IN DEFENSE OF EXTRINSIC ORDERING

ROBERT D. KING

University of Texas at Austin

The question of how rules are organized in a grammar is tightly bound up with a number of lingering problems in phonology, notably abstractness. As long as there is a debate about the permissible 'distance' between underlying and superficial representations, there will be a debate on the number and types of rules needed in a grammar and on how these rules interact. No less vexed, and no less directly relevant to rule ordering, is the question of precisely what kind of evidence should be required to motivate phonological rules. Everyone agrees that alternations matter most, but some alternations matter more than others in the practice of most phonologists; to cite but one example, there is a distinct tendency to prefer inflectional to derivational alternations as evidence for a rule. Zwicky (1972b) raises the question of the worth of phonological variants relative to distributional restrictions as evidence for phonological rules. Skousen (1972) asks whether 'productivity' is a necessary condition for a phonological rule. Further, to make a murky situation even murkier, when one reflects on Stampe's (1969, 1972) penetrating observations and their implications for the overall picture of both synchronic and diachronic phonology, it seems only reasonable to be confused about which of the still unresolved questions in phonology a linguist should confront first, or whether any one topic, such as ordering, can be effectively dealt with in isolation. I remain uncertain as to whether ordering should be given priority over the host of other unsettled questions in phonology.

As sympathetic as I am in principle to recent efforts directed toward reducing or eliminating altogether the amount of (extrinsic) ordering permitted in a theory,[1] I am not convinced that durable advances have been made. Naturally, I am in agreement with the effort to determine, where possible, general principles that sequence rules in grammars. A theory that constrains on principled grounds the number of possible derivations in a grammar is assuredly to be preferred over one that makes no predictions concerning which sequences

of rules are to be expected. My feeling, though, is that we would have to give up too much – too much both in our ability to account for synchronic relationships and in our ability to explain historical change – if we accept the revisionist theories of ordering that I have seen. More specifically, I claim that the attempts to dispense in one way or another with extrinsic ordering have not come to grips with a sufficiently rich assortment of synchronic data,[2] that it has not been adequately considered whether the principles adduced to sequence rules are plausible as UNIVERSAL principles underlying the organization of grammars,[3] and that the consequences the proposed theories have for diachronic change have not been sufficiently thought through. Although each of these objections merits development at length, I will deal mainly only with the latter one in this paper.

1. The basic definition of extrinsic ordering (Chomsky, 1965, 223) is that rules are sequenced by an explicit statement rather than by the formulation of the rules; rules sequenced solely by virtue of their formulation are said to be intrinsically ordered. In the recent literature the definition of intrinsic ordering has been extended to include the application of rules governed by universal principles which permit one sequence but disallow the other.

It would be very easy to score points against a theory claiming that 'Rules are extrinsically ordered, period' since, as Kiparsky (1968b) and Chafe (1968) were the first to point out, preferred rule sequences do in fact exist. A theory of extrinsic ordering that simply ignores the observed asymmetries of rule order relationships would be bad in a number of ways, quite apart from the basic fact that a theory must account for recurrent ordering relationships. Such a theory would not predict that changes in rule order might occur over time, to cite just the one weakness of this kind of theory that is most germane to my considerations.

I would like to develop this point somewhat further. The relationship between synchronic grammars and phonological change that evolved in the late 1960's may be characterized for the present purposes in the following way. It is assumed that rules are extrinsically ordered in the grammar of a language; however, it is recognized that certain ordering relationships (e.g., feeding and nonbleeding relationships in Kiparsky's (1968b) proposal) are more natural than their opposites. Given the assumption, supported by a wide variety of well known facts about language acquisition and attested historical changes, that certain types of internal linguistic evolution are governed by Simplification, various kinds of correct predictions about expected developments can be made. It is generally agreed now that the principle of Maximal Utilization of Rules is incorrect (see, e.g., Kenstowicz and Kisseberth, 1973a, and King, 1973, 553-59) and must be replaced by other principles. Two principles for predicting expected developments that seem promising currently are the principle that opacity is minimized, and a principle stating that it is more likely for

rules that interact in the formation of a paradigm to apply in the order which minimizes allomorphic variation (i.e., the Paradigm Condition).[4] It is probable that other principles, perhaps some of those that have been proposed in the literature on intrinsic ordering, will play a role in judging the complexity of rule sequences. It is also probable that there will be competition between principles which make opposing predictions, so that simplicity of rule sequences cannot be determined by appeal to any one principle alone but is a GLOBAL property of the entire grammar.

Whatever the status and validity of the principles that have been proposed to account for preferred rule orders, the 'standard' view of the relation between synchrony and diachrony is that the rules of a grammar are extrinsically ordered but that certain orders of rules are more highly valued than others. A logical consequence of this theory is that reordering to the more highly valued sequences is to be expected.

The theory of ordering that I am defending is thus not simply the theory stating that rules are extrinsically ordered. Rather it is the theory stating that rules are extrinsically ordered but that there is an evaluation metric which provides a measure of the instrinsic complexity of any particular grammar. One of the better known clauses in the metric assesses the complexity of a rule relative to the complexity engendered by a different formulation of the rule; another clause, the one of interest here, assesses the complexity of pairs of rules, relative to specific inputs, on the basis of opacity, paradigm uniformity, and whatever other factors turn out to have significance.

