclusion of the box labeled "surface structure combinatoric rules" anticipates conclusions arrived at in chap. 10.

I have inadvertently neglected to include in this section discussion of a principle that will play an important role in chapters 17–18 of this book, namely, the principle of **strict cyclicity.** This principle excludes de facto violations of the cyclic principle by imposing a narrow interpretation on what it is for a transformation to apply to a particular domain: if one potential domain *Y* for the application of transformations is contained in another domain *X,* then if all the material relevant to some application of a transformation is contained in *Y,* only *Y* and not *X* can count as the domain for that application of a transformation. This stipulation has the effect of preventing one from "returning to" *Y* after cyclic transformations have applied to *X*. See McCawley 1984 for discussion of alternative conceptions of "strict cyclicity" that have played a role in the literature and of the implications of the version of the principle that is adopted here.



14. Two caveats must be made regarding the interpretation of "main clause." First, when Ss are conjoined, each of the conjunct Ss has the main or subordinate status of the whole coordinate structure, e.g., in *Where did he go and when will he be back?* both *Where did he go?* and *When will he come back?* are main clauses. Second, direct quotations of main clauses count as main clauses, e.g., in *I'm sure that Bill asked Ann, "When will he come back?"* we must treat *"When will he come back?"* as a main clause, notwithstanding its being embedded in a larger structure. See Hooper and Thompson (1973), Green (1976), and Goldsmith (1981) for discussion of whether the transformations that Emonds identifies as applying only in main clauses really are restricted to main clauses.

15. *Do* in (2b) is unstressed; *John dóes enjoy tennis* is of course acceptable.

16. The construction may well be only partial, i.e., the speaker or hearer may well construct only as much of the derivation as he needs to for his purposes of the moment. The derivations given here purport to represent the linguistic structure that is accessible to the language user but do not embody the claim that he avails himself of all of that structure in each instance in which he uses the given sentence.