UNORDERED RULE HYPOTHESES

ANDREAS KOUTSOUNDAS
Indiana University

1. The assumption that rules must be extrinsically ordered, that is, that language-specific restrictions must be imposed on the application of rules, has been accepted as an important and well established insight into the nature of language from the earliest studies in transformational theory. For example, Chomsky and Halle (1968, 342) state:

(1) The hypothesis that rules are ordered ... seems to us to be one of the best-supported assumptions of linguistic theory.

Arguments given in support of the assumption that rules must be extrinsically ordered have the following form: extrinsically ordered rules are contrasted with unordered rules, the latter are shown to be inadequate, and it is therefore concluded that rules must be extrinsically ordered.

Any hypothesis which excludes language-specific constraints on the way rules are applied in a grammar is considered an Unordered Rule Hypothesis. While they have not always been explicitly identified, four such hypotheses have been argued against in the literature. Specifically, the arguments against unordered rules have been against one or more of the following hypotheses:

(2) a. The Direct Mapping Hypothesis
b. The Arbitrary Ordering Hypothesis
c. The Random Sequential Hypothesis
d. The Free Reapplication Hypothesis

The purpose of this paper is to show why, although the arguments given against the four Unordered Rule Hypotheses are valid, the conclusion that rules must therefore be extrinsically ordered does not follow.
2. In this section, the four Unordered Rule Hypotheses listed in section 1 will be briefly discussed, and representative arguments showing the inadequacy of each hypothesis will be given.

2.1 The Direct Mapping Hypothesis

The Direct Mapping Hypothesis, also known as the Simultaneous Application Hypothesis (Chomsky and Halle, 1968, 19, footnote 5), asserts that all rules apply simultaneously to an underlying form to derive a surface form. It thus asserts that there can be no intermediate representations between an underlying form and its corresponding surface form in any derivation. Clearly, if all rules apply simultaneously, it makes no sense to speak of the order in which they apply. The Direct Mapping Hypothesis cannot, therefore, allow language-specific ordering restrictions on the application of rules.

The following quotations illustrate that in some arguments unordered rules have been equated with rules which apply in accordance with the Direct Mapping Hypothesis.

(3) A great decrease in the complexity of rules and a great increase in the insight which they provide into the workings of language is obtained if phonological rules are ordered rather than simultaneously. (McCawley, 1968, 22)

The following statement is by Chomsky and Halle (1968, 19, footnote 5):

(4) ... L₃ supports a different empirical hypothesis concerning rule ordering, namely, that rules be unordered and that they apply simultaneously, so that each derivation has only two steps ... the empirical evidence in natural language rules against the hypothetical situation of L₃, and therefore against the simultaneous application hypothesis and in favor of hypotheses (13), (14) [given below].

(13) It is always possible to order the rules in a sequence and to adhere strictly to this ordering in constructing derivations without any loss of generality as compared to an unordered set of rules or a set ordered on a different principle.

(14) Such linear ordering makes it possible to formulate grammatical processes that would otherwise not be explicable with comparable generality.

Finally, consider the following statements from Postal (1968):

(5) Nothing reveals more completely or clearly the failure of stratification-al grammar to ordinary assum are UNORDERED STRATIFICATIONAL PHONOLOGICAL RULES [emphasis mine]

Arguments showing the following form. It rules can be specified sequentially, but that must be restricted simultaneity. Moreover, part of the structural other rule, which the Direct Mapping McCawley's argument underlyng and surfa Russian that permits pressing the rules of I mutated in (7), while would have to be exp mutations of (8).

(6) Underlying
   Zeg + 1 ≠
   Zeg + 1 + a ≠
   Zeg + 1 ≠ bi ≠
   Zeg + 1 ≠ 1, i ≠

(7) a. L-Drop: velar or lab
    b. Final Devo
    c. Voice Assi

(8) a'. L-Drop: it
    b'. Final Devo or if it is g in additioent.
    c'. Voice Assi
al grammar to provide a serious theory of phonology than the extraordinary assumption, fundamental to this view, that linguistic rules are unordered ... (p. 141)

Stratificational phonology is quite clear on the assertion that all phonological rules are simultaneous ... (p. 142) [emphasis mine: A K]

Arguments showing the inadequacy of the Direct Mapping Hypothesis have the following form. It is demonstrated that the structural descriptions of two rules can be specified in a very general way if the rules are allowed to apply sequentially, but that the structural description of at least one of the rules must be restricted in an ad hoc way if all rules are required to apply simultaneously. Moreover, this ad hoc restriction typically involves the inclusion of part of the structural description of one rule in the structural description of the other rule, with the resulting intersection inherently unexplainable as long as the Direct Mapping Hypothesis is maintained. For example, consider McCawley's argument from Russian (1968, 22-23). He points out that, given underlying and surface forms like those illustrated in (6), a grammar of Russian that permits sequential application of rules would be capable of expressing the rules of L-Drop, Final Devoicing, and Voice Assimilation as formulated in (7), while if the Direct Mapping Hypothesis is assumed, these rules would have to be expressed in the obviously less general and repetitive formulations of (8).

(6) Underlying | Surface
---|---
żeg + 1 $\neq$ | żok
żeg + 1 $\neq$ a $\neq$ | żla
żeg + 1 $\neq$ bi $\neq$ | żog bi
żeg + 1 $\neq$ 1,i $\neq$ | żok i,i

(7) a. L-Drop: word-final /l/ in a verb is deleted after a grave (i.e., a velar or labial) consonant.
b. Final Devoicing: word-final obstruents become voiceless.
c. Voice Assimilation: an obstruent assumes the voicing of an immediately following voiceless obstruent.