There is, of course, nothing novel about the outline of the theory that I have been discussing. I merely wanted to make it clear how the theory of generative phonology that I wish to defend, i.e., the theory that has extrinsically ordered rules but with a metric that places higher value on certain orders, differs from the theories of no extrinsic ordering which have been proposed.[5] I will refer to the theory I support as the 'standard theory'. (I recognize that this is somewhat inaccurate, but I am prepared to argue that the theory I have sketched here is essentially the one that, implicitly at least, emerged at the end of the 1960's.) The major difference is that the latter theories build into the organization of the grammar principles that in the standard theory would be regarded as merely evaluative, and not absolutely binding.

In principle, since a theory that does not allow extrinsic ordering (or, as in Kisseberth (1973a), allows it only in a small set of cases) makes far more restrictive claims than the standard theory (which would ALLOW virtually any sequence, at a cost determined by the evaluation metric), it should be very easy to falsify such a theory: one would simply have to produce a sequence of two rules not predicted by the universal principle. In practice, the theories of intrinsic ordering that I have seen contain subsidiary principles that effectively neutralize the force of counterexamples. In Kisseberth's (1973a) theory, Minimization of Opacity is taken as the universal sequencing principle. Two rules

applying in an order which does NOT minimize opacity should therefore be sufficient to disprove the theory. This does not necessarily follow, however, because in such cases the universal principle is abandoned and a global rule reanalysis proposed; and if there is not a phonological basis for a global rule, then the two rules are sequenced by a priority statement, i.e., they are extrinsically ordered. In the Koutsoudas, Sanders, and Noll (1971) (henceforth KSN) theory, whose fundamental underlying assumption is that rules apply random sequentially,[6] one should be able to disprove the theory, in its pristine form, by producing cases – there are innumerable ones[7] – not correctly sequenced by this convention of rule application. But as work in the KSN framework has progressed, subsidiary constraints have been proposed (e.g., simultaneous application if two more rules could apply to a single representation, Proper Inclusion Precedence, Counterbleeding Precedence) which have the effect of extending the limited range of derivations that are possible on the assumption that rules are ordered random sequentially. These extensions make it far more difficult, of course, to disprove the theory, perhaps even impossible.[8]

What I have been calling the standard theory makes weaker claims in that rule sequences violating an ordering restriction in the metric are tolerated; but they are marked as nonoptimal, implying that they are exceptional. The prediction, then, is that such sequences will be reversed in contexts where we have come to expect simplifications, e.g., in children's speech, and perhaps in casual speech as well (Zwicky, 1972a, 279-81), and that the nonoptimal sequences will disappear (most commonly through reordering or loss, though there are other mechanisms) given enough time. I assert that it is exactly this kind of weaker claim that the theory SHOULD make, and that anything stronger makes wrong predictions about phonological change or makes no prediction where one is clearly warranted.[9]

I am also asserting, implicitly, that NO principle of intrinsic ordering – at least none of those found in the literature to date – is sufficient to enforce exceptionless sequencings and is still plausible as a universal mechanism of grammar (see footnote 3). If Minimization of Opacity is taken as a universal principle, then why, in case after case, is opacity not only tolerated in a grammar but is actually CREATED through internal linguistic change? Instances of phonological change in which this has happened are legion, and it will suffice here to discuss an example from the history of German (see King, 1972b).

Middle High German had a Syncope rule which deletes unstressed *e* in various environments, in particular between *h, m,* and *n* and *t* (see Paul and Mitzka, 1966 §47). From the paradigm of *wonen* 'to live, dwell', we have: 3rd per. sg. pres. ind. *wonet* → *wont,* 1st per. sg. past ind. *wonete* → *wonte* (stress falls on the first vowel in the examples we discuss here, so that *e* in post-tonic syllables is unstressed); but there is no syncope of unstressed *e* in, for example, the infinitive of 'to live, dwell' (*wonen*) or in 1st per. sg. pres.ind. *wone.* Subsequent to Classical Middle High German, a rule was added which

lengthens vowels in stressed open syllables (A. Bach, 1965 § 118; Paul and Mitzka, 1966, § 46); at the same time, original long vowels in closed syllables were shortened. At this stage of the language, the two rules of concern to us here applied in the order Syncope–Lengthening. Derivations for the cited forms of the model verb *wonen* goes as follows:

(1)	Underlying:	/wonen	wone	wonet	wonete/
	Syncope:	---------	-------	wont	wonte
	Lengthening:	wōnen	wōne	-------	----------
	Phonetic:	[wōnen	wōne	wont	wonte]

Subsequently, the rules were reordered, giving:

(2)	Underlying:	/wonen	wone	wonet	wonete/
	Lengthening:	wōnen	wōne	wōnet	wōnete
	Syncope:	---------	-------	wōnt	wōnte
	Phonetic:	[wōnen	wōne	wōnt	wōnte]

The chronology of the two rules is well established; and a number of North German dialects apparently continue to have the older order as in (1) – see Behagel (1928, 276). The majority of dialects today, Standard German in particular, are descended from (2).

