The Application of Phonological Rules

Andreas Koutsoudas; Gerald Sanders; Craig Noll

Language, Vol. 50, No. 1 (Mar., 1974), 1-28.

Stable URL:
http://links.jstor.org/sici?sici=0097-8507%28197403%2950%3A1%3C1%3ATAOPR%3E2.0.CO%3B2-S

Language is currently published by Linguistic Society of America.

THE APPLICATION OF PHONOLOGICAL RULES

ANDREAS KOUTSOUDAS   GERALD SANDERS   CRAIG NOLL

Indiana University   *University of Minnesota*   *Indiana University*

Phonological evidence is presented in support of the hypothesis that all restrictions on the relative order of application of grammatical rules are determined by universal rather than language-specific principles. For a systematically representative set of synchronic and diachronic facts, previously accounted for by means of non-universal extrinsic-ordering constraints, it is shown that there are alternative explanations of equal or greater generality in which the relative order of application of rules is either entirely unrestricted, or else fully predictable from the forms of the rules by a universal principle of proper inclusion precedence.*

1. The purpose of this paper is to present evidence in support of the following hypothesis:

> (1) All restrictions on the relative order of application of grammatical rules are determined by universal rather than language-specific principles.

It follows from this hypothesis that no grammatical rules are extrinsically ordered,[1] and hence that there are no possible pairs of natural languages which differ in any way that can be accounted for only by assuming that the languages have the same rules but different constraints on their relative order of application in derivations.

We will argue for this hypothesis by showing that the power of extrinsic ordering is empirically unmotivated with respect to facts about the phonology and phonological changes of natural languages. Specifically, we will attempt to show that any phonological fact that can be accounted for by means of a theory with the power of extrinsic ordering can be accounted for with equal or greater generality by means of a theory without this power.[2] We will thereby provide support for the more general hypothesis that the empirically correct order of application of grammatical rules is determined entirely by universal principles.

[1] Rule A is extrinsically ordered before rule B with respect to language L if and only if the grammar of L includes a set of one or more non-universal constraints which exclude from the set of well-formed derivations of that grammar any derivation in which, within the same structural or cyclic domain, a pair of lines related by the application of rule B precedes a pair related by the application of rule A (cf. Chomsky 1965:67, 1967:103). The arguments presented here against extrinsic-ordering restrictions are of course independent of the particular notational devices that are used for their expression, and thus apply equally to theories that express extrinsic-ordering constraints by abbreviatory numerical indices on rules (cf. Chomsky 1965), by explicit constraints on derivations (cf. Lakoff 1971), or by diacritic features or application markers of any sort (cf. Kenstowicz & Kisseberth 1970).

[2] Since this paper was written (September 1971), a number of refinements have been made in the particular theory assumed here. See, in particular, Iverson 1973 and Ringen 1972. All these refinements are consistent with the basic hypothesis of this paper—that all restrictions on rule application are determined by universal principles.

In theories with the power of extrinsic ordering, rules are assumed to apply sequentially, thereby determining a number of distinct application relations for pairs of rules. If a rule A is applied before a rule B in a given derivation, rule A will necessarily stand in one and only one of the following relations to rule B:

    (a) A FEEDS B if and only if the application of A INCREASES the number of forms to which B can apply.

    (b) A BLEEDS B if and only if the application of A DECREASES the number of forms to which B can apply.

    (c) A DOES NOT AFFECT B if and only if A neither feeds nor bleeds B.

Similarly, if A is applied before B, B will stand in one and only one of the following relations to A:

    (a) B COUNTER-FEEDS A if and only if the application of B WOULD INCREASE the number of forms to which A could apply IF B were to apply before A.

    (b) B COUNTER-BLEEDS A if and only if the application of B WOULD DECREASE the number of forms to which A could apply IF B were to apply before A.

    (c) B DOES NOT AFFECT A if and only if B neither counter-feeds nor counter-bleeds A.

Given that the first rule of a sequentially-applied pair must either feed, bleed, or not affect the second, and that the second must either counter-feed, counter-bleed, or not affect the first, any given rules A and B, where A is applied before B, will necessarily be related to each other in just one of the following nine ways:[3]

    FEEDING: A feeds B; B does not affect A.

    BLEEDING: A bleeds B; B does not affect A.

    COUNTER-FEEDING: A does not affect B; B counter-feeds A.

    COUNTER-BLEEDING: A does not affect B; B counter-bleeds A.

    MUTUALLY NON-AFFECTING: A neither feeds nor bleeds B; B neither counter-feeds nor counter-bleeds A.

    FEEDING AND COUNTER-FEEDING: A feeds B; B counter-feeds A.

    BLEEDING AND COUNTER-BLEEDING: A bleeds B; B counter-bleeds A.

    FEEDING AND COUNTER-BLEEDING: A feeds B; B counter-bleeds A.

    BLEEDING AND COUNTER-FEEDING: A bleeds B; B counter-feeds A.

While theories with extrinsic ordering permit pairs of rules to be related in any one of these nine ways, theories without extrinsic ordering permit pairs of rules to stand only in those relations that are predictable from the forms of the given rules by universal principles. The particular theory of applicational precedence that will

---

[3] The terms 'feeding' and 'bleeding' were used by Kiparsky 1968, and adopted later by Wang 1969 and Newton 1971. Wang's REPLENISHING and VOIDING relations are special cases of what Newton and we have called 'counter-feeding' and 'counter-bleeding', respectively. Chafe 1968 uses the terms APPROPRIATE and INAPPROPRIATE ADDITIVE INTERFERENCE (similar to our 'feeding' and 'counter-feeding', respectively), APPROPRIATE and INAPPROPRIATE SUBTRACTIVE INTERFERENCE (similar to our 'counter-bleeding' and 'bleeding', respectively), SYMMETRIC ADDITIVE and SUBTRACTIVE INTERFERENCE (similar to our compound relations 'feeding and counter-feeding' and 'bleeding and counter-bleeding', respectively), and NON-INTERFERENCE (similar to our 'mutually non-affecting'). Feeding and counter-bleeding relations generally determine a relation between rules and phonetic representations which is called TRANSPARENCY by Kiparsky 1972; counter-feeding and bleeding generally determine what he calls OPACITY.

be proposed here will in fact exclude all rule-ordering restrictions other than those determining the compound relation of bleeding and counter-bleeding; and these restrictions will be predicted in all cases by a single universal principle of proper inclusion precedence. In §2 we will provide support for this hypothesis of universally determined rule application with respect to the explanation of synchronic phonological data about natural languages. In §3, we will provide support for this hypothesis with respect to the explanation of facts about linguistic change.

2.1. The assertion that two rules are extrinsically ordered in a feeding relation is empirically equivalent to the assertion that the two rules are entirely unrestricted in their relative order of application, i.e. that each rule simply applies to every representation that satisfies its structural description.[4] The equivalence between rules ordered in a feeding relation and ones with no restrictions whatever on their order of application follows from the fact that, in both cases, each rule is applicable to ALL POSSIBLE representations which satisfy its structural description. Neither rule, in other words, is prevented by any specified applicability condition from applying to any possible representation which meets its structural description. To impose any language-specific restriction on the relative order of application of any such pair of rules would be to add to the grammar that includes these rules a statement which is wholly redundant.

The equivalence between the hypothesis of extrinsic ordering and the hypothesis of unrestricted rule application with respect to rules standing in a feeding relation can be illustrated, e.g. by considering Kiparsky's proposed rules of Consonant deletion and Diphthongization for certain dialects of Finnish (1968:177):

(2) a. CONSONANT DELETION: $\gamma \rightarrow \emptyset$ / V ____ V
    b. DIPHTHONGIZATION: $ee \rightarrow ie$

In these dialects, Kiparsky claims, forms like *vie* are derived from underlying representations like *vee* by means of 2b, which diphthongizes long (= geminate?) mid vowels. He also claims that the grammars of these dialects include 2a, which deletes 'certain medial voiced continuants', and that underlying representations like *teγe* will have phonetic forms like *tie*, as a result of the application of both the deletion and the diphthongization rules. He accounts for these facts by proposing that Consonant deletion is extrinsically ordered before Diphthongization, an ordering restriction which permits the former rule to feed derived representations into the domain of the latter. These assumptions thus allow for the following derivations:

(3) *vee*   *teγe*
    —   *tee*   (2a)
  vie *tie*   (2b)

But it is obvious that exactly the same derivations will be generated by grammars in which Diphthongization and Consonant deletion are not extrinsically ordered.

---

[4] All rules under consideration here are obligatory. From the meaning of 'obligatory', it follows that any such rule MUST be applied to ANY representation to which it CAN be applied, and that it can be prevented from applying to a representation that satisfies its structural description only by the explicit postulation of some constraint which restricts its domain of applicability to some well-defined proper subset of the set of representations that satisfy its structural description. The most common constraints of this sort are extrinsic-ordering ones. See Ringen 1972 and Norman 1972 for detailed discussion of the application of optional rules in syntax and phonology.

That is, given that these rules are obligatory, then if each rule has no applicability restrictions whatever and is simply applied to every possible representation that satisfies its structural description, Diphthongization will correctly apply to *vee* and *tee*, Consonant deletion will correctly apply to *teγe*, and the derivations will correctly terminate with the forms *vie* and *tie*, which satisfy the structural description of neither rule. In other words, the assertions that Kiparsky's grammar makes about the order of application of these rules follow necessarily from the tautology that rules can only be applied to representations to which it is possible for them to apply, and from the fact that obligatory rules MUST be applied to any representations to which they CAN be applied.

The same is true for all other pairs of rules standing in a feeding relation. For example, consider the following two rules of English proposed by Chomsky (1964:89):

(4) a. SPIRANTIZATION: $t \rightarrow s$ / _____ $+ i$
    b. PALATALIZATION: $si \rightarrow \check{s}$ / _____ V
        prezident + i > *prezidens + i*
        prezident + i + al > *prezidenš + al* (**prezidens + i + al*)

Given any representation such as *prezident+i+al*, it is clear that only 4a is applicable; and only after 4a has been applied, yielding *prezidens+i+al*, can 4b possibly be applied in the derivation of this form, yielding *prezidenš+al*. Thus there is no empirical basis for imposing any constraints whatever on the relative order of application of these rules; the facts thus argue against rather than for the claim that Chomsky explicitly intended them to support: 'As soon as the attempt to construct explicit rules to determine the phonetic shape of a string of formatives passes the most superficial and introductory stage, it becomes obvious that a fairly strict ordering must be imposed on phonological processes, if they are to be describable in full generality' (1964:88).