(8) a'. L-Drop: identical to (7a)
b'. Final Devoicing: an obstruent becomes voiceless if it is word-final or if it is grave and followed by word-final /l/ in a verb, provided, in addition, that it is not immediately followed by a voiced obstruent.
c'. Voice Assimilation: an obstruent assumes the voicing of an obstru-
ent which either follows it immediately or, if the first obstuent is grave, is separated from it by word-final /l/ in a verb.

2.2 The Arbitrary Ordering Hypothesis

The Arbitrary Ordering Hypothesis requires that:

(9) a. There are no language-specific ordering restrictions on the rules of a grammar.
b. All rules apply sequentially.
c. For each derivation, the rules are arranged in some arbitrary sequence and tested for application in that sequence.
d. Whether a rule applies or not, once it is tested for application, this rule is no longer available for further testing (within the same cycle).
e. A derivation terminates when the end of the sequence is reached.

The interpretation of unordered rules as rules that apply in accordance with the Arbitrary Ordering Hypothesis can be surmised from the following statement in Chomsky and Halle (1968, 342).

(10) Convention (29) is evidently not the only possible condition on the ordering of rules. It is possible, for instance, to require that rules apply in an ARBITRARY ORDER or that they apply simultaneously.

[emphasis mine: A.K.]

where convention (29) is stated as follows:

(11) Rules are applied in linear order, each rule operating on the string as modified by all earlier applicable rules. (p. 341)

Just as it is possible to give a general characterization of arguments against the Direct Mapping Hypothesis, it is also possible to give a general characterization of arguments against the Arbitrary Ordering Hypothesis. It is argued that the Arbitrary Ordering Hypothesis allows an input to be converted into any of several different outputs at least one of which is incorrect. Rules which must apply in a feeding order are generally used in such arguments; specifically, it is demonstrated that since this hypothesis allows rules to be tested for application in both a feeding and a counterefeeding order, both correct and incorrect outputs can be generated. For example, consider once more McCawley’s rules of L-Drop and Final Devoicing in Russian, stated in (7a) and (7b), respectively. Given an underlying representation like /\#eg1/\#/, these rules can, according to the Arbitrary Ordering Hypothesis, be tested for application in the sequence L-Drop first, then Final Devoicing (i.e., in a feeding order), or vice in the sequence L-Drop first, then Final Devoicing.

2.3 The Random 

The Random 

(12) a. There is a gran 
b. All ri 
c. Rules: 
d. When the same 
e. A der 

The following arguments, unor 

(13) ... all of sequen 

rather p 
entirely stags, v
2.3 The Random Sequential Hypothesis

The Random Sequential Hypothesis asserts that:

(12) a. There are no language-specific ordering restrictions on the rules of a grammar.

b. All rules apply sequentially; that is, no two rules can ever apply simultaneously.

c. Rules are tested for application randomly.

d. Whether or not a rule applies, once it has been tested for application, this rule is still available for further testing (within the same cycle).

e. A derivation terminates when there are no more rules which are applicable.

The following statement from Postal (1968, 143) illustrates that in some arguments, unordered rules are equated with the Random Sequential Hypothesis.

(13) ...all of the arguments given [below] in effect disconfirm the random sequential assumption as well. In fact, random sequential rules are rather pointless since the advantages of sequential application reside entirely in the possibility of constraining rule applications to certain stages, which is impossible with random application.
Like the arguments against the Arbitrary Order Hypothesis, those against the Random Sequential Hypothesis show that this hypothesis allows an input to be converted into several different outputs at least one of which is incorrect. Any example of a pair of rules that must apply in a counterbleeding order can be used to construct an argument of this type. Specifically, it can be shown that since this hypothesis allows the rules to be tested for application in both a counterbleeding and a bleeding order, both correct and incorrect outputs can be generated. Consider, for example, the two rules in (14), posited by Vennemann (1970, 77) for collective neuter nouns in Modern Standard German:

\[ \begin{align*}
&\text{Umlaut} \\
&V \rightarrow [\text{-back}] / \quad C_0 + \begin{bmatrix} V \\
&\text{-back} \end{bmatrix} \\
&\text{Apostrophe} \\
&e \rightarrow \emptyset / \quad \begin{bmatrix} \text{-voice} \\
&\text{-sonorant} \end{bmatrix} \quad \#
\end{align*} \]

The Umlaut rule fronts a vowel which precedes a front vowel; and Apostrophe deletes a word-final vowel which is immediately preceded by a voiceless obstruent.

According to the Random Sequential Hypothesis, Umlaut and Apostrophe can be tested for application in any order. Given an input like \( \text{ge + wolk} + \text{e} \), 'cloud formation', testing the application of these rules in the order Umlaut, then Apostrophe (i.e., in a counterbleeding order) gives the intermediate form \( \text{ge + wolk} + \text{e} \) and then the correct surface form \( \text{gewolk} \). Testing the application of these rules in the opposite order (i.e., in a bleeding order), we find that Apostrophe applies to \( \text{ge + wolk} + \text{e} \), to give \( \text{ge + wolk} \), but that Umlaut cannot be applied to this output because its structural description is not met; thus, the incorrect surface form \( \text{*ge + wolk} \) is derived. The conclusion for such cases is clear: the Random Sequential Hypothesis is inadequate.