This example is relevant to the claim that opacity is minimized in the following way. Observe that Lengthening is transparent in the original order (1): long vowels are found in and only in open syllables; short vowels are found in and only in closed syllables. In (2), however, forms like [wōnt] and [wōnte], which have long vowels in closed syllables, demonstrate that Lengthening has become opaque. Notice, though, that the surface paradigm in (2) is more regular than in (1) to the extent that a long/short vowel variation has been replaced by allomorphs with an invariant long vowel.

Most reorderings proceed in just this way, i.e., from transparency to opacity; paradigm regularity is gained at the expense of making a transparent rule opaque (most of the reorderings in Kiparsky (1968b) are of this type). The basic point, made at greater length in Dinnsen and King (1973), is that if Minimization of Opacity is indeed a universal sequencing principle for languages, as Kisseberth (1973a) argues, then it should be a target of preference in phonological changes that are due to reordering. But contrary to the theoretical prediction, cases of the sort just described abound, whereas it is exceedingly difficult to find good cases of the kind that support Kisseberth's theory; that is, it is difficult to find good cases of reordering that MINIMIZE opacity, say at the expense of a paradigm (see King (1973, 565-6) for a possible but disputed case from Icelandic). Is it really justified, with so dramatic a disparity between what the theory proposes and what actually happens in so many in-

stances, to base a theory on the assumption that Minimization of Opacity is a UNIVERSAL sequencing principle?

In summary, a theory like Kisseberth's (1973a), which regards transparency as the 'unmarked' condition for rule interaction, makes the wrong prediction in the large number of cases (the German example is representative) in which transparency gives way to opacity for the sake of the paradigm. In other instances, discussed more fully in Dinnsen and King (1973), the theory makes no prediction of change where a prediction seems clearly warranted. I would now like to consider how the theory of intrinsic ordering proposed in KSN would treat this kind of example, but before doing so I will digress and dispose of a question that is often raised when reordering is advanced as the explanation of some historical change.

Notice that in (2) there is a long vowel in every form of a verb like *wōnen* after the reordering occurs. There are no longer any long/short vowel alternations in the paradigm; and it would be more than reasonable to propose that the underlying forms have been relexicalized in the innovating dialect. The underlying forms contain long vowels, the Lengthening rule is no longer needed, and the derivation would be simply:

(3)	Underlying:	/wōnen	wōne	wōnet	wōnete/
	Syncope:	---------	-------	wōnt	wōnte
	Phonetic:	[wōnen	wōne	wōnt	wōnte]

It is completely plausible to me that something of this sort is the correct solution ONCE THE REORDERING HAS OCCURRED.[10] I claim, though, that the passage from (1) to (3) presupposes the existence of the intermediate stage (2). Consider the data presented to the learner of the language at stage (1): long and short vowels alternate under conditions that are completely transparent. Later learners of the language with system (3) have overridden the data, obviously; but what mechanism of change accounts for the exact way in which the data were overridden? If one is an adherent of the traditional theory of proportional analogy, one might simply say that the short vowel forms were leveled to agree with the long vowel forms in the paradigm. But we are then entitled to ask why weren't the long vowel forms leveled to agree with the short vowel forms? Without additional qualifications, the traditional theory of proportional analogy assigns equal probabilities to the two kinds of leveling. However, recent work has made it possible to narrow somewhat the possibilities for analogical change. The change from system (1) which did NOT occur, namely the generalization of the short vowel throughout the paradigm, would be the result of losing, in (1), the Lengthening rule. But Lengthening is a completely transparent rule in (1), and the loss of such a rule would be incompatible with a certain principle about rule loss that has been tentatively established: opacity, by the first branche of Kiparsky's definition, is a neces-

sary condition for rule loss; transparent rules cannot be lost (except, as seems to be true, under the force majeure of a competing language).[11]

If this constraint on rule loss is correct, then we have an explanation for why the data at the phonetic level in (1) could not have been overridden in one of the two ways that the traditional idea of analogy would predict, namely, by extending the short vowel to the whole paradigm. But we now have to explain what set the stage for the way the data were in fact overridden, as represented by the end result (3). I suggest that the transition from (1), with lexically short vowels and a transparent Lengthening rule, to (3), with lexically long vowels and no Lengthening rule, is possible only through an intermediate stage. This stage must be minimally different in the productive mechanisms and underlying forms from (1), in which the new learner of the language was presented only with long vowels in the output and in which the rule that has been lost in (3) – the Lengthening rule – is opaque. The only system that satisfies these conditions is (2), which is minimally different in underlying forms and rules from (1), with a Lengthening rule which has become opaque by reordering and hence subject to loss, and which has the same output as (3).

The point of this digression was to anticipate one of the objections that often is expressed when a case of reordering is proposed, namely, that rather than reordering, the progression is from an earlier system with two rules in one order to a later system with different underlying forms and only one rule. Aside from the fact that there are cases (I will discuss one below from Old English, and there are others) in which the alternations involved do not permit a lexical reanalysis because no single form is generalized throughout the paradigm, in general, the argument fails because of the 'transition problem': there is no motivated way to proceed from the earlier to the later stage without passing through an intermediate stage with reordering. My argument for this conclusion is based on what I consider the only reasonable assumption regarding the system – in the German example, (2) – which could have produced the kind of phonetic data from which the final grammar – in this example, (3) – could have been 'projected' (see Peters, 1972).