Similar counter-evidence to Chomsky's claim about extrinsic ordering is provided by rules 5a–b, proposed in this order for Southern Paiute by Chomsky & Halle (1968:349).

(5) a. VOWEL DEVOICING: $V \rightarrow [-\text{voice}]$ / _____ $\left\{ \begin{matrix} \# \\ [-\text{son}]\ V \end{matrix} \right\}$

    b. GLIDE/NASAL DEVOICING: $\begin{bmatrix} +\text{son} \\ -\text{voc} \end{bmatrix} \rightarrow [-\text{voice}]$ / _____ $\begin{bmatrix} -\text{cons} \\ -\text{voice} \end{bmatrix}$

Thus, although Chomsky & Halle assume that an ordering of 5a before 5b must be imposed on these rules, this assumption has no empirical basis, since the elimination of their proposed ordering constraint has no effect on the form or derivational function of their rules. E.g., given a form such as *paawa*, if rules 5a–b are not extrinsically ordered, the only derivation POSSIBLE from this input and these rules is that of 6, resulting from the application of 5a before that of 5b:

(6) *paawa*
    paawạ   (5a)
    paaw̥ạ   (5b)

Similar examples could be cited from almost every serious phonological study, since a very large number of the actual pairs of rules which have been proposed

(even by those linguists who are explicitly arguing for the necessity of extrinsic ordering) are rules that are in feeding relations, and hence require no ordering restrictions whatever to assure their proper application in any particular derivation. All facts that are accounted for by such feeding pairs can thus be accounted for with equal generality by theories which prohibit any language-specific restriction on the relative order of application of phonological rules. This appears to be equally true, as we will now see, for extrinsically ordered rules standing in relations other than feeding.

2.2. Kiparsky (1968:199) has proposed the extrinsically ordered rules 7a–b to account for the occurrence of alternations such as *tāx* 'day', *tāyə* 'days', as claimed for certain Low German dialects:

(7) a. SPIRANTIZATION: $\begin{bmatrix} +\text{stop} \\ +\text{voice} \end{bmatrix} \rightarrow [-\text{stop}] / V \underline{\quad\quad}$

    b. DEVOICING: $[+\text{obstruent}] \rightarrow [-\text{voice}] / \underline{\quad\quad} \#$

Kiparsky asserts that underlying forms like *tāg* undergo 'first spirantization (*tāg* > *tāγ*) and then devoicing (*tāγ* > *tāx*)', thus positing a counter-bleeding relation between 7a, which applies to voiced stops, and 7b, which devoices stops and would thus bleed 7a if it were ordered before it.

It is clear, however, that *tāg* meets the structural descriptions of both Spirantization and Devoicing, and that both rules can be applied simultaneously here without contradiction to yield the correct derived form *tāx*:

(8)   t   ā   g   #
           [+stop, +voice, ...]
              (7a)   (7b)
               ↓   ↓
      t   ā   x   #
           [−stop, −voice, ...]

Thus if rules are allowed to apply simultaneously, no language-specific restriction on the relative order of application of these rules can be justified.[5]

It should be noted that the simultaneous application of rules in counter-bleeding relations and the sequential application of rules in feeding relations follow necessarily from exactly the same principle—namely, that every obligatory rule MUST be applied to every representation to which it CAN be applied. The difference between the feeding and the counter-bleeding situations is simply that in the former

---

[5] There are a number of well-known arguments presenting empirical evidence against the hypothesis that ALL rules must be applied simultaneously, and hence that all derivations must consist of exactly two lines (cf. Chomsky & Halle, 342–3, 348–9; McCawley 1968:22–3; Postal 1968:140–52). But these arguments obviously do not rule out the assumption that SOME rules are applied simultaneously, and hence that the number of lines in a derivation may be greater than two but less than the number of rules that are applied in determining that derivation. The most natural initial hypothesis about simultaneous and sequential application would seem to be simply that those rules that CAN apply simultaneously to a given representation DO apply simultaneously; those that can't, don't. Evidence against complete simultaneity is of no value at all, of course, for deciding between this hypothesis of predictable simultaneity and sequentiality and the more commonly accepted hypothesis that ALL rules are applied sequentially, regardless of whether they could or could not be applied simultaneously.

case there are no representations to which both rules CAN apply, while in the latter case there are. In both cases, every rule is applied to every structure to which it is applicable.

The fact that the extrinsically ordered counter-bleeding relation of 7a–b can be reduced to an empirically equivalent relation of simultaneous application is not accidental to these particular rules. This reduction is possible for all other known cases of rules which have been extrinsically ordered in a counter-bleeding relation. E.g., consider one of the traditional analyses of vowel nasalization and nasal consonant deletion in French:

(9) a. NASALIZATION: $V \rightarrow \tilde{V} /$ ____ N

   b. CONSONANT DELETION: $N \rightarrow \emptyset /$ ____ $\left\{ \begin{matrix} \# \\ C \end{matrix} \right\}$

To account for derivations such as *grande* > *grãd* (not **grad*), grammars which conform to theories that require that ALL rules be applied sequentially must impose a restriction that 9a be extrinsically ordered before 9b. This language-specific restriction on rule application is necessary for such grammars in order to prevent 9b from applying to representations like *grande* to yield ungrammatical surface representations like **grad*. In other words, such grammars require that 9a–b be asserted to stand in a counter-bleeding relation.

For theories without the power of extrinsic ordering, on the other hand, the requirement that every obligatory rule MUST be applied to every representation to which it CAN be applied is sufficient to determine the simultaneous application of 9a–b to all representations like *grande*, since such representations meet the structural descriptions of both rules, and since the structural changes specified by these rules are mutually consistent with respect to these inputs. The simultaneous application of these rules is illustrated here:

(10) g   r   *a*   *n*   *d*   *e*
             | |
          (9a) (9b)
             ↓   ↓
     g   r   *ã*   *∅*   *d*

Another example of the reduction of counter-bleeding relations to relations of simultaneous application is provided by the following pair of rules for Uruguayan Spanish posited by Saporta (1965:223):

(11) a. VOWEL LOWERING: $e \rightarrow \varepsilon /$ ____ C

   b. FINAL *s*-DELETION: $s \rightarrow \emptyset /$ ____ #

If 11a is extrinsically ordered before 11b, underlying *klases* 'classes' will first be converted to **klasɛs* by 11a, which will then be converted to the correct surface form *klasɛ* by 11b. If the order of application of these two rules were reversed, 11b would convert underlying *klases* to **klase*, to which 11a would not be applicable; i.e., this bleeding order of application would be inconsistent with the fact that the correct derived representation of *klases* in Uruguayan Spanish is *klasɛ* rather than **klase*. Thus the correct ordering relation between 11a and 11b in grammars in which all rules are applied sequentially is the counter-bleeding relation of 11a before 11b.

It is evident, however, that representations like *klases* satisfy the structural descriptions of both Vowel lowering and Final *s*-deletion, and that these rules could

thus be simultaneously applied to such representations. Such application will always yield the correct derived forms:

(12) k   l   a   s   e   s   #
                    | |
                 (11a) (11b)
   ↓   ↓
     k   l   a   s   ε   ∅   #
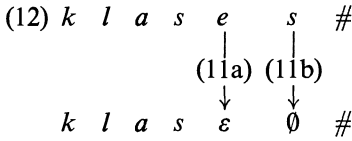
As a final example of reducing counter-bleeding relations to simultaneous application relations, consider the following rules for Latin American Spanish from Saporta (220, 222):

(13) a. *k*-INSERTION: $\emptyset \rightarrow k / V\theta$ ____ $\left\{ \begin{matrix} o \\ a \end{matrix} \right\}$

b. STRIDENCY: $\theta \rightarrow s$

Saporta asserts that 13a must be extrinsically ordered before 13b. This counter-bleeding relation is required to account for the fact that, e.g. underlying *kreθo* 'I grow' yields *kresko* rather than **kreso* as the correct derived representation. But again, representations like *kreθo* satisfy the structural descriptions of both 13a and 13b; and given the requirement that obligatory rules must be applied whenever possible, 13a–b will automatically be applied simultaneously to *kreθo* to give the correct derived form *kresko*:

(14) k   r   e   θ   o
                 | |
              (13b) (13a)
                 ↓   ↓
     k   r   e   s   k   o
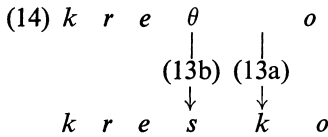
It thus appears that the explanatory power of any pair of rules extrinsically ordered in a counter-bleeding relation will always be precisely equivalent to that of the same pair of rules without any language-specific restriction on their order of application.[6]

---

[6] It should be noted that there is a possible alternative treatment of the counter-bleeding cases dealt with here which is consistent with the assumption of complete sequentiality, but also fully consistent with the hypothesis that there is no extrinsic ordering of phonological rules. Instead of simultaneous application, this treatment uses the following universal precedence principle:

COUNTER-BLEEDING 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 there is some string that is included in the inputs of both A and B, but not in the output of B.

The adequacy of this counter-bleeding principle is exemplified by Saporta's rules 13a–b, where the correct applicational precedence of 13a over 13b with respect to forms such as *kreθo* follows from the fact that the structural descriptions of both rules include the string θ, and the output of 13b does not include θ. Alternative treatments involving simultaneous application and the counter-bleeding principle are possible for all counter-bleeding cases known to us, and we thus have no empirical basis for choosing between them here.

The kind of evidence which would suffice to falsify counter-bleeding precedence, and support simultaneous application, is illustrated by the following hypothetical example:

Rules: (a) $\breve{V} \rightarrow \overline{V} /$ ____ $\begin{bmatrix} +\text{cons} \\ +\text{voice} \end{bmatrix}$

(b) $[+\text{voice}] \rightarrow [-\text{voice}] / \breve{V}$ ____ #

Underlying form: *kăd*

Derived form: *kāt*

2.3. Another relation which may hold between two extrinsically ordered rules is the compound one of bleeding and counter-bleeding. This relation is exemplified by the following pair of rules for Latin American Spanish discussed by Saporta (222):

(15) a. FINAL DEPALATALIZATION: $\tilde{l} \rightarrow l / \underline{\quad\quad} \#$
    b. DELATERALIZATION: $\tilde{l} \rightarrow y$

Saporta proposes that 15a, which holds for Castilian as well as Latin American Spanish, is extrinsically ordered before 15b, which holds only for Latin American Spanish. Given this ordering, correct derivations will be effected for any underlying forms containing palatal laterals, e.g. *akeĺ* 'that', *akeĺos* 'those':

(16)   CASTILIAN   LATIN AMERICAN
       akeĺ   *akeĺos*   *akeĺ*   *akeĺos*
       akel   ——————   *akel*   ——————   (15a)
       ——————   ——————   ——————   *akeyos*   (15b)

Given this order of application, 15a will bleed 15b of only those palatal laterals which are in word-final position. The latter rule will then apply to all remaining representations which include palatal laterals, converting them into non-lateral glides. If the order of application were reversed here, 15b would apply to ALL representations which include palatal laterals; and, by converting them all to non-lateral glides, would bleed 15a of all possible representations which might satisfy its structural description. Therefore, given the restriction that 15a is extrinsically ordered before 15b, 15a stands in a bleeding relation to 15b, and 15b stands in a counter-bleeding relation to 15a.