2.4 The Free Reapplication Hypothesis

The Free Reapplication Hypothesis asserts that:

\[ \begin{align*}
&\text{a. There are no language-specific ordering restrictions on the rules of} \\
&\quad \text{a grammar.} \\
&\text{b. An obligatory rule must apply whenever its structural description} \\
&\quad \text{is met. (It follows from this requirement that rules will apply} \\
&\quad \text{simultaneously if possible; otherwise, they will apply sequentially.)} \\
&\text{c. All rules are scanned for applicability to each new representation} \\
&\quad \text{in a derivation.}
\end{align*} \]
esis, those against 
esis allows an input 
of which is incor-
overbleeding or-
cifically, it can be 
ted for application 
et and incorrect 
ules in (14), pos-
s in Modern Stan-

\[ \begin{align*}
  d. & \quad A \text{ derivation is terminated when no obligatory rules are applicable}.^4 \\

  \text{The following statement from Kenstowicz, Kim and Kisselberth (1972) illus-}
  \text{trates that in some arguments unordered rules are equated with rules that apply in }
  \text{accordance with the Free Application Hypothesis:}

  (16) & \quad \text{Having seen that there are grounds for rejecting both the direct map-}
  \text{ping hypothesis and the free reapplication hypothesis, let us consider }
  \text{a third alternative, which we can refer to as the Ordered Rule Hypothesis. (p. 30-31, mimeo)}

  \text{Unlike the arguments against the Unordered Rule Hypotheses discussed so far in this paper, the arguments against the Free Reapplication Hypothesis show that, given this hypothesis, there are cases in which an input can be converted only into an incorrect output. Pairs of rules that are in a mutual bleeding relation can be used to construct such an argument. Consider, for example, the following two rules of Palestinian Arabic adopted from Cole (1973, 56-57 57).}^5

  (17) & \quad \text{a. Geminate Reduction}
  \[
  \begin{array}{ccc}
    C & C & C \\
    1 & 2 & 3
  \end{array}
  \rightarrow
  \begin{array}{c}
    \emptyset \\
    3
  \end{array}
  \quad \text{(where } i = 2 \text{ or } 2 = 3) \\

  \text{b. } \iota\text{-Epenthesis}
  \[
  \begin{array}{ccc}
    \emptyset & \iota & C \\
    \iota & C
  \end{array}
  \rightarrow
  \text{C } \{\#\}

  \text{Geminate Reduction deletes one of two identical consonants under certain }
  \text{conditions, and } \iota\text{-Epenthesis inserts a high front vowel between two word }
  \text{final consonants or between the first two consonants of a three consonant }
  \text{cluster.}

  \text{A representation like } \text{jarrb} + u \text{ 'try' satisfies the structural description of }
  \iota\text{-Epenthesis since it has a three consonant cluster. It also satisfies the structural }
  \text{description of Geminate Reduction since a geminate cluster forms part of }
  \text{the three consonant cluster. Since } \text{jarrb} + u \text{ satisfies the structural descriptions }
  \text{of both rules, according to the Free Reapplication Hypothesis, both these }
  \text{rules must be applied to this representation. The only way that both rules can }
  \text{apply to this representation is simultaneously; but applying the rules simulta-}
  \text{neously to } \text{jarrb} + u \text{ gives the incorrect surface form } \text*jaribu.}

  (18) & \quad \text{jarrb + u}
  \[
  \begin{array}{c}
    \text{j} \\
  \end{array}
  \rightarrow
  \begin{array}{c}
    \text{a} \\
  \end{array}
  \quad \text{(17b) (17a)}

  \downarrow

  \text{*a} \quad \text{bo + u}

  \text{A new representation for the word 'try' can be proposed:} \\
  \text{\textit{Jari}b + u}
In other words, there is no way to derive the correct surface form *jarbu here because, according to the Free Reapplication Hypothesis, there is no way to make Geminate Reduction and not /-Epenthesis apply to jarrb + u. This being the case, the Free Reapplication Hypothesis is disconfirmed.

3. In the preceding section we saw that there are valid arguments against the Unordered Rule Hypotheses of Direct Mapping, Arbitrary Ordering, Random Sequential, and Free Reapplication. Although these arguments are valid, the inference that rules must therefore be extrinsically ordered is invalid. This inference would be valid only if it were the case that these four hypotheses were the only logically possible hypotheses of rule application that disallow extrinsic order. But this is not the case.

Koutoulouas, Sanders, and Noll (1971) proposed an Unordered Rule Hypothesis of UNIVERSALLY DETERMINED RULE APPLICATION. This Universally Determined Rule Application Hypothesis, henceforth the KSN hypothesis, claims the following:

(19)  
a. All restrictions on the application of rules are determined by universal principles (hence there are no language-specific ordering restrictions on the rules of a grammar) 6  
b. An obligatory rule must apply whenever its structural description is met, unless its application is precluded by some universal principle. (It follows from this requirement that rules will apply simultaneously if possible; otherwise, they will apply sequentially).  
c. All rules are scanned for applicability to each new representation in a derivation.  
d. A derivation is terminated when no obligatory rules are applicable. 4  

The KSN hypothesis differs from the Direct Mapping Hypothesis in that it allows some rules to apply sequentially. It differs from the Arbitrary Ordering Hypothesis and the Random Sequential Hypothesis in that it allows some rules to apply simultaneously. Finally, the KSN hypothesis differs from the Free Reapplication Hypothesis in that the former but not the latter can prevent a rule from applying to a representation that meets its structural description through the use of a universal principle.

It can easily be demonstrated that none of the standard arguments against the Unordered Rule Hypotheses discussed earlier can be used against the KSN hypothesis. Consider first the argument against the Direct Mapping Hypothesis. Since the KSN hypothesis does not require that all rules apply simultaneously, it is never the case that the structural description of one rule must be included in the structural description of another rule just so that these rules can apply simultaneously.