Let us now consider how the theory of intrinsic ordering proposed in KSN would deal with the German example. Their fundamental assumption that rules apply random sequentially gives rise to a dilemma in that both rules, Syncope and Lengthening, apply to an output like *wonet*. The KSN convention for handling this situation is to allow the rules to apply simultaneously; this will produce *wōnen wōne wōnt wōnte,* which is the correct output for the innovating dialects. But how do we account for the conservative dialects with *wōnen wōne wōnt wōnte*? I assume, based on the characteristics of the alternative analysis that KSN propose in similar cases (specifically, in reference to Kiparsky's examples of reordering from Swiss German dialects), that KSN would argue that the conservative dialects had a THIRD rule, Shortening, which shortens vowels in closed syllables.[12] If all three rules – Syncope,

Lengthening, and Shortening – apply to the underlying forms /wonen wone wonet wonete/ whenever their structural analysis is met, i.e., in the random sequential mode, then the correct outputs will be obtained: [wōnen wōne wont wonte]. The two stages in the development of German are then characterized as follows:

(4)	OLD STAGE	NEW STAGE
	Syncope	Syncope
	Lengthening	Lengthening
	Shortening	

We could then say that the change here was simply the loss of the Shortening rule. I assume that something of this kind would be close to the KSN analysis for the phonological change in question.

But there is a fatal flaw in this argument, quite apart from the fact that the 'extra rule' reanalysis so favored by KSN is, in general, highly suspect (see note 12). In the earlier stage of German for which the Shortening rule is proposed, the rule is completely transparent: there are NO exceptions to the distributional generalizations that long vowels occur only in open syllables and short vowels occur only in closed syllables. The research to date on rule loss (see note 11) suggests that transparent rules cannot be lost (does anyone anticipate that English will lose the rule aspirating voiceless stops in initial position?). Rather than it being the case that rule loss is, like rule addition, a kind of change that 'just happens', as the discussion in King (1969, 46-51) makes it appear to be, it is only after a rule has become opaque that it CAN be lost. (Furthermore, opacity by branch one of Kiparsky's definition seems to contribute far more importantly to the potential for loss than opacity by the second branch of the definition) If this diachronic constraint is correct, then rule loss cannot be invoked to explain the change that the KSN theory would require.[13]

I would now like to consider another example of a similar sort but with an interesting difference. The paradigm for 'day' in Old English is:[14]

(5)	SINGULAR	PLURAL
Nom.-Acc.	dæj	dag as
Gen.	dæjes	dag a
Dat.	dæje	dag um

Earlier, this paradigm had the form:[15]

(6)	SINGULAR	PLURAL
Nom.-Acc.	dæx	dag as
Gen.	dæjes	dag a
Dat.	dæje	dag um

The only difference between the two paradigms is the nominative-accusative singular [dæx] > [dæj]. In a theory with extrinsic ordering the explanation for what has happened to change this form is straightforward. The underlying form of the root for 'day' is /dæg / in the grammars of both (5) and (6). (The front/back vowel alternation between singular and plural is conditioned by the back vowels in the plural suffixes.) Notice that [ɡ] is manifested on the surface in both (5) and (6), so that unlike the German example, there can be no disagreement about whether the second stage of the language has undergone relexicalization: it cannot have. Two rules, both of them independently motivated, interact in the formation of the paradigm we are discussing:

(7) Devoicing: Fricatives are devoiced word-finally.

(8) Vocalization: The voiced velar fricative [ɡ] becomes the glide [j] following front vowels.

Selecting the nominative-accusative and genitive singular, and the nominative accusative plural as representative types in the paradigm, the system of rules underlying the earlier paradigm (6) was:

(9)				
	Underlying:	/dæg	dæg+es	dag+as/
	Devoicing:	dæx	----------	----------
	Vocalization:	------	dæjes	----------
	Phonetic:	[dæx	dæjes	dag as]

Devoicing and Vocalization are subsequently reordered, as in (10), producing the later and attested forms in the paradigm (5):

(10)				
	Underlying:	/dæg	dæg+es	dag+as/
	Vocalization:	dæj	dæjes	----------
	Devoicing:	-----	---------	----------
	Phonetic:	[dæj	dæjes	dag as]

The reason for the reordering is analogical: the amount of allomorphic variation is reduced from [dæx ~ dæj ~ dag] to [dæj ~ dag]; the difference in the root-final consonant now correlates symmetrically with the singular-plural distinction. The reordering is predicted from the Paradigm Condition (Kiparsky 1971a, 596-99) which, I have claimed, is a component of the evaluation metric of a grammar of extrinsically ordered rules.[16]

The relation between Devoicing and Vocalization is one of mutual bleeding with respect to the nominative-accusative singular form /dæg/: Devoicing converts [ɡ] to [x], thus blocking Vocalization; Vocalization converts [ɡ] to [j], thus blocking Devoicing. It is for just this situation of mutual

bleeding that Kisseberth (1973a) proposes the use of extrinsic ordering of a restricted type, i.e., a priority statement requiring one order in one grammar but the opposing order in the other grammar. Minimization of opacity fails here because the degree of opacity is the same in either order, and as far as I can see, there is no 'phonological basis' (in Kisseberth's (1973a) sense of the term) for a global rule. In such cases, his theory requires priority statements. To account for the historical development that has been observed, then, we will have to say that one priority statement ('Devoicing precedes Vocalization') is replaced by another ('Vocalization precedes Devoicing').