It is clear that, given any representation such as *akeĺ* to which both rules are applicable, simultaneous application would yield an incorrect output. The correct application of all such pairs of rules is fully determined, nevertheless, by the following universal principle of proper inclusion precedence proposed by Sanders 1970:

(17) 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 with respect to R if and only if the structural description of A properly includes the structural description of B.[7]

Given these rules and forms, it is clear that the counter-bleeding principle would generate a pair of mutually contradictory predictions about the application of (a) and (b): the precedence of (a) over (b) would follow from the fact that [+voice] is included in both structural descriptions, but not in the output of (b); the converse precedence of (b) over (a) would follow from the inclusion of V̆ in the structural descriptions of both rules, but not in the output of (a). The simultaneous application of (a) and (b) is possible, however, and yields the correct results:

$$
\begin{array}{ccc}
k & \breve{a} & d \\
| & | & | \\
 & (a) & (b) \\
 & \downarrow & \downarrow \\
k & \bar{a} & t
\end{array}
$$
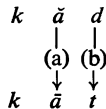
On the other hand, an example of the kind of evidence which would suffice to support the counter-bleeding precedence principle and falsify simultaneous application is provided by the hypothetical rules 25a–b of §2.6.

For present purposes, however, it makes no difference which of these two alternatives turns out to be correct, since both are consistent with the hypothesis that rules are not extrinsically ordered.

This principle determines the correct application of the two Latin American Spanish rules 15a–b, since the structural description of the former is the string $l\#$ and that of the latter is the string l, and $l\#$ properly includes l. There is no need, therefore, for any language-specific constraints on the application of these rules.

The correctness of this principle can also be demonstrated with respect to numerous other pairs of bleeding and counter-bleeding rules proposed for various languages. Consider the following pair of rules proposed for Caddo by Chafe (1968:124):[8]

$$\text{(18) a.} \quad \begin{bmatrix} +\text{cons} \\ +\text{back} \\ +\text{stop} \end{bmatrix} \begin{bmatrix} -\text{voc} \\ +\text{round} \end{bmatrix} \rightarrow \begin{bmatrix} +\text{cons} \\ -\text{back} \\ +\text{stop} \\ +\text{round} \end{bmatrix} \qquad k\,w \rightarrow p$$

$$\text{b.} \quad \begin{bmatrix} +\text{cons} \\ +\text{back} \\ +\text{stop} \end{bmatrix} [-\text{voc}] \rightarrow \begin{bmatrix} +\text{cons} \\ +\text{back} \\ -\text{stop} \end{bmatrix} [-\text{voc}] \qquad k \rightarrow h\,/ \rule{1cm}{0.4pt}\ C$$

[7] 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.

It should be noted that the structural description of any rule of the form $X \rightarrow Y / W \rule{1cm}{0.4pt} Z$, like that of its notational alternate $W\,X\,Z \rightarrow W\,Y\,Z$, is the symbol string $W\,X\,Z$ and not merely X; and that the proper application of 17, like that of the counter-bleeding principle stated in fn. 6, requires that all rules be expressed in terms of features rather than alphabetic symbols. It will also be observed that the proposed precedence principle subsumes as a special case the familiar ordering of a context-sensitive rule before its corresponding context-free 'elsewhere' rule. It should be noted finally that if a structural description X PROPERLY INCLUDES a structural description string Y, then the set of representations which meets structural description string X is PROPERLY INCLUDED in the set of representations that meet structural description Y.

For all the cases of proper inclusion precedence considered here, the related rules are intrinsically disjunctive, since application of either rule yields a representation that fails to satisfy the structural description of the other. However, there is some evidence that for syntactic rules, at least, there can be proper inclusion precedence between rules which do not destroy each other's contexts, but which must nevertheless be prevented from applying conjunctively in derivational reflexes of the same structure. It thus seems appropriate to consider disjunctive application as part of the meaning of proper inclusion precedence; this could be effected, e.g., by adding to 17 the condition that rule B cannot be applied to any (direct or indirect) product of the application of rule A. Further research is required, however, before the precise nature of the relationship between disjunctiveness and applicational precedence can be adequately specified.

[8] These rules could also have been formulated in terms of Chomsky & Halle's feature system as:

$$\text{(18) a'.} \quad \begin{bmatrix} +\text{cons} \\ -\text{ant} \\ -\text{cor} \\ -\text{cont} \\ -\text{voice} \end{bmatrix} \begin{bmatrix} -\text{cons} \\ -\text{voc} \\ +\text{back} \end{bmatrix} \rightarrow \begin{bmatrix} +\text{cons} \\ +\text{ant} \\ -\text{cor} \\ -\text{cont} \\ -\text{voice} \end{bmatrix}$$

$$\text{b'.} \quad \begin{bmatrix} +\text{cons} \\ -\text{ant} \\ -\text{cor} \\ -\text{cont} \\ -\text{voice} \end{bmatrix} [-\text{voc}] \rightarrow \begin{bmatrix} -\text{cons} \\ -\text{voc} \\ +\text{low} \end{bmatrix} [-\text{voc}]$$

Although such a reformulation would not show that 18a is a contraction rule in Caddo, it nevertheless would still allow principle 17 to predict the correct applicational precedence.

Since the input of 18a properly includes that of 18b, it follows from principle 17 that 18a and not 18b must be applied to any representation of the form X*kw*Y which satisfies the structural descriptions of both rules. This precedence determines the correct application of Chafe's rules for Caddo.

Examples such as these demonstrate that the application of rules in the compound relation bleeding and counter-bleeding can be correctly predicted by a universal principle of proper inclusion precedence. The correct prediction of rule applications has also been shown to follow from universal principles for all rules standing in feeding relations and in counter-bleeding relations. For all the rules that have been discussed, no loss of generality has been associated with the elimination of extrinsic-ordering restrictions, since these restrictions have been removed without changing any of the rules themselves. This means that for a vast number of phonological rules that have actually been proposed for various languages, there is no justification whatever for imposing any language-specific restriction on their relative order of application.

We wish to argue now that this is also true with respect to facts which have been accounted for previously by pairs of rules which do not stand in any of the three relations discussed thus far, and whose order of application cannot be determined by any known universal principle. We will do this by showing that, for each of these cases, there is a reasonable alternative explanation which uses natural rules whose relative order of application is fully determined by the proposed universal principles of rule application.

2.4. Consider the following two rules proposed for Modern Polish and Old Church Slavic by Kiparsky (1968:197–8):

$$(19) \text{ a. DEAFFRICATION: } \begin{bmatrix} +\text{voiced} \\ -\text{grave} \\ +\text{strident} \end{bmatrix} \rightarrow [+\text{continuant}] \qquad\qquad \begin{matrix} d\check{z} \rightarrow \check{z} \\ dz \rightarrow z \end{matrix}$$

b. SECOND PALATALIZATION:

$$\begin{bmatrix} +\text{obstruent} \\ -\text{grave} \\ -\text{strident} \\ -\text{diffuse} \end{bmatrix} \rightarrow \begin{bmatrix} +\text{strident} \\ +\text{diffuse} \end{bmatrix} / \underline{\quad} \check{e} \qquad \begin{matrix} k,\check{e} \rightarrow ts\check{e} \\ g,\check{e} \rightarrow dz\check{e} \end{matrix}$$

In these languages the affricates resulting from Second palatalization do not deaffricate; cf. derivations such as *g,ĕlo* 'very' > *dzĕlo* (not **zĕlo*). Kiparsky, having a theory with the power of extrinsic ordering, proposes to account for this fact by requiring that the grammars of Polish and Old Church Slavic include not only rules 19a–b, but also the restriction that 19b must never be applied before 19a, i.e. by asserting that 19b is in a counter-feeding relation with respect to 19a.

However, the fact that *dz* in forms like *dzĕlo* does not deaffricate can be accounted for without any extrinsic-ordering restriction whatever, simply by restricting Deaffrication so that it applies only to non-diffuse stridents, i.e. to *dž* but not to *dz*. All that is required, therefore, is that Kiparsky's 19a be replaced by the following:

$$(19) \text{ a}'. \begin{bmatrix} +\text{voiced} \\ -\text{grave} \\ +\text{strident} \\ -\text{diffuse} \end{bmatrix} \rightarrow [+\text{continuant}] \qquad\qquad\qquad d\check{z} \rightarrow \check{z}$$

Thus the facts which Kiparsky accounts for by two rules and one ordering restriction can also be accounted for by the two rules alone at the cost of only one additional feature specification. Furthermore, 19a′ expresses a more revealing generalization about deaffrication in these languages than 19a, since it follows correctly from 19a′ (but not from 19a) that phonetic *z* is never derived from *dz* in these languages (see Shevelov 1965:635, Ivanov 1964:134 ff.)

The hypothesis that there is no extrinsic ordering of phonological rules implies that similar explanations will be possible for facts accounted for by any other pair of rules extrinsically ordered in a counter-feeding relation. We know of no clear evidence that would contradict this implication.[9]

2.5. The hypothesis of universally-determined rule ordering appears to be equally consistent with the explanation of any fact that has been accounted for by pairs of rules extrinsically ordered in a bleeding relation. Consider the following two rules proposed by Kiparsky (1968:176, 178) for the Schaffhausen dialect of Swiss German:

(20) a. UMLAUT: $V \rightarrow [-\text{back}] / \dots$ Umlaut context

b. BACK VOWEL LOWERING: $\begin{bmatrix} V \\ -\text{high} \\ +\text{back} \end{bmatrix} \rightarrow [+\text{low}] / \underline{\quad} \begin{bmatrix} +\text{cons} \\ -\text{grave} \\ -\text{lateral} \end{bmatrix}$

$o \rightarrow \mathrm{\partial} / \underline{\quad} \{t, d, \dots\}$

[9] This includes Kiparsky's proposal that some Finnish dialects have rules 2a–b (of §2.1) with the restriction that 2b is extrinsically ordered before 2a. There are several ways of accounting for the facts about these dialects without the assumption of extrinsic ordering, but we are unable to determine which of these alternatives are most highly valued independently of the particular data cited by Kiparsky. (We are indebted to Robert Harms, Meri Lehtinen, and Thomas Perry for information about diphthongization and consonant alternations in Finnish dialects.)