The example used to illustrate the general type of argument against the Arbitrary Ordering Hypothesis in Russian. Recall it is possible, given that the incorrect form *jarrb found to be applicable the input /jarrb + u/ to be applicable L-Drop to /jarrb + u/ and once it is  

The example used Random Sequential I of Umlaut and Apocope with the Rule /ge + walk + #/, not also possible to deriv Sequential Hypothes counter bleeding order which outputs no hypothesis, the rules case from which only con ge + walk + e # meet Apocope, and since t esis, they must both simultaneously gives the does not meet the str impossibility to derive .  

Finally, recall that Hypothesis involved /-Epenthesis. Specification like jarrb + u * form *jarbu, it is not because, given the Fr rules can apply to jarr rules in this manner aga the KSN hyp precedence of not only mutual bleeding relat principle:
face form /arbu
thesis, there is no way o jarrb + u. This be-

ments against the Ordering, Random
ents are valid, the d is invalid. This in-
out hypotheses that disallow ordered Rule Hy-
ion. This Univer-
the KSN hypot-

determined by uni-
natural ordering re-
some universal prin-
el will apply simul-
new representation
rules are applicable.
thesis in that it is Arbitrary Ordering at it allows some
sis differs from the t the latter can pre-
s arguments against used against the KSN Mapp-
, rules apply simulta-
on of one rule must just so that these
ment against the

Arbitrary Ordering Hypothesis involved the rules of L-Drop and Final Devoicing in Russian. Recall that from an underlying representation like /2og + 1 #/, it is possible, given this hypothesis, to derive both the correct form *zok and the incorrect form *2og. This is because the Arbitrary Ordering Hypothesis allows these rules to be tested for application in both the (correct) feeding order and the (incorrect) counterfeeding order. If the KSN hypothesis is adopted, however, no such multiplicity of outputs is possible since, given this hypothesis, the rules can be applied only in a feeding order, whereby only the correct output *zok is derived. Specifically, if the rules are scanned for applicability to the input /2og + 1 #/, as required by (19c), only the L-Drop rule is found to be applicable, and by (19b), it must be applied. The result of applying L-Drop to /2og + 1 #/ is 2og #. When the rules are scanned for applicability to this new representation, only Final Devoicing is found to be applicable, and once it is applied, the correct output *zok is derived.7

The example used to illustrate the general type of argument against the Random Sequential Hypothesis involved the Modern Standard German rules of Umlaut and Apocope. As stated earlier, when these rules are applied in accordance with the Random Sequential Hypothesis to a form such as /ge + wolk + #/, not only is it possible to derive the correct gewölk, but it is also possible to derive the incorrect *gewolk. This is because the Random Sequential Hypothesis allows these rules to be applied in both the (correct) counterbleeding order and the (incorrect) bleeding order. Again, no such multiplicity of outputs results with the KSN hypothesis, because given this hypothesis, the rules can only apply simultaneously to forms like ge + wolk + e # from which only correct forms will be derived. Specifically, since ge + wolk + e # meets the structural descriptions of both Umlaut and Apocope, and since these rules are obligatory, according to the KSN hypothesis, they must both apply to this input (see (19b)). Applying these rules simultaneously gives the correct surface form gewölk. Furthermore, since gewölk does not meet the structural descriptions of either Umlaut or Apocope, it is impossible to derive the incorrect surface form *gewolk.

Finally, recall that the argument given earlier against the Free Reapplication Hypothesis involved the Palestinian Arabic rules of Geminate Reduction and $-\text{Epenthesis.}$ Specifically, it was argued that, given these rules and a representation like jarrb + u 'try', although it is possible to derive the incorrect surface form *jarbu, it is not possible to derive the correct surface form jarbu. This is because, given the Free Reapplication Hypothesis, there is no way that these rules can apply to jarrb + u other than simultaneously, and applying these rules in this manner results in *jarbu. Such an argument cannot be advanced against the KSN hypothesis, however, since the correct applicational precedence of not only these rules, but also of numerous other pairs of rules in a mutual bleeding relationship, can be predicted by the following universal principle:
(20) PROPER INCLUSION PRECEDENCE: For any representation R, which meets the structural descriptions of each of two rules A and B, A takes applicational precedence over B if the structural description of A properly includes the structural description of B.

The structural description of a rule B is PROPERLY INCLUDED in the structural description of a rule A if and only if the STRUCTURAL DESCRIPTION of B can be placed upon the STRUCTURAL DESCRIPTION of A with some part of the structural description of A left over.

For the example under discussion, the relevant expansions of the structural descriptions of Geminate Reduction and i-Epenthesis are the following:

(21) SD of Geminate Deletion: C C C (where F = all features in the feature system)
     SD of i-Epenthesis: aF aF C C

Since the structural description of Geminate Deletion properly includes that of i-Epenthesis, Proper Inclusion Precedence predicts that Geminate Deletion and not i-Epenthesis must be applied to a representation like jarrb + u, which satisfies the structural descriptions of both rules. Applying Geminate Deletion to jarrb + u gives the direct surface form jarbu. Furthermore, since i-Epenthesis cannot apply to jarbu, the incorrect surface form *jariwu will never be generated. Thus, Proper Inclusion Precedence determines the correct application of these rules for Palestinian Arabic.