But WHY should one priority statement be replaced by another one? Since by definition, the basis for a priority statement is 'nonphonological' in Kisseberth's theory, why should one order be preferable in any way to the other? Unless one is willing to allow the Paradigm Condition into the theory as one of the universal sequencing principles, in which case a hierarchy of preference between the Paradigm Condition and the Opacity Condition would have to be established, since the two principles often make opposing predictions (and the Paradigm Condition would obviate the need for most global rules), Kisseberth's theory cannot render a satisfactory explanation of this change which, as we have seen, is accounted for in a perfectly reasonable and straightforward way in a theory that allows extrinsic ordering.

How would this change be accounted for in the KSN theory? Their basic move, given two rules that stand in a mutual bleeding relationship ('bleeding and counterbleeding' in their terms), is to sequence the application of the rules by the principle of Proper Inclusion Precedence. Rewriting the two rules in question, (7) and (8), in feature format, we obtain (7') and (8'):

(7') $\begin{bmatrix} +\text{obstruent} \\ +\text{continuant} \end{bmatrix} \rightarrow [-\text{voice}] \;/\; \underline{\quad} \;\#$

(8') $\begin{bmatrix} +\text{obstruent} \\ -\text{anterior} \\ +\text{continuant} \\ +\text{voice} \end{bmatrix} \rightarrow \begin{bmatrix} -\text{consonantal} \\ -\text{back} \end{bmatrix} / \begin{bmatrix} V \\ -\text{back} \end{bmatrix} \underline{\quad}$

To apply Proper Inclusion Precedence we compare the structural analysis of (7') with that of (8''):

(7'') $\begin{bmatrix} +\text{obstruent} \\ +\text{continuant} \end{bmatrix} \#$

(8'') $\begin{bmatrix} V \\ -\text{back} \end{bmatrix} \begin{bmatrix} +\text{obstruent} \\ +\text{continuant} \\ -\text{anterior} \\ +\text{voice} \end{bmatrix}$

(8'') $\begin{bmatrix} V \\ -\text{back} \end{bmatrix}$ $\begin{bmatrix} +\text{obstruent} \\ +\text{continuant} \\ -\text{anterior} \\ +\text{voice} \end{bmatrix}$

The order of application should be determined by whether (7'') properly includes (8'') or vice-versa. It appears to me that neither relationship obtains; rather, the structural analyses of the two rules are disjoint, and it is not possible, from Proper Inclusion Precedence alone, to sequence the rules.[17]

No matter what Proper Inclusion Precedence does or fails to do in this case, it should be clear that there is a serious dilemma, for whichever order Proper Inclusion Precedence dictates, the opposite order must be accounted for as well. If anything about ordering is clear, surely it is that a given principle should not predict two opposite orders. Any move to evade the dilemma by reanalysis seems ill advised to me, for the rules are motivated, and the requisite alternations are there; hence there is no abstractness problem. In short, I do not see how this example – or any other case of mutual bleeding in which different dialects or different historical stages of a language have rules in opposite orders – would be explained in the KSN theory.

Let me assume for the sake of argument that there is a way of handling the two stages of Old English in the KSN theory, say by positing a rule for the earlier stage that later gets lost. I would still claim that there is a very serious flaw of a general type in their explanation of historical changes (section III of KSN). There are, of course, a wide range of possible changes that may occur as a language develops: sound changes, morphological analogy, morphophonemic analogy, and restructurings of various kinds. After all the extraneous issues are subtracted, most of the cases discussed in section III of KSN are fundamentally of a single type: they involve morphophonemic analogy.[18] This uniformity is simply not expressed in the KSN theory. In a theory with extrinsic ordering, for these cases of morphophonemic analogy, there is ONE motivation – paradigm regularity, and there is ONE mechanism – reordering.[19] This was seen in my discussion of the examples from German and Old English. In the KSN theory, completely different kinds of explanations must be invoked, each of them dependent on accidental facts that happen to be at hand in the different instances: rule loss (more precisely, loss of a transparent well-formedness condition) in the two Swiss German examples, an explanation whose fragility was commented on earlier, and rule generalization in the Alsatian case.[20] But exactly the same fundamental process, morphophonemic analogy, is going on in all three cases, as well as in the two cases discussed here. I submit that a theory providing a single motivation and a single mechanism to account for cases that are exactly the same in their fundamental nature is preferable to

one that must reach for different motivations and mechanisms from case to case.

2. So far I have omitted from the discussion another alternative to the standard theory, Vennemann's NATURAL GENERATIVE GRAMMAR (Vennemann, 1972, MS). His approach to phonology represents a rather more radical departure from the usual theory in a number of ways. One of its principles is that no extrinsic ordering is allowed; to this is coupled an extremely concrete version of the abstractness condition. Other conditions (see Vennemann, MS) are proposed to enforce phonological uniqueness. The result is a highly concrete phonology in which (it comes as no surprise) rules need not be extrinsically ordered.