E.g., all the facts that Kiparsky accounts for by rules 2a–b and the posited counter-feeding ordering of 2b before 2a can also be accounted for by assuming that these dialects have underlying long vowels rather than geminates, and that their grammars include, in place of the geminate diphthongization rule 2b, this equally plausible long-vowel diphthongization rule:

(2) b′. $\bar{e} \rightarrow ie$

These assumptions determine derivations such as the following:

$$\begin{array}{cc} v\bar{e} & te\gamma e \\ | & | \\ (2b') & (2a) \\ \downarrow & \downarrow \\ vie & tee \end{array}$$

The phonetic forms that Kiparsky gives for the dialects in question are thus readily derivable without the assumption of any extrinsic-ordering constraints. A more strongly supported alternative has been proposed by Perry 1971, who provides evidence that only one of the two dialect groups has a phonological rule of diphthongization (2b). Additional support for Perry's analysis is provided by Kiparsky's (1972) constraint on absolute neutralization, which effectively precludes Kiparsky's own extrinsic order-based analysis of the Finnish dialect data (1968). There are also a number of other extrinsic-ordering-free accounts of these purportedly counter-feeding dialects, differentiating them from the purportedly feeding dialects in terms of differences in syllable structure or syllabic constraints on diphthongization, or in terms of the completeness of their respective consonant reduction processes. The phonetic data available to us are not sufficiently detailed, however, to permit a choice from among these alternative treatments.

Kiparsky exemplifies the alternations to be accounted for by these rules by the following forms:
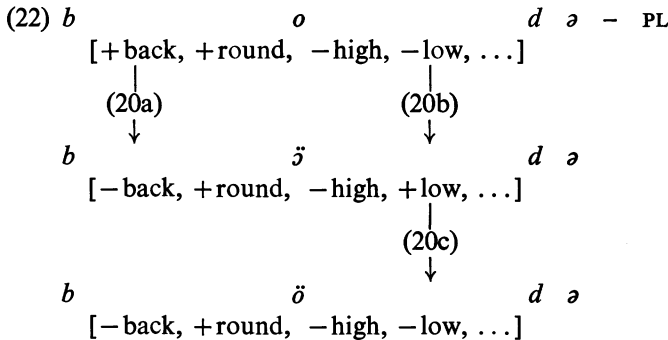
(21)   SINGULAR   PLURAL (UMLAUT CONTEXT)
      underlying: *bogə bodə*   *bogə*-PL *bodə*-PL
      derived: *bogə bɔdə*   *bögə*   *bödə*

To account for the fact that the umlauted alternant of *ɔ* is *ö* rather than **ɔ̈*, Kiparsky proposes that 20a is extrinsically ordered before 20b, so that Umlaut will bleed Back vowel lowering of all mid back vowels that are also in an umlauting context.

However, the fact that the plural of a form like *bodə* is *bödə*, rather than **bɔ̈də*, can clearly also be accounted for by assuming that the Schaffhausen dialect has a general rule to the effect that all front rounded vowels are non-low:[10]

(20) c. $\begin{bmatrix} V \\ -\text{back} \\ +\text{round} \end{bmatrix} \rightarrow [-\text{low}]$

Given this general rule, which is consistent with all the Schaffhausen data provided by Kiparsky, it is clear that extrinsic-ordering restrictions are entirely superfluous here, since the unrestricted application of 20a–c to any representations that satisfy their structural descriptions yields the correct derived forms for this dialect. Consider, e.g., the following derivation:

(22) b   o   d   ə   —   PL
      [+back, +round, −high, −low, ...]
            | |
          (20a)   (20b)
            ↓   ↓
     b   ɔ̈   d   ə
      [−back, +round, −high, +low, ...]
                              |
                            (20c)
                              ↓
     b   ö   d   ə
      [−back, +round, −high, −low, ...]

Since the structural description of 20c is not satisfied by the underlying form *bodə*-PL, this rule cannot apply. But the structural descriptions of 20a–b are satisfied by this representation; and since the fronting and the lowering of a back vowel are non-contradictory structural changes, both rules will apply simultaneously to *bodə*-PL to yield the intermediate form **bɔ̈də*. But this includes a front rounded vowel and thus satisfies the structural description of 20c, which will apply to change the low front rounded vowel *ɔ̈* to the non-low front rounded vowel *ö*, yielding the correct derived form *bödə*.

Since it is true that the Schaffhausen dialect has no low front rounded vowels at all (Stickelberger 1881:18–19), rule 20c accounts for a significant fact about this

[10] Like all other phonological rules, a segment-structure rule like 20c will apply to all representations to which it is (non-vacuously) applicable. It will thus serve the functions of a morpheme-structure rule or constraint with respect to underlying lexical representations, as well as serving the functions of a phonological rule or phonetic redundancy rule with respect to derived representations.

dialect which needs to be stated in its grammar, independently of any facts about umlauting and back vowel lowering. A grammar including the extrinsically unordered rules 20a–c, therefore, is clearly more adequate with respect to the language as a whole than a grammar including only the extrinsically ordered rules 20a–b.

A case precisely parallel to that of Schaffhausen is provided by the rules of Umlaut and Back vowel rounding proposed for certain Low German and Swiss German dialects by Kiparsky (1968:176, 199):

(23) a. UMLAUT: V → [−back] / ... Umlaut context

b. BACK VOWEL ROUNDING: $\begin{bmatrix} V \\ +back \\ +low \\ +long \end{bmatrix}$ → [+round] $\bar{a} \to \bar{ɔ}$

To account for the fact that the plural of *šwɔn* 'swan' is *šwæn*, rather than **šwɔ̄n*, Kiparsky proposes that the underlying form for 'swan' is *šwān*, and that 23a is extrinsically ordered before 23b. Given this ordering, 23a will bleed 23b of all low back vowels in umlaut contexts, thereby preventing the rounding of the low vowel in the plural form *šwæn*, while allowing such rounding for the non-fronted low vowel of the singular *šwɔn*.

As in the Schaffhausen case, however, all the data cited by Kiparsky indicate that these dialects have no low front rounded vowels at all, a restriction not uncommon in natural languages. Where the restriction was accounted for in the Schaffhausen dialect by a rule asserting that all front rounded vowels are non-low, the occurrence of *æ* rather than *ö* as the umlaut alternant of *ɔ* in the Low German and Swiss dialects suggests that this restriction follows in these dialects from the principle that all low front vowels are unrounded, a generalization that is explicitly expressed as follows:

(23) c. $\begin{bmatrix} V \\ -back \\ +low \end{bmatrix}$ → [−round]

Thus, given underlying *šwān*-PL, after 23a–b have applied simultaneously to yield **šwɔ̄n*, 23c will apply to this form to give the correct derived form *šwæn*. Clearly, then, 23a–c, without any extrinsic-ordering restrictions, account for all the facts about these dialects that are accounted for by Kiparsky's extrinsically ordered rules, and in addition express a significant generalization about their phonological structure which is not revealed by Kiparsky's rules.

These examples illustrate two general points concerning the extrinsic ordering of phonological rules. First, arguments suggesting the necessity of extrinsic ordering based on a very limited range of facts often fail completely as soon as additional facts about the language are brought into consideration. Second and more important, in a theory that excludes the possibility of extrinsic ordering, the linguist is forced at the outset to look for general explanatory principles which there would otherwise be little reason to look for. Theories of grammar which prohibit language-specific restrictions on the application of phonological rules thus provide a degree of stimulation and direction in the search for significant linguistic generalizations which is lacking in those theories which permit such restrictions.

2.6. We have attempted to show that a theory which prohibits extrinsic ordering, simply requiring that all phonological rules be applied according to a very small number of universal principles, is capable of explaining with equal or greater generality all natural language data which can be accounted for by means of rules extrinsically ordered in feeding, counter-bleeding, bleeding and counter-bleeding, counter-feeding, and bleeding relations. The remaining relations in which a pair of extrinsically ordered rules can stand are the following:

(24) a. Mutually non-affecting
  b. Feeding and Counter-bleeding
  c. Bleeding and Counter-feeding
  d. Feeding and Counter-feeding

Since the outputs of mutually non-affecting rules (24a) have no effect upon their respective inputs, there can be no possible empirical consequence of different orders of application. There are no problems, therefore, concerning the application of such rules in theories without extrinsic ordering, each rule being simply applied whenever it can be applied.

Rules standing in relations 24b and 24c are illustrated by the following artificial-language rule pairs:

(25) Feeding and Counter-bleeding
  a. ab → abab
  b. ab → ac.
(26) Bleeding and Counter-feeding
  a. ab → ac
  b. ab → abab.

To the best of our knowledge, no such rule pairs have been proposed for any natural language.[11]

Finally, the only known examples of rules standing in the feeding and counter-feeding relation are the subrules of 'alpha-switching' rules, such as the following proposed by Chomsky & Halle (356) for Biblical Hebrew:

$$(27) \quad \begin{bmatrix} + \text{ voc} \\ - \text{ cons} \\ \alpha \text{ low} \end{bmatrix} \rightarrow \begin{bmatrix} -\alpha \text{ low} \\ \alpha \text{ round} \\ + \text{ back} \end{bmatrix} / \underline{\quad} C + \text{IMPERFECT} \qquad \begin{array}{l} a \rightarrow o \\ o, e \rightarrow a \end{array}$$

If such rules are motivated at all, then it is clear, as Chomsky & Halle observe (357), that their subrules cannot be ordered with respect to each other; they must be restricted in such a way that, if one rule is applied to a given segment, the other rule must not be applied to any reflex of that segment in the same derivational cycle. Given this universal restriction, it is clear that there is no problem with respect to the proper application of such rules in grammars without extrinsic-ordering restrictions.

2.7. The claim in this section has been that, for all possible ways in which each member of a pair of extrinsically ordered rules can affect the application of the

---

[11] It should be noted that the order of application for 25 is predictable by the principle for counter-bleeding given in fn. 6, while the order of application in 26 is apparently not predictable by any principle proposed here or by any other known principle.

other, any actual natural-language data that can be accounted for with such an extrinsically ordered pair can be accounted for with equal or greater generality by rules that are free of any language-specific restriction on their order of application. There is every reason to believe that, if extrinsic ordering is unnecessary for pairs of rules, it will be equally unnecessary for triples, quadruples, or sets of rules of any number whatever.