4. In section 3, the KSN hypothesis was contrasted with the four Unordered Rule Hypotheses. It was shown that none of the standard arguments against these four hypotheses are valid arguments against the KSN hypothesis. Since the examples used in constructing those arguments involved only phonological rules, and since the KSN hypothesis requires that restrictions on the order of application of BOTH phonological and syntactic rules be predicted by universal principles, it will now be demonstrated that there also are syntactic examples in which the empirically correct application of the rules involved can be predicted by universal principles.

4.1 The CYCLE is a generally accepted universal principle of grammar. Its function is to constrain transformational rules to apply from the most deeply embedded sentence to the next most deeply embedded sentence, and so on, until every rule has had a chance to apply to every sentence. Once this principle is assumed, the restrictions on the order of application of a number of syntactic rules can be predicted without extrinsic ordering. Consider, for example, Ross's claim (1968, 81) that Passive and Relative Clause Formation must be extrinsically ordered:

(22) a. Passive
     b. Relative

The argument is sentence (23) cannot:

(23) The man w

It is obvious that must precede the argument, however, that know of no evidence is not necessary to; Passive will be tested.

formation of Relative C this, assume that th ordered, and that th (24)

Clearly, Relative Cl the S cycle, and be however, can apply therefore, Passive w this sequence of ap

Now consider Ro that relates pairs of

(25) He is writing
(26) He is writing
extrinsically ordered as in (22).

(22)  a. Passive (optional)
      b. Relative Clause Formation (obligatory)

The argument is that, unless these rules are extrinsically ordered as in (22), sentence (23) cannot be derived.

(23) The man who was arrested by Officer McNulty went mad.

It is obvious that to derive sentences like (23), the application of Passive must precede the application of Relative Clause Formation. It is equally obvious, however, that if these rules apply cyclically in such derivations (and I know of no evidence that is inconsistent with this assumption), extrinsic order is not necessary to guarantee this order of application; if these rules are cyclic, Passive will be tested for application on an earlier cycle than Relative Clause Formation; thus the application of Passive will necessarily precede the application of Relative Clause Formation in deriving sentences like (23). To see this, assume that these rules apply cyclically, that they are not extrinsically ordered, and that the structure underlying sentence (23) is (24).

(24)

```
                  S2
                 /\
               /   \      VP
              /     S1    went mad
             /       /\
            /       /   \      NP
           /       /     S2    the man
          /       /       /\
         /       /       /   \      NP
        /       /       /     S1    Officer McNulty VP arrested
       /       /       /       /\
      /       /       /       /   \      NP
     /       /       /       /     S1    the man
```

Clearly, Relative Clause Formation can apply to this phrase marker only on the $S_2$ cycle, and being obligatory, it must be applied on this cycle. Passive, however, can apply to (24) only on the $S_1$ cycle. In deriving sentence (23), therefore, Passive will necessarily apply before Relative Clause Formation; and this sequence of application is precisely the sequence Ross wanted guaranteed.

Now consider Ross’s (1969b) rule of Sluicing. Sluicing is a deletion rule that relates pairs of sentences like (25) and (26).

(25) He is writing something, but you can’t imagine what he is writing.

(26) He is writing something, but you can’t imagine what.
Ross claims that the application of Sluicing must be restricted to follow the application of Wh-Fronting (Question Formation in his terminology) in any derivation requiring the application of both these rules, e.g., in the derivation of sentences like (25). He further states that this applicational restriction must be guaranteed by extrinsically ordering these rules as in (27).

(27)  
   a. Wh-Fronting (obligatory)  
   b. Sluicing (optional)

This applicational restriction can also be guaranteed without imposing extrinsic order on the rules, however - simply by requiring them to be cyclic. Once this is done, it will always be the case that, in deriving sentences like (26), Sluicing will apply on a later cycle than Wh-Fronting, thereby ensuring that the application of the rules will always be Wh-Fronting followed by Sluicing. For example, assume that the structure underlying (26) is (28).

(28)  
\[ S_4 \quad \text{but} \quad S_3 \quad \text{you can't imagine} \quad S_1 \]

he is writing something  
Q: he is writing Wi+something

If we assume further that the rules are not extrinsically ordered and that they apply cyclically, Sluicing cannot apply to (28) before we reach the S_4 cycle. The structural description of Wh-Fronting will be met either on the S_1 cycle (if this rule is defined on Q) or on the S_3 cycle (if this rule is defined on a higher verb), and being obligatory, it must be applied on one of these cycles. This means that, in deriving sentence (26), Wh-Fronting must be applied either on the S_1 or the S_3 cycle, and Sluicing on the S_4 cycle, which is equivalent to saying that in the derivation of this sentence, the application of Sluicing must follow that of Wh-Fronting.

4.2 In addition to the principle of the Cycle, a linguistic metatheory must include definitions of the notions 'obligatory rule' and 'optional rule'. In accordance with the KSN hypothesis these notions may be understood as follows.

(29) An obligatory rule is a rule that must apply if its structural description is met (unless some universal principle prevents its application).

(30) An optional rule is a rule that does not necessarily apply if its structural description is met.

Ringen (1972a) ha of obligatory rules is

(31) If on any one and an option rule must spg

In this paper, I will PRECEDENCE princi.
Because of the Obl the application of sor
dering. Consider, for known as Topicalizat
good reasons for resti sentences that requi
sentences like (22).

(32) Himself. Har

Postal claims that i
question must be ord

(33) a. Reflexive
   b. Y-Movemen

It is true, as Postal
of Y-Movement cann
that extrinsic orderin
hypothesis and that t
the Cycle cannot be e
cation of these rules. Optional Preceden

(34)

NP

Harry_1
Ringen (1972a) has pointed out that a consequence of this characterization of obligatory rules is the following.