To criticize Vennemann's theory would involve, among other things, debating whether his adherence to so concrete a phonology does not force him to give up valid generalizations at some point. For example, in his recent discussion of the German velar nasal, Vennemann (MS,14) cheerfully abandons the considerable advantages of deriving it from underlying /ng/ – advantages he had sedulously cultivated in his earlier article on the subject (Vennemann, 1970) – and says simply that he prefers an analysis which recognizes a phoneme /ŋ/ and "describes the morpheme structure and morphophonemics of German with SOMEWHAT MORE COMPLICATED RULES" (emphasis mine, RDK). The appropriate questions here are, how much has been given up and at what cost? But I stated at the outset of this paper that I would avoid discussing the abstractness problem, and it is exactly this issue that must be confronted first in analyzing Vennemann's theory (which, I would like to add, I admire in certain respects).

But there is another reason why I have chosen not to discuss Vennemann's theory in this paper. I have mainly been concerned with arguing that available theories of intrinsic ordering do not account very well for certain facts about change. I would claim that the same is true of Vennemann's theory, but he has been careful to take into account the implications of his theory for phonological change. In this domain, my argument with him would revolve around questions of diachronic theory proper, such as: is 'rule inversion' in fact a *bona fide* phonological change?; and, does this theory give a unified, coherent account of the changes – all of them – that in the standard theory are described by reordering? My discussion of these issues will have to await another opportunity.

3. Ultimately, the substance of the dispute discussed in this paper can be concentrated in a simple question: do dialects ever have the same underlying forms and the same rules but in different orders? Theories of intrinsic ordering predict that they cannot, and theories with extrinsic ordering predict that they can. Obviously, it becomes very important to determine whether or not dia-

lects can differ in just this way. The advent of generative phonology brought forth a spate of examples purporting to show that dialects quite typically differed in rule order and nothing else. It must be recognized, of course, that many such examples are simply wrong, vitiated by superficiality.[21] Even so, the number of putative instances of dialects differing solely in rule ordering is very large and cannot be simply wished away. Bailey (1974) has an almost overwhelming number of such examples; and there remains, of course, the well known case from German dialects involving Final Devoicing and *g*-Deletion (see Kiparsky, 1971a, 599-600).[22] Some examples seem to be due to nothing more arcane than geographical rule spread (see, for example, Keyser, 1963), and not Opacity or the Paradigm Condition. Until all these cases have been convicingly handled in a theory of intrinsic ordering, it seems to me extremely premature to abandon the old theory, despite all its shortcomings.

4. In the preceding sections I was primarily concerned with pointing out ways in which theories of intrinsic ordering have failed to provide an adequate basis for explaining phonological changes.[23] This is a serious defect in any synchronic theory, since linguistic change offers one of the very few reliable EXTERNAL means of assessing the correctness of a phonological theory: there really isn't much difference between phonology and historical phonology. I have also called attention to other weaknesses in these theories. Having made an argument for extrinsic ordering – actually, to put it more precisely, for a theory that provides a measure of the naturalness of rules sequenced extrinsically – I would like to relax my stance somewhat and point out areas of phonology where I believe intrinsic ordering is correct.

I think it is now relatively clear that sequencing priorities between certain kinds of phonological processes do in fact follow from universal principles and need not be stated as a control which must be learned by each speaker of the language. The clearest cases are those involving 'low level' phonetic processes, in particular, priority relationships between language-specific phonological processes and articulatory constraints imposed by universal phonetic theory. A typical example (taken from Harms, 1972) is the relationship between Schwa-Epenthesis and Progressive Devoicing in English, whose interaction accounts for the affixal alternations observed in (11):

(11)	[sip]	'sip'	[sips]	'sips'	[sipt]	'sipped'
	[kis]	'kiss'	[kisəz]	'kisses'	[kist]	'kissed'
	[fit]	'fit'	[fits]	'fits'	[fitəd]	'fitted'
	[biyn]	'bean'	[biynz]	'beans'	[biynd]	'beaned'

Assuming that the affix series [s ~ z ~ əz] and [t ~ d ~ əd] are derived from underlying /z/ and /d/,[24] a rule of Schwa-Epenthesis is needed (I forego formulating it here), as well as Progressive Devoicing, (12):

(12) [+ obst

That Schwa-E
grammar that
seems beyond
nounceable o
can be resolve
obstruents (e
mon in other
separate the t

It is eq
English gram
constraint th
lable, voicing
(That is to sa
point which i
syllable.) Thi
in which taut
débâcle [deb

In shor
but is a requi
Epenthesis is
articulatory c
Norman, 197
one of the m
NATURE (rat
and disallow
to the requir
in the langua
universal phc
sically order

But th
have been m
with two hig
sarily very re
while to disc
the area of c
Progressive I
cation is, in
and I think t
ciple which
troversial se
olated; it mu

(12) [+ obstruent] → [- voice] / [- voice] —— #

That Schwa-Epenthesis is a rule of English phonology, i.e., a rule of English grammar that a speaker must 'learn' to control in his acquisition of the language, seems beyond doubt. Its function, of course, is to break up certain unpronounceable obstruent clusters; but the dilemma of unpronounceable clusters can be resolved in other ways, among them, deletion of one of the clustered obstruents (e.g., hypothetical /kis + z/ → [kiz]), a process that is fairly common in other languages. But this is not done in English; the speaker learns to separate the two obstruents by a vowel.