Consider in this respect the set of seven extrinsically ordered phonological rules for Mohawk (28a) proposed by Postal (143–52),[12] and the alternative rules 28b which provide a simpler account of Postal's data without requiring any extrinsic-ordering restrictions at all.

28a. POSTAL'S RULES 28b. ALTERNATIVE RULES

TRUNCATION:

(1) $V \rightarrow \emptyset \; / \; \underline{\quad} \; V$ $V \rightarrow \emptyset \; / \; \underline{\quad} \; V$

PROTHESIS:

(2) $\emptyset \rightarrow i \; / \; _{\text{verb}}[\# \; \underline{\quad} \; D_0^n \, V \, D_0^n \, \#]_{\text{verb}}$ $\emptyset \rightarrow i \; / \; _{\text{verb}}[\# \; \underline{\quad} \; D^n \, V \, D_0^n \, \#]_{\text{verb}}$

EPENTHESIS:[13]

(<3) $\emptyset \rightarrow a \; / \; \dots \; \underline{\quad} \; \dots$ $\emptyset \rightarrow a \; / \; \dots \; \underline{\quad} \; \dots$

STRESS

(3) $V \rightarrow \hat{V} \; / \; \underline{\quad} \; D_0^n \, V \, D_0^n \, \#$ $V \rightarrow \hat{V} \; / \; \left\{ {\# \atop \hat{V}} \right\} D_0^n \; \underline{\quad} \; D^n \, V \, D_0^n \, \#$

STRESS JUMP:

(4) $V \, D_0^n \, \acute{a} \rightarrow \hat{V} \, D_n^0 \, \acute{a}$ (not necessary)

TONE:

(5) $\hat{V} \rightarrow \acute{\hat{V}} \; / \; \underline{\quad} \; \left\{ {P \atop h\,R} \right\}$ $\hat{V} \rightarrow \acute{\hat{V}} \; / \; \underline{\quad} \; \left\{ {P \atop h\,R} \right\}$

LENGTH:

(6) $\hat{V} \rightarrow \hat{V}{:} \; / \; \left\{ {\underline{\quad} \, C \, V \atop [\overset{\frown}{\quad}]} \right\}$ $\hat{V} \rightarrow \hat{V}{:} \; / \; \left\{ {\underline{\quad} \, C \, V \atop [\overset{\frown}{\quad}]} \right\}$

C = any consonant, resonant or not, including systematic $w \, y$.
D = any non-vowel, i.e. consonant or $h \, \text{?}$.
R = any resonant, i.e. $w \, y \, n \, r$.
$\overset{\frown}{\quad}$ = falling tone.

It should be noted that there are only two respects in which 28b differs from 28a—namely, in its treatment of the restriction in Mohawk against the stressing of epenthetic a's, and in the formulation of the environment of Prothesis and Stress. Neither modification results in any increase in complexity relative to Postal's formulations, or in any loss of generalization with respect to the given facts about Mohawk.

---

[12] These rules are used by Postal (151) to argue against the stratificational position that all rules of a grammar must apply simultaneously, and are taken to support the notion that 'without order, linguistic generalizations are necessarily lost'.

[13] Epenthesis is not explicitly formulated by Postal. It is clear, however, that he intends it to be ordered somewhere BEFORE Stress (an ordering symbolized here as '< 3'), since it is only under the assumption of such an ordering that his rule of Stress jump could have any motivation at all.

In Postal's rules, Stress applies to the penultimate vowel of the word, without distinction as to whether or not that vowel has been introduced by the previously applying rule of Epenthesis. However, since it is a fact about Mohawk that only non-epenthetic vowels may bear stress, a Stress jump rule must be assumed to shift the stress off epenthetic a and onto the first vowel on its left. The slash through the a in the output of Postal's Epenthesis rule then has the function of allowing Stress jump to distinguish between the two kinds of phonetically identical a, underlying versus epenthetic.

The revised rules account for the same restriction in a somewhat simpler and more natural manner. Unlike Postal's rules, the a introduced by Epenthesis is given no diacritic to distinguish it from a's which are stressed regularly. The stress rule itself, however, is formulated to capture directly the observation that a vowel can never be stressed if there is a stressed vowel preceding it in the same word. One effect of this quite natural modification of the context for Stress is that there is now no motivation for Stress jump, so it can be eliminated completely from the grammar. This modification also eliminates any need for Postal's diacritic marking of epenthetic a's, thereby simplifying the rule of Epenthesis and the vocabulary for element-types which must be assumed for the grammar of Mohawk.

The second difference between the two sets of rules concerns the environment for Prothesis and Stress. In Postal's rules, the vowel in the environment may be preceded by an indefinite number of non-vowels (hence the superscript n), or preceded by no non-vowels at all (hence the subscript 0); but we have found empirical justification for only the former specification, and accordingly have formulated the rules to be applicable before D^n (i.e. a string of one or more D's) rather than before D_0^n.

Consider now the derivation of certain representative Mohawk forms cited by Postal in arguing for the necessity of extrinsic ordering. In every case, the rules of 28b operate on the same underlying representations as proposed by Postal, and give the same forms generated by his extrinsically ordered rules 28a. The application of rules in the derivations below is governed only by the universal principle requiring that all (obligatory) rules that can apply to an underlying form do so, and that the structure thus derived be operated on in turn by all rules whose structural descriptions are satisfied by it, until a form is derived to which no rule is applicable, at which point the derivation reaches its natural termination.[14]

For an underlying form such as *tnieks* 'you and I eat it', the derivation would thus proceed as follows:[15]

(29) *tnieks*
    tneks   (Truncation)
    itneks   (Prothesis)
    ítneks   (Stress)

The fact that all rules apply sequentially here is not a consequence of any independent condition of the theory requiring that all phonological rules apply sequen-

---

[14] It is assumed that a rule cannot apply to a representation that meets its structural description if the application of that rule would yield an output identical to the input. In other words, vacuous application of grammatical rules is prohibited.

[15] The morpheme boundaries which are included in the underlying representations given by Postal are irrelevant to any of the rules of 28, and hence have been omitted here.

tially; rather, it is simply the case that, at every stage in this derivation, only one of the rules of 28b is applicable. The underlying form *tnieks* satisfies the structural description of Truncation only, hence that is the only rule to apply. The form derived by the application of Truncation, *tneks*, can be operated on only by Prothesis. The output of this rule, *itneks*, again satisfies the structural description of only one of the rules of 28b, i.e. Stress; the form *ítneks* is derived by the application of this rule. At this point the derivation terminates, for there are no rules in 28b which can apply to *ítneks*. (The final phonetic form *ídeneks* then re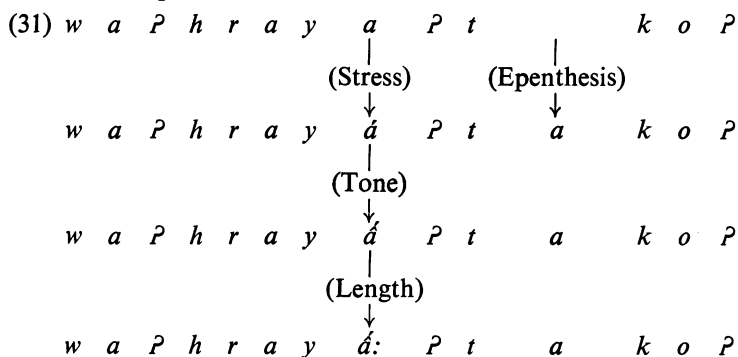sults from the application of rules other than those listed in 28, rules which Postal does not state and which are presumably irrelevant to any of the ordering arguments that he gives.)

Consider next the underlying form *nikanuhske* 'houses'. In the derivation of this form there are no intermediate stages (at least when only the rules of 28b are involved), but simply an underlying structure and a final derived structure:

(30) $n \quad i \quad k \quad a \quad n \quad\quad u \quad\quad h \quad s \quad\quad\quad k \quad e$

$\quad\quad\quad\quad\quad\quad\quad\quad\quad | \quad\quad\quad\quad\quad\quad\quad |$

$\quad\quad\quad\quad\quad\quad\text{(Stress)} \quad\quad\text{(Epenthesis)}$

$\quad\quad\quad\quad\quad\quad\quad\downarrow \quad\quad\quad\quad\quad\quad\downarrow$

$\quad\quad n \quad i \quad k \quad a \quad n \quad\quad ú \quad\quad h \quad s \quad\quad a \quad\quad k \quad e$
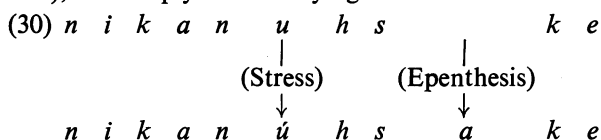
All the rules relevant to this derivation apply simultaneously, a fact which again is not determined by any independent statement in the theory, but is merely a consequence of the particular underlying forms, the structural descriptions of the particular rules, and the fact that the rules are obligatory. Of all the rules of 28b, only two are applicable to underlying *nikanuhske*, Stress and Epenthesis; accordingly, they apply simultaneously. Since the output includes a stressed vowel in the antepenultimate position, Stress cannot re-apply to the derived form. None of the other rules can apply either, and the derivation thus terminates. (The actual phonetic form *niganúhzage* would of course also require the application of other rules not listed in 28.)

As a final example, consider the underlying form *waʔhrayaʔtkoʔ* 'he picks bodies'. The phonetic form *waʔhrayá:ʔtakoʔ* is derived by the application of Postal's extrinsically ordered rules (and additional rules not considered convert this to *wahayá:dagoʔ*). Given the theory of rule application adopted here, the derivation would proceed as follows:

(31) $w \quad a \quad ʔ \quad h \quad r \quad a \quad y \quad\quad a \quad\quad ʔ \quad t \quad\quad\quad\quad k \quad o \quad ʔ$

$\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad | \quad\quad\quad\quad\quad\quad\quad\quad\quad |$

$\quad\quad\quad\quad\quad\quad\quad\quad\quad\text{(Stress)} \quad\quad\text{(Epenthesis)}$

$\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\downarrow \quad\quad\quad\quad\quad\quad\quad\downarrow$

$\quad\quad w \quad a \quad ʔ \quad h \quad r \quad a \quad y \quad\quad á \quad\quad ʔ \quad t \quad\quad a \quad\quad k \quad o \quad ʔ$

$\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad | $

$\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\text{(Tone)}$

$\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\downarrow$

$\quad\quad w \quad a \quad ʔ \quad h \quad r \quad a \quad y \quad\quad á̌ \quad\quad ʔ \quad t \quad\quad a \quad\quad k \quad o \quad ʔ$

$\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad | $

$\quad\quad\quad\quad\quad\quad\quad\quad\quad\text{(Length)}$

$\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\downarrow$

$\quad\quad w \quad a \quad ʔ \quad h \quad r \quad a \quad y \quad\quad á̌: \quad\quad ʔ \quad t \quad\quad a \quad\quad k \quad o \quad ʔ$

As in 30, Stress and Epenthesis are both applicable to the underlying form, and hence apply simultaneously. The output of these two rules satisfies the structural

description of Tone, but of no other rule; hence Tone is the only rule to apply at this point. After the application of Tone, again a single rule, Length, can and must be applied. The form *waʔhrayá:ʔtakoʔ*, then, is the final line of the derivation, since none of the rules of 28b are applicable to this form.