(31) If on any one cycle, the structural descriptions of all obligatory and an optional rule are met by a given representation, the obligatory rule must apply to this representation.

In this paper, I will refer to this consequence as the OBLIGATORY-OPTIONAL PRECEDENCE principle.

Because of the Obligatory-Optional Precedence principle, restrictions on the application of some syntactic rules can be predicted without extrinsic ordering. Consider, for example, the rules of Reflexive and Y-Movement (also known as Topicalization). Postal (1971, 146-48) points out that there are good reasons for restricting Reflexive to apply before Y-Movement to derive sentences that require the application of both these rules, e.g., in deriving sentences like (32).

(32) Himself, Harry loves.

Postal claims that to guarantee this applicational restriction, the rules in question must be ordered as in (33).

(33) a. Reflexive (obligatory)
    b. Y-Movement (optional)

It is true, as Postal says, that in deriving sentences like (32) the application of Y-Movement cannot precede that of Reflexive. It is equally true, however, that extrinsic ordering is not necessary to guarantee this. Assuming the KSN hypothesis and that the underlying structure of (32) is (34), the principle of the Cycle cannot be evoked to guarantee the desired restriction on the application of these rules, for there is only one cycle in (34); but, the Obligatory-Optional Precedence principle will guarantee this restriction.

(34) 

```
S
  /
 NF
  /
 Harry1
      /
 V
 /     /
 NP    NP
    /
    VP
    /
   loves
   /
 Harry1
```
The structural descriptions of both the Reflexive and Y-Movement rules are met by (34). But since Reflexive is an obligatory rule and Y-Movement is an optional rule, Obligatory-Optional Precedence requires that Reflexive must apply to this phrase marker. If, in addition, Y-Movement is applied, sentence (32) is derived. Thus, since Obligatory-Optional Precedence guarantees the application of Reflexive to (34), it is not necessary to extrinsically order these rules.

4.3 The Complex NP Constraint is a universal principle of grammar proposed by Ross (1968, 70) to prevent elements from being moved out of relatives. Ross (1969b, 276-277) provides evidence that this constraint should be defined as a derivational rather than as a local constraint; that is, it should be defined on any pair of lines of a derivation whether they are adjacent or not. Since Ross does not actually reformulate this constraint, I have proposed the following formulation (Koutsoudas, 1973 p. 74-5).

(35) THE COMPLEX NP CONSTRAINT
If, in a line in a derivation, there is a complex noun phrase NP₁ consisting of a lexical head NP₂ and a sentence S₁, there cannot be another line in this derivation in which an element that was contained in S₁ is not dominated either by NP₁ or by S₁.

Once the Complex NP Constraint is assumed, the restrictions on the order of application of a number of rules can be predicted without extrinsic ordering. For example, as Ross (1968, 69) has pointed out, unless Topicalization is restricted so that it does not apply to an output of Relative Clause Reduction, ungrammatical sentences like (36) will be derived.

(36) The woman who was pestering the dog was Marsha.

(37) The woman pestering the dog was Marsha.

(38) *The dog, the woman pestering was Marsha.

Clearly, however, (35) eliminates the need for imposing extrinsic ordering to guarantee this restriction. Specifically, if (35) is adopted, the fact that constituents in reduced relative clauses cannot be topologized is explained by the fact that once Topicalization is applied to move a constituent out of a reduced relative clause, this constituent is no longer dominated by either the complex noun phrase or the sentence that contained it, thus violating the Complex NP Constraint; consequently, the sentence in question will be marked as ungrammatical. For example, assume that the structure underlying sentences (36)-(38) is (39).
Movement rules are applied after the reflexive must be satisfied. Sentence (39) guarantees that the order of these concepts is

(39) 

\[ \text{NP}_1 \rightarrow \text{VP} \]

\[ \text{NP}_2 \rightarrow \text{\textit{the woman}} \]

\[ \text{S}_2 \rightarrow \text{\textit{who was pestering the dog}} \]

Relative Clause Reduction can apply to (39) to give (40),

(40) 

\[ \text{S} \rightarrow \text{NP}_1 \rightarrow \text{VP} \rightarrow \text{\textit{was Marsha}} \]

\[ \text{NP}_2 \rightarrow \text{\textit{the woman}} \rightarrow \text{\textit{pestering the dog}} \]

If Topicalization were to apply to (40), the result would be (41),

(41) 

\[ \text{S} \rightarrow \text{NP}_1 \rightarrow \text{NP}_2 \rightarrow \text{VP} \rightarrow \text{\textit{was Marsha}} \]

\[ \text{the dog} \rightarrow \text{\textit{the woman}} \rightarrow \text{\textit{pestering}} \]

But now the Complex NP Constraint is violated because the noun phrase \textit{the dog} is not dominated by either \textit{NP}_1 or by \textit{S}_1. This being the case, sentence (38) will be blocked.

Consider now the fact that just as constituents in reduced relative clauses cannot be topicalized, neither can constituents in extraposed relative clauses; that is, sentences like (44), for example, are ungrammatical.

(42) \textit{The boy who was hugging the dog came in.} \\
(43) \textit{The boy came in, who was hugging the dog.} \\
(44) \textit{*The dog, the boy came in who was hugging.}
To account for this fact, it seems that Topicalization must be prevented from applying to an output of Extraposition from NP. (35) again guarantees this restriction, without imposing extrinsic ordering on the rules in question. Thus, assuming (35), and assuming that there is no extrinsic ordering, once Topicalization has applied to an output of Extraposition from NP, (35) will be violated and the resulting sentence will be marked as ungrammatical. For example, assume (35), no extrinsic order, and that the structure underlying sentences (42)-(44) is (45).