It is equally beyond doubt that Progressive Devoicing is not a rule of English grammar; rather, it is one particular exemplification of the universal constraint that "once voicing ceases following the nucleus (vowel) of any syllable, voicing can no longer resume in that same syllable" (R. Harms, 1972, 10). (That is to say, if within a syllable, voicing begins and then is stopped at that point which is after the nucleus of the syllable, it cannot resume in that same syllable.) This universal constraint is also manifested in a language like French, in which tautosyllabic resonants are devoiced after voiceless obstruents: *débâcle* [debakL], *théâtre* [teatR].

In short, Progressive Devoicing is not a rule of English grammar at all, but is a requirement of universal phonetics; and its position relative to Schwa-Epenthesis is a consequence of the fact that any rule expressing a universal articulatory constraint must follow the rules for specific languages (see also Norman, 1973). I would claim that these two rules are intrinsically ordered in one of the most fundamentally correct senses of the term, namely that the NATURE (rather than the formulation) of the two processes requires one order and disallows the other. The effects of a language-specific rule are subordinate to the requirements of universal phonetic constraints: no matter what happens in the language, no matter what kinds of sequences arise during derivation, the universal phonetic constraints must be observed. In this sense, they are intrinsically ordered after any of the language-specific rules.

But this kind of intrinsic ordering does not generalize to the cases that have been most often discussed in the literature, which typically have to do with two higher level, language-specific phonological rules. Though not necessarily very relevant to the controversy over extrinsic ordering, I think it is worthwhile to discuss this type of example because it sets a kind of lower limit to the area of common agreement: certainly processes like Schwa-Epenthesis and Progressive Devoicing represent ordered processes whose sequence of application is, in one sense of the word, intrinsic. But what else does? Moreover, and I think this is most important in view of my comments in note 3, the principle which serves to sequence the two processes is universal in the least controversial sense of the word: an articulatory constraint simply CANNOT be violated; it must have the last say in any derivation. What is crucial is the recog-

nition that two fundamentally different kinds of processes are involved. It would not be persuasive, therefore, to appeal to the notion, for example, that feeding order must be maximized in order to sequence the two rules here (though this would do so correctly since Schwa-Epenthesis creates the environment in which Progressive Devoicing becomes applicable). The argument that feeding order is the sequencing principle here would be valid only if, in principle, some language could apply the rules in the opposite, nonfeeding order. But this is precisely what cannot be done; so I see no merit in trying to account for the sequencing observed here by calling on a putative universal principle that rules must apply in a feeding order. Notice also that another putative universal principle, that opacity must be minimized, could be invoked to sequence the rules correctly: the order Progressive Devoicing followed by Schwa-Epenthesis would produce opaque outputs like *[kisz], making Progressive Devoicing opaque. Again, since there is no possibility that a language could ever apply the rules in this order, I would say that nothing has been explained by bringing in the notion of Opacity.

While the sequencing of Schwa-Epenthesis and Progressive Devoicing follows from a principle very different from either of those just mentioned, i.e., feeding order must be maximized and opacity must be minimized, it does seem to me that such principles have explanatory value in the sequencing of low level phonetic rules.[25] One example will suffice to illustrate the point. Zwicky (1972a, 279) discusses variant pronunciations of words like *winter* in his speech. In careful speech the pronunciation is [wɪ̃ntr]. The [t] in [wɪ̃ntr] is not affected by Flapping, the well known rule of American English that converts [t] to [D] following a stressed vowel and before a sonorant, since [t] follows [n] and not a vowel. In fast speech, however, the word is pronounced [wɪ̃Dr]. This could be accounted for by ordering the rule of Nasal Dropping (which comes after the rule nasalizing vowels in nasal environments) before Flapping:[26]

(13)	Underlying:	/wintr/
	Nasalization:	wɪ̃ntr
	Nasal Dropping:	wɪ̃tr
	Flapping:	wɪ̃Dr
	Phonetic:	[wɪ̃Dr]

However, the correct derivation could also be enforced by the requirement that rules apply in the sequence that maximizes feeding order (and minimizes bleeding order).[27] It has often been observed that 'lowest level' phonetic rules tend to be sequenced in just this way. If there are no counterexamples to the claim, and if it is possible and desirable to specify exactly what is meant by my vague term 'lowest level phonetic rule', then there would of course, be a theoretical advantage in sequencing rules in this part of the grammar intrinsi-

cally, say by requiring that feeding order be maximized and bleeding order minimized.