It has been shown that a set of six rules having no restrictions on their relative order of application can generate correctly, and without loss of generality, the forms derivable by Postal's proposed set of seven rules and six extrinsic-ordering constraints. There is every reason to believe that similar evidence of the non-necessity of language-specific constraints on rule application could be provided with respect to still larger sets of rules for Mohawk and other languages, and that the hypothesis of universally determined rule application is ultimately defensible with respect to all motivated phonological rules of all languages.

3. We have presented evidence suggesting that language-specific restrictions on rule application are not necessary for the principled explanation of any facts about particular natural languages. However, even if the hypothesis of extrinsic ordering is empirically indefensible with respect to particular languages, it would still be possible to maintain that the hypothesis is correct if it could be shown that the power of extrinsic ordering is necessary for the explanation of some facts about ALL languages, e.g. about the ways in which all temporally adjacent, genetically related languages resemble each other and differ from each other. Thus, if it can be shown that certain explanatory laws of language change depend essentially on the assumption that grammars include extrinsic-ordering restrictions, and if the facts accounted for by these laws are unexplainable in any other way, then the extrinsic-ordering hypothesis is strongly supported, even though it is apparently unnecessary with respect to data from any single language.

One such theory of language change has in fact been proposed, namely Kiparsky's (1968) theory of increasing feeding and decreasing bleeding. If this theory is correct, it would constitute a strong empirically based argument for the hypothesis of extrinsic ordering—in fact, the only strictly empirical argument that appears to have been presented thus far by the proponents of this hypothesis.[16] Hence it is necessary to consider Kiparsky's theory in some detail now, and to show that to the extent it generates correct predictions about language change, these predictions can

---

[16] Most of the explicit arguments that have been presented in support of extrinsic-ordering restrictions are based on questions concerning the relative simplicity of grammars with and without such restrictions, rather than on any differences in their (weak or strong) generative capacity, or in their ability to provide explanations of any actual facts about languages. Moreover, the simplicity arguments themselves are totally void of empirical import, since they are all crucially dependent upon the question-begging, a-priori assumption that non-universal RULES are counted as part of the axiomatic basis of a particular grammar, but non-universal CONSTRAINTS ON RULE APPLICATION are not. These arguments in fact simply ignore the crucial point, repeatedly made by Chomsky and others, that the 'choice of an evaluation measure is an empirical matter' (Chomsky 1965:37), and that it is impossible to use a single evaluation measure as a basis for choosing between alternative grammars that are governed by different theories of grammar—e.g. one theory that permits extrinsic-ordering constraints and another that does not. Or, as Chomsky has put it: 'It is also apparent that evaluation measures of the kinds that have been discussed in the literature of generative grammar cannot be used to compare different theories of grammar; comparison of a grammar from one class of proposed grammars with a grammar from another class, BY SUCH A MEASURE, is utterly without sense' (Chomsky 1965:38).

also be generated by an alternative theory which does not depend upon the hypothesis of extrinsic ordering.

In a theory with extrinsic ordering, rules standing in a feeding or counter-bleeding relation clearly have a more natural, more general, and more law-like character than those standing in a bleeding or counter-feeding relation. This follows from the fact that, in the case of feeding and counter-bleeding, each of the related rules expresses a true generalization which holds with respect to ALL the possible linguistic representations of some language; but in the case of bleeding and counter-feeding, the generalizations which are expressed by the related rules hold only with respect to SOME of the possible representations of a language, namely with respect to some proper subset of the set of representations which constitutes the derivational product of the application of some specified proper subset of the rules of its grammar. Other things being equal, therefore, a grammar which is free of bleeding and counter-feeding relations among its rules will provide a more general and scientifically more highly valued explanation of any body of linguistic data than any grammar which includes rules standing in either or both of the domain-reducing relations of bleeding and counter-feeding.

Moreover, if we may assume that the principles of hypothesis formation and hypothesis evaluation which govern the actual process of human language acquisition include principles determining the selection of hypotheses of the form 'All X are Y' over all descriptively equivalent hypotheses of the less general form 'All X which are Z are Y', then it follows that a child will consistently select, from the set of possible alternative hypotheses consistent with any given body of primary linguistic data, those which are maximally free of bleeding or counter-feeding relations. From this it follows that children might acquire grammars for the language of their community which include FEWER instances of bleeding and counter-feeding relations than the grammars possessed by the adult members of that community, but a child's grammar could never include MORE instances of bleeding and counter-feeding relations than those of the adult grammars which generate the sample of linguistic data to which the child has access. It thus also follows that languages might change over time in such a way that the grammars of their temporally subsequent dialects include fewer cases of bleeding and counter-feeding relations than those of their temporally antecedent dialects, but changes characterized by an increase in such domain-diminishing relations should never occur.[17]

Kiparsky presents a number of actual cases of historical change which are presumably consistent with the latter implication, and he cites no instances which are inconsistent with it. The facts he discusses would, therefore, count as empirical evidence in support of the premises from which this principle of historical change is deductively derived, namely (1) that it is necessary for grammars to include language-specific restrictions on the relative order of application of grammatical rules, and (2) that the actual process of language acquisition is restricted in such a way that, for any set of alternative hypotheses that are equally consistent with a given

[17] The version of Kiparsky's theory of language change outlined here differs in certain respects from the somewhat weaker, less explicit, and strictly probabilistic versions of Kiparsky 1968, 1972. These differences, however, are irrelevant to any of the issues of present concern.

body of linguistic data, the most general of these hypotheses is always selected. We take the second of these two premises to be established beyond reasonable doubt, since it is consistent with everything we know about language acquisition, and also with everything known about the acquisition and maintenance of other types of human knowledge and rule-governed behavior. It is only the first premise, therefore, which can reasonably be questioned. We intend to show here that this extrinsic-ordering assumption is not in fact a necessary premise for the deductive explanation of the facts about linguistic change that Kiparsky cites.

It will be possible to show that the assumption of extrinsic ordering is unnecessary only if we can show that all the dialect differences that Kiparsky discusses can be accounted for as a result of differences between grammars without extrinsic-ordering restrictions; and also that theories of grammar which prohibit extrinsic-ordering specifications are capable, in conjunction with their associated simplicity metrics, of generating the correct predictions about the types of linguistic change which do and do not occur in the domain of natural language. If this can be done, then the strongest of all known arguments for extrinsic ordering will have been successfully refuted.

3.1. The first test of the relative explanatory values of the hypotheses of extrinsic and non-extrinsic ordering with respect to facts about language change can be carried out with respect to Kiparsky's example (1968:197–8) involving the differences in deaffrication between Polish and Old Church Slavic on the one hand, and Russian and (most?) other modern Slavic languages on the other. Kiparsky claims that these differences provide support for his hypothesis that 'Feeding order tends to be maximized' (1968:197), since it is apparently possible to account for the differences as a consequence of different extrinsic orderings of the same two rules, where the ordering for the more conservative languages (Old Church Slavic and Polish) is counter-feeding, and that for the more innovative languages (Russian etc.) is feeding. The rules that Kiparsky posits are the following:

(32) a. DEAFFRICATION: $\begin{bmatrix} +\text{voiced} \\ -\text{grave} \\ +\text{strident} \end{bmatrix} \rightarrow [+\text{continuant}]$ $\quad\quad dž \rightarrow ž$
$\quad\quad dz \rightarrow z$

b. SECOND PALATALIZATION:

$\begin{bmatrix} +\text{obstruent} \\ -\text{grave} \\ -\text{strident} \\ -\text{diffuse} \end{bmatrix} \rightarrow \begin{bmatrix} +\text{strident} \\ +\text{diffuse} \end{bmatrix} / \underline{\quad\quad} ě$ $\quad\quad k,ě \rightarrow tsě$
$\quad\quad g,ě \rightarrow dzě$

For the more conservative languages, in which the reflex of *g,ělo* 'very' is *dzělo*, Kiparsky posits the extrinsic ordering of 32a before 32b, so that the latter rule will counter-feed rather than feed the former with respect to all instances of *dz* arising as a result of Second palatalization (see §2.4 above). For the more innovative languages, in which the reflex of *g,ělo* is *zělo*, Kiparsky posits the opposite extrinsic ordering, i.e. 32b before 32a, so that Second palatalization will always feed rather than counter-feed Deaffrication.

But as we have shown in §2.4, the relevant facts about Polish and Old Church Slavic can also be accounted for by positing, in place of 32a, a more restricted deaffrication rule:

$$(32)\ a'.\ \begin{bmatrix} +\text{voiced} \\ -\text{grave} \\ +\text{strident} \\ -\text{diffuse} \end{bmatrix} \rightarrow [+\text{continuant}] \qquad\qquad d\check{z} \rightarrow \check{z}$$

Thus the difference between the conservative and innovative Slavic languages here can be fully accounted for by assuming that the conservative languages have rules 32a' and 32b, while the innovative languages have 32a and 32b, where there is no ordering restriction on the application of either pair of rules. The rules of each dialect simply apply whenever their structural descriptions are satisfied. The derivation of *g,ělo* thus proceeds as follows:

(33) CONSERVATIVE SLAVIC   INNOVATIVE SLAVIC

g,ělo		*g,ělo*	
dzělo	(34b)	*dzělo*	(34b)
————	(34a')	*zělo*	(34a)

It will be observed that in this treatment, unlike Kiparsky's, the relevant difference between the more conservative Slavic languages and the more innovative ones is fully accounted for by a difference in the generality of their respective deaffrication processes: the conservative languages have rule 32a', which deaffricates only non-diffuse affricates like *dž*, while the innovative languages have the more general rule 32a, which deaffricates not only the non-diffuse affricates like *dž* but also the diffuse ones like *dz*. For theories which exclude the power of extrinsic ordering, therefore, the correct prediction about this change in the historical development of a language like Old Church Slavic into a language like Modern Russian follows from the very general and independently well-supported principle that grammatical rules become more but not less general during the transmission of a language from one generation to the next. Although Kiparsky's theory of phonological change also includes this general principle of rule generalization, which he refers to as 'rule simplification' (1968:176), he is unable to use this principle in accounting for the observed change in Slavic; instead, he must resort to an entirely distinct principle about the dominance of feeding over counter-feeding extrinsic-rule orderings during the course of language transmission. With respect to these facts, therefore, it is clear that a theory of grammar that excludes extrinsic ordering is consistent with a simpler and more general theory of linguistic change than that proposed by Kiparsky. His Slavic data thus serve to confirm rather than disconfirm the hypothesis that there is no extrinsic ordering of rules.