(45)  
```
S          
  /\         
 NP1 NP2     VP
  /\           came in
 S1 who was hugging the dog
```

Extraposition from NP can apply to (45), yielding (46).

(46)  
```
S          
  /\         
 NP1 NP2     VP
  /\           came in
 S1 who was hugging the dog
```

If Topicalization were subsequently to apply to (46), the result would be (47).

(47)  
```
S          
  /\         
 NP NP1     VP
  /\           came in
 S1 who was hugging
```

But now the Complex NP constraint is violated because the noun phrase "the dog" is dominated neither by $S_1$ nor by NP$_1$. This being the case, sentence (44) will be marked as ungrammatical.

5. In section 4, we NP Constraint, and given in (29) and (15) the rules involved in the KSN hypothesis.

In addition to these examples, S. Anderson, 1969, shows that specific restrictions on the cession of rules, incubation, and the specific ordering of ungrammatical language.

Although my analysis of particular hypotheses is not based on extrinsic ordering, these two alternative hypotheses have been found to be consistent with the samples of facts that ordering constraints in natural language.

Although the predicted ordering of applicability has been tested with respect to a specific hypothesis, this does not mean that all other constraints that rule hypotheses are incorrect, as shown in this paper.

6. In conclusion, the Direct Mapping Hypothesis, Free Hypothesis, and the KSN hypothesis are consistent with the examples provided in this paper.
5. In section 4, we saw that the universal principles of the Cycle, the Complex NP Constraint, and the characterization of ‘obligatory’ and ‘optional’ rules given in (29) and (30) suffice to predict the empirically correct application of the rules involved in many syntactic examples in a manner consistent with the KSN hypothesis.

In addition to the KSN hypothesis, a number of other hypotheses about how rules apply in derivations have recently been proposed (see, for example, S. Anderson, 1969; Kiparsky, 1971a; and Kisseberth, 1973a). These hypotheses are similar to the KSN hypothesis in that they eliminate some language-specific restrictions on the application of rules; but they differ from the KSN hypothesis in that they allow some language-specific restrictions on the application of rules, including extrinsic ordering.

Although my main purpose here is not to argue for the superiority of any particular hypothesis, clearly, a hypothesis that asserts that no rules are extrinsically ordered is to be more highly valued than any hypothesis that asserts that some rules are extrinsically ordered and some are not. This is the case because the former type of hypothesis is falsifiable by a single clear case that necessitates extrinsic order, while the latter type is not falsifiable at all. Given these two alternatives, the choice is clear: the hypothesis that no rules are extrinsically ordered must be assumed to be true until disconfirmed. Since this hypothesis has been subjected to, and survived, tests regarding representative samples of facts that have been previously accounted for by means of extrinsic ordering constraints, it can be considered a well supported hypothesis of natural language.

Although the particular universal principles for the prediction of the correct order of application of rules proposed thus far have obviously not been tested with respect to every fact that has been previously accounted for by means of extrinsic ordering constraints, and although further investigation of these and other cases will undoubtedly lead to new principles, there is no reason to believe that these developments will not be consistent with the hypothesis of Universally Determined Rule Application.

6. In conclusion, we have seen that of the many logically possible hypotheses of rule application that exclude extrinsic order, only four hypotheses appear to have been discussed since the inception of transformational grammar: the Direct Mapping Hypothesis, the Arbitrary Ordering Hypothesis, the Random Sequential Hypothesis, and the Free Resapplication Hypothesis. Clearly, if these were the only possible alternatives to extrinsic order, then the assumption that rules must be extrinsically ordered would be correct, since these four hypotheses are, in fact, inadequate. But this is not the case for, as I have shown in this paper, there is at least one more Unordered Rule Hypothesis, the KSN hypothesis, which is not disconfirmed by any of the standard argu-
ments against the (other four) Unordered Rule Hypotheses. The assumption, therefore, that rules must be extrinsically ordered is unwarranted.

NOTES

1. This should not be confused with the claim that multiple applications of a single rule are simultaneous (see Chomsky and Halle, 1968, 344), which means that whenever the structural description of a rule is met at two or more places in a representation, the rule is required to apply to all of them simultaneously. Whether this is the case or not has no bearing on the present discussion.

2. Postal's assertion here that "the advantages of sequential application resides ENTIRELY in the possibility of constraining rule applications to certain stages" (emphasis mine: A X) is clearly false since, as Postal himself must be aware, the main advantage of sequential application is that we are not forced to include the structural description of one rule in that of another, as we must if all rules apply simultaneously.

3. I have formalized Yeneman's prose statement of these rules.

4. In actuality, a derivation terminates when the optional rules have had a chance to apply and no obligatory rules are applicable.

5. I have simplified Cole's Geminat Reduction rule for ease of exposition.

6. The set of principles constituting the KSN hypothesis that has been proposed to date are illustrated in the remainder of this section and in section 4. Although this set of principles predicts the correct application of rules in many cases, it is still incomplete.

7. Notice, however, that in order to terminate this derivation, we must prevent vacuous application. For further discussion of vacuous application see Ringen's paper in this volume.

8. Depending on whether or not total sequentiality is assumed for optional rules (the Obligatory-Optional Precedence principle is neutral with regard to this assumption), Y-Movement could apply either simultaneously with Reflexive or after Reflexive has been applied.

9. I have simplified the formulation of the Complex NP Constraint given in the reference cited for ease of exposition.

10. This restriction is necessary only if we assume, with Ross, that pronominalization is included in the theory of grammar.