3.2. Another example of the adequacy of the hypothesis of non-extrinsic ordering with respect to diachronic facts is provided by Kiparsky's data (1968:178–9) about the difference between the Schaffhausen and Kesswil dialects of Swiss German, with respect to the quality of umlauted vowels preceding palatal and non-lateral dental or alveolar consonants. Kiparsky asserts that the phonetic forms for Schaffhausen

(34a) reflect an older stage of the language, while the contrasting phonetic forms for Kesswil (34b) represent a later innovative development:

    (34) a. Schaffhausen (conservative)

          Singular   Plural (umlauted)
            bogə   *bögə*
            bɔdə   *bödə*

        b. Kesswil (innovative)

          Singular   Plural (umlauted)
            bogə   *bögə*
            bɔdə   *bɔ̈də*

To account for these facts, Kiparsky posits identical underlying forms for both dialects (*bogə, bogə*-PL, *bɔdə, bɔdə*-PL), identical rules of Umlaut (35a) and Back vowel lowering (35b), and two different extrinsic-ordering restrictions on the application of these rules, the precedence of 35a over 35b being specified for the grammar of the Schaffhausen dialect, with the opposite precedence being specified for the Kesswil dialect:

    (35) a. Umlaut: V → [−back] / ... Umlaut context

$$\text{b. Back vowel lowering:} \begin{bmatrix} V \\ -\text{high} \\ +\text{back} \end{bmatrix} \rightarrow [+\text{low}] / \underline{\quad} \begin{bmatrix} +\text{cons} \\ -\text{grave} \\ -\text{lateral} \end{bmatrix}$$

$$o \rightarrow ɔ / \underline{\quad} \{t, d, \ldots\}$$

These rules and ordering restrictions are capable of generating the correct derivations for both dialects; and the restriction for the presumably more conservative Schaffhausen dialect determines a bleeding relation between Umlaut and Back vowel lowering, while the restriction for the more innovative Kesswil dialect determines a counter-bleeding relation. Hence Kiparsky's proposals are consistent with his general diachronic principle that 'Bleeding order tends to be minimized' (1968:199).

    However, as shown above in §2.5, the relevant facts about the Schaffhausen dialect can be adequately accounted for without any extrinsic-ordering restrictions whatever, simply by assuming that the grammar of this dialect includes, in addition to the umlaut and lowering rules of 35, an independently motivated rule 35c (= 20c), which expresses the generalization that all well-formed front rounded vowels in this dialect are non-low:

$$\text{(35) c.} \begin{bmatrix} V \\ -\text{back} \\ +\text{round} \end{bmatrix} \rightarrow [-\text{low}]$$

This rule is independently motivated for the Schaffhausen dialect, which has no *ɔ̈* vowels, but does not hold for the Kesswil dialect, which does have such vowels; thus a theory without extrinsic ordering is capable of accounting for the difference between these two dialects as a consequence simply of the presence of a certain segment-well-formedness restriction, present in the conservative dialect but lost in the innovative one. This reflects a type of linguistic change by rule loss which is quite natural and well-attested in other languages and language families.

By positing rules 35a–c for the grammar of Schaffhausen, and rules 35a–b alone for the grammar of Kesswil, the correct phonetic forms for both dialects can be derived from the same underlying forms without any language-specific restrictions on rule application for either grammar:

(36) SCHAFFHAUSEN   KESSWIL
    bogə-PL *bodə*-PL   *bogə*-PL *bodə*-PL
    bögə   *bɔ̌də*   *bögə*   *bɔ̌də*   (35a–b)
    ————   *bödə*   ————   ————   (35c)

Thus, where Kiparsky's treatment of the given facts about vowel changes in Swiss German requires us to assume ad-hoc, language-specific restrictions on rule application, as well as an ad-hoc principle of preference for certain ordering restrictions over certain others, the alternative treatment proposed here requires us only to assume that one of the ways in which a language can change is by a loss of a rule. Since this assumption is independently necessary for any adequate theory of linguistic change, it is clear that the power of extrinsic ordering is unnecessary for the correct prediction of the illustrated changes in Swiss German.

3.3. Another of Kiparsky's arguments for the explanatory utility of extrinsic ordering in diachronic linguistics is of exactly the same type as that involved in the Schaffhausen and Kesswil case. Here, instead of Umlaut and Back vowel lowering, the two rules which are claimed to participate in a reversal of ordering over time are Umlaut and Back vowel rounding:

(37) a. UMLAUT: $V \rightarrow [-back] / \ldots$ Umlaut context

b. BACK VOWEL ROUNDING: $\begin{bmatrix} V \\ +back \\ +low \\ +long \end{bmatrix} \rightarrow [+round]$ $\bar{a} \rightarrow \bar{ɔ}$

Kiparsky proposes that these two rules are both included in the grammars of each of two distinct classes of German dialects. For one dialect group, which he implies to be the more conservative, Kiparsky posits the bleeding ordering of 37a before 37b; for the other and presumably more innovative dialect group, he posits the counter-bleeding ordering of 37a after 37b. Given these two rules, and two distinct restrictions on their order of application, it is possible to derive the correct phonetic forms for each dialect group:

(38) a. Conservative dialects

    SINGULAR   PLURAL (umlaut context)
      šwān   *šwān*-PL
      ————   *šwæn*   (37a)
      šwɔn   ————   (37b)

  b. Innovative dialects

    SINGULAR   PLURAL (umlaut context)
      šwān   *šwān*-PL
      šwɔn   *šwɔn*-PL   (37b)
      ————   *šwɔ̈n*   (37a)

The apparent innovation here would thus again be consistent with the principle of minimization of bleeding orderings over time.[18]

However, as shown in §2.5, Kiparsky's treatment of the conservative dialects once again ignores the fact that these dialects apparently have no low front rounded vowels at all, which is naturally accounted for by this rule (= 23c):

$$(37)\ \text{c.}\quad \begin{bmatrix} V \\ -\text{back} \\ +\text{low} \end{bmatrix} \rightarrow [-\text{round}]$$

Given 37c, which expresses the generalization that all low front vowels are un-rounded (true for the conservative dialects here but false for the innovative ones), it follows again that the relevant facts about both dialect groups can be accounted for without extrinsic-ordering restrictions, the change underlying their observed differences being subsumed under the general principle of rule loss. Thus if the grammars of the conservative dialects include 37a–c, while those of the innovative dialects include only 37a–b, the correct surface forms for both dialects are derivable by theories that exclude extrinsic-ordering restrictions, as follows:

(39) CONSERVATIVE DIALECTS   INNOVATIVE DIALECTS

šwān	šwān-PL	šwān	šwān-PL	
šwɔ̄n	šwɔ̄n	šwɔ̄n	šwɔ̃̄n	(37a,b)
——	šwǣn	——	——	(37c)

3.4. In one more case cited by Kiparsky of purported phonological change by rule reordering (1968:199), the relative conservatism of the dialects involved is presumably known on linguistically independent grounds, and hence the diachronic adequacy of the non-extrinsic-ordering hypothesis can be put to further test. This involves the contrast in voicing and spirantization of post-vocalic consonants in a group of Alsatian, Bavarian, and Middle German dialects, which Kiparsky asserts to be the more conservative group, and in a group of other Modern German (especially Low German) dialects, which Kiparsky claims to be the more innovative group. This contrast is exemplified by the following paradigm:

(40) ALSATIAN (conservative)   LOW GERMAN (innovative)

tāk	tāx	'day'
tāɣə	tāɣə	'days'

To account for these data, Kiparsky proposes that Alsatian and Low German have the same underlying representations (*tāg* and *tāgə*) and the same two rules, 41a–b, but differ in that Alsatian has an extrinsic-ordering restriction requiring the bleeding application of 41a before 41b, while Low German has a restriction requiring the counter-bleeding application of 41a after 41b:

(41) a. DEVOICING: [+obstruent] → [−voice] / ___ #

b. SPIRANTIZATION: $\begin{bmatrix} +\text{stop} \\ +\text{voice} \end{bmatrix} \rightarrow [-\text{stop}] / V$ ___

---

[18] As suggested by Chafe (131), there is an alternative explanation which treats the change from the conservative to the innovative dialects simply as a consequence of the change of underlying *šwān* to underlying *šwɔ̄n*, with consequent loss of rule 37b. If this hypothesis is correct, then the question of accounting for the difference between these dialect groups in terms of extrinsic ordering differences obviously never arises.

These assumptions permit the derivation of the correct phonetic forms for both dialect groups, as shown in 42, and are in accordance with Kiparsky's prediction of decreased bleeding-ordering restrictions over time:

(42) a. ALSATIAN (conservative)

tāg	*tāgə*	
tāk	_____	(41a)
_	*tāγə*	(41b)

b. LOW GERMAN (innovative)

tāg	*tāgə*	
tāγ	*tāγə*	(41b)
tāx	_____	(41a)

There is, however, an equally reasonable account for these data which does not require the power of extrinsic ordering, and which permits the correct prediction of direction of change here to be derived from the ordinary principle of rule generalization over time. First, it can be observed that, for the innovative dialect group represented by Low German, 41a–b require no extrinsic-ordering restrictions at all, since for any representation such as *tāg* which satisfies the structural descriptions of both rules, the rules can always be applied simultaneously to that representation without contradiction, and the result will always be phonetically correct. That is, since the devoicing and spirantization of the *g* of *tāg* are mutually consistent processes, and since 41a and 41b are both obligatory, they apply simultaneously to this string to yield the correct *tāx* for Low German. The plural *tāγə* can also be correctly derived without extrinsic ordering, since it meets the structural description only of Spirantization.

For Alsatian and its sister conservative dialects, however, it is clear that, if there are no extrinsic-ordering restrictions, then the grammars of these dialects cannot possibly contain the same rules as those of the innovative dialects. Given Kiparsky's data, a very natural type of rule difference is indeed immediately suggested by the fact that spirantized stops occur POST-vocalically in Low German (*tāx*, *tāγə*), but they occur only INTER-vocalically in Alsatian and the other conservative dialects (*tāγə*, but not **tāγ* or **tāx*). Thus all the presently relevant facts about the conservative dialect group can be quite naturally accounted for by assuming that the grammars of these dialects include, rather than the (New) Spirantization rule 41b, the following rule:

(41) b'. OLD SPIRANTIZATION: $\begin{bmatrix} +\text{stop} \\ +\text{voice} \end{bmatrix} \rightarrow [-\text{stop}] / V \_\_\_ V$

It is logically impossible for 41b', which applies to segments followed as well as preceded by vowels, ever to be applicable to the same structure as 41a, which applies only to segments followed by non-vowels; hence their relative order of application cannot possibly have any empirical implications, and cannot be restricted in any empirically defensible way. These rules are wholly consistent with the hypothesis of non-extrinsic ordering.

It follows, therefore, that these two groups of German dialects differ only with respect to the relative generality of their respective spirantization processes: the conservative dialects define this process over the relatively more restricted domain

V [+stop, +voice] V, while the innovative dialects define it over the relatively more general domain V [+stop, +voice] X. The correct prediction of direction of change follows, in other words, from the general principle of rule generalization.

4. Evidence has been presented here showing that the power of extrinsic ordering is empirically unmotivated with respect to the explanation of synchronic and diachronic phonological facts. In §2, it was shown that, for representative facts which have been accounted for by each of the logically possible types of rule-ordering relations determined by extrinsic-ordering constraints, there are alternative explanations in which the relative order of application of rules is either entirely unrestricted, or else fully predictable from the forms of the rules by universal principles. In §3, it was shown that the power of extrinsic ordering is equally unnecessary for the explanation of various facts about phonological change. All predictions generated by Kiparsky's principles of diachronic rule reordering were shown to follow simply from the more general, independently well-motivated principles of rule generalization and rule loss over time. By showing that there is neither synchronic nor diachronic support for the hypothesis of extrinsic ordering, we have provided empirical support for the more restrictive hypothesis that all constraints on the relative application of phonological rules are determined by universal rather than language-specific principles of grammar.[19]

[19] A number of claims about extrinsic ordering have actually been based, however, not on any questions of fact or explanation, but rather on the use of different orderings of the same rules as an elegant or economical way of characterizing the difference between closely related dialects. For such proposals to have any empirical significance, it would be necessary to show that there is no reasonable way to characterize the observed dialect differences in terms of theories which prohibit extrinsic ordering. This has never been shown, as far as we can determine, for any of the proposed cases of rule-ordering alternations that have been presented in the literature.

Consider, e.g., the argument presented by Chomsky & Halle (342–3), in which certain observations reported for two dialects of English-based Pig Latin are adduced to support the hypothesis that dialectal variation may be a function of two extrinsically ordered rules standing in opposite orders. The crucial facts cited are the Pig Latin forms, in the two dialects, for $\bar{a}ys$ ('ice') and $s\bar{a}y$ ('sigh'): in Dialect I these forms are kept distinct, while in Dialect II they are homophonous. Chomsky & Halle propose to account for these facts by assuming rules (a) and (b), and by asserting that rule (a) is extrinsically ordered before (b) for Dialect I, but that for Dialect II these same rules are restricted to apply in the reverse order:

    (a) DIPHTHONG LAXING: $\bar{a}y \rightarrow \Lambda y \; / \; \underline{\hspace{1cm}} \; [-\text{voice}]$
    (b) PIG-LATINIZATION: $\# \# \, C_0 \, V \, X \, \# \# \rightarrow \# \# \, V \, X \, C_0 \, \bar{e}y \, \# \#$

According to Chomsky & Halle, the derivations of these forms in their respective dialects are as follows:

DIALECT I			DIALECT II		
$\bar{a}ys$	$s\bar{a}y$		$\bar{a}ys$	$s\bar{a}y$	
Λys	——	(a)	$\bar{a}ys\bar{e}y$	$\bar{a}ys\bar{e}y$	(b)
$\Lambda ys\bar{e}y$	$\bar{a}ys\bar{e}y$	(b)	$\Lambda ys\bar{e}y$	$\Lambda ys\bar{e}y$	(a)

These facts can be explained, however, with no loss of generality or complication of rules, if it is assumed that the Pig Latinization rule as given is a rule of Dialect II only, and that Dialect I has a different Pig Latinization rule which shifts the word-initial consonants to the RIGHT of the right-hand word boundary, instead of to the LEFT. (We owe this observation to Catherine Ringen.) That is, we assume rule (b) to be a rule of Dialect II only, with the equally general and well-motivated rule (b') holding for Dialect I:

    (b) Dialect II: $\# \# \, C_0 \, V \, X \, \# \# \rightarrow \# \# \, V \, X \, C_0 \, \bar{e}y \, \# \#$
    (b') Dialect I: $\# \# \, C_0 \, V \, X \, \# \# \rightarrow \# \# \, V \, X \, \# \# \, C_0 \, \bar{e}y$

In many of the cases discussed here, the extrinsic-ordering restrictions of a previously proposed grammar have been eliminated without changing any of the rules of that grammar or any of the empirical claims that it generated. In all other cases, the elimination of the original ordering restrictions was accompanied by only very minor changes in the originally proposed rules; these changes resulted in no loss of generality or explanatory power, and the proposed extrinsically unordered rules were no less natural or plausible than the original extrinsically ordered ones. Furthermore, in every case dealt with here, the facts have been accounted for without any recourse to diacritic elements, rule features, derivational constraints referring to more than two lines in a derivation, or special applicational markings of any sort. The evidence against language-specific constraints on rule application that is provided by these cases thus counts equally against theories which express such constraints by any one or more of these various notational alternatives.[20]

Although the arguments presented here in support of the hypothesis of universally determined rule application have been based solely on facts about phonology and phonological change, there is every reason to believe that there are parallel arguments with respect to syntactic facts, and that the appropriate application of syntactic rules is determined by the same set of universal principles that determine the proper application of all phonological rules.[21] We believe, therefore, that extrinsic ordering is not necessary for the explanation of any facts whatever about natural languages.

REFERENCES

CHAFE, WALLACE L. 1968. The ordering of phonological rules. IJAL 34.115–36.

CHOMSKY, NOAM. 1964. Current issues in linguistic theory. The structure of language, ed. by Jerry A. Fodor and Jerrold J. Katz, 50–118. Englewood Cliffs, N.J.: Prentice-Hall.

——. 1965. Aspects of the theory of syntax. Cambridge, Mass.: MIT Press.

——. 1967. Some general properties of phonological rules. Lg. 43.102–28.

——, and MORRIS HALLE. 1968. The sound pattern of English. New York: Harper & Row.

IVANOV, V. V. 1964. Istoričeskaja grammatika russkogo jazyka. Moscow: Prosveščenie.

IVERSON, GREGORY K. 1973. A guide to sanguine relationships. Paper presented at the Indiana University Conference on Rule Ordering.

KENSTOWICZ, MICHAEL, and CHARLES KISSEBERTH. 1970. Rule ordering and the asymmetry hypothesis. Papers from the 6th Regional Meeting of the Chicago Linguistic Society, 504–19.

Assuming these to be the rules in the respective dialects, it is clear now that we need not consider any of these rules extrinsically ordered. Consider the derivations of *sāy* in Pig Latin I and II, where in both cases the rules apply whenever their structural descriptions are satisfied (note that the derived form *āy##sēy* of Dialect I cannot undergo Diphthong laxing, since that rule is noncyclic and hence applies only within words):

DIALECT I		DIALECT II	
sāy		*sāy*	
āy##sēy	(b′)	*āysēy##*	(b)
		ʌysēy##	(a)

[20] See fn. 1 for specific references.

[21] See Koutsoudas 1971a,b, 1972, Lehmann 1972, and Ringen 1972 for evidence of the nonnecessity of extrinsic ordering in syntax. See Sanders 1970 for examples of the use of the Proper Inclusion principle for predicting applicational precedence relations between syntactic rules.

KIPARSKY, PAUL. 1968. Linguistic universals and linguistic change. Universals in linguistic theory, ed. by Emmon Bach and Robert T. Harms, 170–202. New York: Holt, Rinehart & Winston.
——. 1972. Historical linguistics. A survey of linguistic science, ed. by W. O. Dingwall, 576–649. College Park, Md.: University of Maryland, Linguistics Program.
——. 1973. Abstractness, opacity, and global rules. Paper presented at the Indiana University Conference on Rule Ordering.
KOUTSOUDAS, ANDREAS. 1971a. Gapping, conjunction reduction, and coordinate deletion. Foundations of Language 7.337–86.
——. 1971b. On the non-sufficiency of extrinsic order. Mimeographed, Indiana University Linguistics Club.
——. 1972. The strict order fallacy. Lg. 48.88–96.
LAKOFF, GEORGE. 1971. On generative semantics. Semantics, ed. by Danny D. Steinberg and Leon A. Jakobovits, 232–96. Cambridge: University Press.
LEHMANN, TWILA. 1972. Some arguments against ordered rules. Lg. 48.541–50.
McCAWLEY, JAMES. 1968. The phonological component of a grammar of Japanese. The Hague: Mouton.
NEWTON, B. E. 1971. Ordering paradoxes in phonology. Journal of Linguistics 7.31–53.
NORMAN, LINDA J. 1972. Rule addition and intrinsic ordering. Paper presented at the Winter meeting of the Linguistic Society of America. Minnesota Papers in Linguistics and the Philosophy of Language, 1.
PERRY, THOMAS. 1971. On a spurious case of counter-feeding order. MS.
POSTAL, PAUL M. 1968. Aspects of phonological theory. New York: Harper & Row.
RINGEN, CATHERINE. 1972. On arguments for rule ordering. Foundations of Language 8.266–73.
SANDERS, GERALD A. 1970. Precedence relations in language. Paper presented at the Summer meeting of the Linguistic Society of America. (To appear in Foundations of Language.)
SAPORTA, SOL. 1965. Ordered rules, dialect differences, and historical processes. Lg. 41.218–24.
SHEVELOV, GEORGE Y. 1965. A prehistory of Slavic. New York: Columbia University Press.
STICKELBERGER, HEINRICH. 1881. Lautlehre der lebenden Mundart der Stadt Schaffhausen. Aarau.
WANG, WILLIAM S-Y. 1969. Competing changes as a cause of residue. Lg. 45.9–25.