DISCUSSION

CHARLES KISSEBERTH (University of Illinois): This question is sort of a general one. In Koutsoudas, Sanders, and Noll (1971) (KSN), as I understand the situation, in cases where you had in past treatments bleeding—not mutual bleeding, just straight one-sided bleeding—according to the principles that you espouse, counterbleeding is the only interaction permitted. In this paper you develop a strategy for dealing with examples which in the past were treated as being bleeding. I refer to this strategy as the over-application strategy. For instance, in Kiparsky's (1968b) example—there are two of them in the book that follow the same pattern—in the example that you discuss, you let the rules which have been treated as bleeding in the past go ahead and apply: and
that derives an otherwise nonoccurring structure. Then, another rule, which you argue is independently motivated, fixes up the structure. So in past
analyses, where the bled rule did not apply at all, you let it apply and come
up with an independently motivated rule which corrects it. Now this strategy
wouldn't apply to all cases of bleeding but just to those in which letting the rule
not be bled creates an inadmissible structure, which can be patched up or cor-
corrected by an independently motivated rule. Now there are other kinds of
bleeding, or things that have been called bleeding in the past, which don't in-
volves the creation of inadmissible structures; and I'll give you an example. The
question is, do you have a strategy for handling these?

A typical example would be that you'd get a neutralization phenomenon,
say in syllable-final position, maybe deglottalization before a consonant. Now,
say you have a sequence of a glottalized consonant and a consonant in the
underlying structure, and you have an insertion rule that puts a vowel be-
tween the glottalized consonant and the following consonant and preserves
the glottalization of the first consonant. Now, using the over-application strat-
yeg, you would have to let Deglottalization operate, because its structural
description is satisfied by the underlying representation. You'd have to deglottalize
and insert the vowel at the same time, simultaneously. But then you'd be left
with an unglottalized consonant followed by a vowel, and there's nothing inad-
missible about that structure. So it's different in character from the two examples
that you discuss in your paper. I know that other people have developed strat-
egies to deal with this sort of situation but do you have any cases that you have
in mind where your strategy could not be applied because the result of over-
application is an admissible structure? If you can make sense of that.

ANDREAS KOUTSOUDAS (Indiana University): Okay. ... (laughter) ... Well,
you know the way you talked about the first kind of bleeding, you sounded
as if there was something unpleasant about letting a rule do something bad
and then fixing it up by an independently motivated rule. I mean, you don't
want to make the assumption that every line of a derivation has cognitive sig-
ificance or anything like that?

CHARLES KISSEIBERTH: Well, if the rule is independently motivated, it's a
possible out.

ANDREAS KOUTSOUDAS: Okay. Well, as you're going to hear from the next
paper, the question of whether rules should be allowed to bleed or not is dis-
puted. What is important, though, is that the way we're going to handle
the type of example you mention must be by a universal principle. So while I
have no strategy to cope with this type of example, other proposals have been
made to do just that.
JOHN JENSEN (University of Colorado): It occurs to me that the KSN hypothesis commits you to a fairly concrete level of phonology; that is, it would not allow any analysis of a language in which some abstract underlying form had to be maintained through several rules before it was neutralized. Do you have any comments on that?

ANDREAS KOUTSODAS: I think what Paul Kiparsky will be proposing has tremendous implications for that, in that one of the things that his universal principle does is to undercut one of the two main functions of extrinsic ordering, namely, ordering two rules so that one can apply only to underlying forms. With Kiparsky's principle, you're taking away one of the strongest arguments for extrinsic ordering. Whether this would lead to a more concrete phonology or not I don't know; it certainly would not necessarily lead to a concrete phonology in the sense of Vennemann.

ROBERT HOWREN (University of Iowa): I have in mind what seems to be a case of counterfeeding order, which the Proper Inclusion Precedence principle doesn't take care of.

ANDREAS KOUTSODAS: NOTHING takes care of that. I'm predicting that counterfeeding order is excluded as a possible rule interaction. If there are any such cases, and if the rules are good, and the facts are correct, they are counterexamples. Until recently, it was difficult to find genuine cases of counterfeeding order, so now what I'm doing is forcing other people to give me such cases, and maybe something will come out of it.

PAUL KIPARSKY (Massachusetts Institute of Technology): Counterfeeding cases are not exotic ... they're all over the place in Sound Pattern of English (Chomsky and Halle, 1968). For example, k in Romance words palatalizes before nonlow front vowels, so that we get cirrhostis and centipede but helical. Yet before nonlow front vowels that come from Vowel Shift, we don't get palatalization; we get hurricane and not *huriseyn. Clearly, this is counterfeeding. Then, secondly, I don't think all cases of mutual bleeding are taken care of by Proper Inclusion. A well known example is the German dialects where you get lanke-lange or lage-lange, depending on the ordering of g-Deletion with respect to Final Devoicing. Here, the rules are not related by proper inclusion, and even if they were, at least one of the two orders would be wrongly excluded by the KSN theory.

ANDREAS KOUTSODAS: I believe that Iverson (1973) tries to show that in the German example you mention it is not a reordering that is involved but a rule simplification.
The counterfeeding example that Kiparsky cites in this discussion — the example involving Chomsky and Halle's rules of Velar Softening and Vowel Shift in English — is not very convincing. Given Kiparsky's own Neutralization principle, Velar Softening must be ruled out as a possible rule of English. For an extensive discussion of this case, see Kouzoudas and Sanders (1974). [Added by AK during the preparation of this volume for publication.]