A further example, in part quite similar to the one above, of this kind of sequencing concerns variant pronunciations of a phrase like *burnt him*. For me, the pronunciation in careful speech is [bɚ̃nt#him]; normally, in casual speech, it is [bɚ̃ Dim]. Without going into details, the derivation is about as follows:

(14)

Underlying:	/bʌrnt#him/
r-Coloring:	bɚ rnt#him
r-Dropping:	bɚ nt#him
Nasalization:	bɚ̃ nt#him
Nasal Dropping:	bɚ̃ t#him
#-Deletion:	bɚ̃ thim
h-Dropping:	bɚ̃ tim
Flapping:	bɚ̃ Dim
Phonetic:	[bɚ̃ Dim]

But even if there seem to be plausible arguments that rule sequences are restricted by universal principles in the 'lower' parts of the grammar, it does not follow necessarily that the principles, which are valid there generalize to higher levels of the phonology. As I have tried to show in earlier sections of this paper, there are serious difficulties in such an extrapolation.

As I said in the introduction to this paper, I can't escape the suspicion that it is premature and somewhat unrealistic to focus on rule ordering without having achieved a better understanding of other problem areas in phonology: abstractness; the differences between derivational, morphological, morphophonemic, phonological, and phonetic rules; and so on. It is equally important to get a better understanding of the diachronic events that participate in the formation of a synchronic grammar. In rule ordering, as in every other critical area of language study, one is again struck by the immense complexity of the ways that languages are organized and by the intricate connections between a grammar and its history. We should not forget that if a synchronic theory makes a hodgepodge of linguistic change, then that theory is not much use. The performance of a synchronic theory in the face of diachronic change remains one of the few reliable tests of synchronic theories that we possess.

NOTES

1 The 'recent efforts' alluded to here have become so extensive that it would be difficult to list them exhaustively. The following are representative of the major trends: S. Anderson (1969); Kenstowicz and Kisseberth (1970); Newton (1971); Koutsoudas

Sanders, and Noll (1971); Kisseberth (1973a); Vennemann (1972, MS); Norman (1973); Sommerstein (MS). I do not include here papers dealing primarily with syntax since, given the substantial differences of so many kinds between the two linguistic domains, there is no *a priori* reason why ordering relationships in syntax should parallel those in phonology.

2 What I mainly have in mind here is the inclination in the literature to deal solely with 'two rule' examples, totally ignoring how the analysis for such examples bears on the rest of the grammar. If anything has become clear since the advent of generative phonology, it is that even innocent appearing little clumps of data in a language can turn out, on deeper investigation, to be incredibly complex and bound up in exceedingly intricate ways with other clumps of data. See Brame (1971) on Arabic stress for a recent, typical example. Yet almost none of the recent work on intrinsic ordering has seriously tried to assess what additional complications in the theory are required to obtain the correct derivations when substantially MORE than two rules interact (the treatment in Koutsoudas, Sanders, and Noll (1971) of a set of more than two rules from Mohawk is a rare exception to the trend). In an extended set of rules containing nonoptimal orders, there will be difficulties for all of the theories excluding extrinsic ordering that I have seen.

A representative example of the sort of difficulties I have in mind is the derivation of *regicide* in Chomsky and Halle (1968, 220), which requires the sequence Velar Softening or Trisyllabic Laxing – Diphthongization – Vowel Shift. None of the usual principles in theories of intrinsic ordering (Proper Inclusion Precedence, Minimization of Opacity, Random Sequential Application) produces the correct results in any obvious way. For an example presenting even greater complications, see Harris (1972). Furthermore, how would the intricate metrical relationships that Kiparsky (1972b) relates to intermediate levels in derivations be handled in a theory without extrinsic ordering?

3 The problem here is one of the central dilemmas in phonological theory: how do we distinguish principles which WORK from principles that are linguistically CORRECT? Sommerstein (MS, 11-12) proposes a theory of intrinsic ordering whose principles are not only extremely convoluted and arcane but require, given n rules, approximately n^2 acts of screening to arrive at the correct derivation for a particular input. Is it realistic to believe that a natural language is organized along these lines? Or consider Kisseberth's (1973a) proposal that rules be universally sequenced (when possible) by the principle that opacity be minimized. Since, in principle, opacity is a property derived not from the interaction of just two rules but of all the rules that can participate in a derivation, in general it would be necessary to scan as many as ALL the possible sequences of the n rules in a grammar to determine the correct derivation. The magnitude of the number of derivations to be checked is $n!$, which of course is extremely large: for ten rules it is 3,628,800. It seems unlikely to me that a genuine universal principle of language organization in human beings could require anywhere near this amount of processing. Compared with a figure like this one, the 'cost' of extrinsic ordering, i.e., the cost of memorizing the order of n rules, doesn't seem as exorbitant as many linguists have tacitly assumed it to be.

4 See Kiparsky (1971a, 621-3) on Opacity. On the Paradigm Condition, see Harris (1970b), Kiparsky (1972a, 206-13) and King (1973, 553-67). On possible conflict between the two principles see King (1972b); against the latter see Bailey (MS).

5 It should be pointed out that there is a concept which is equivalent to what I call here the 'standard theory' that could be used to induce partial intrinsic ordering on the grammar. Suppose it were possible to determine exactly what sequencing principles belong in the evaluation metric (the Paradigm Condition and Opacity have been mentioned, and no doubt there are others). Suppose further that conflicting evaluations made by these principles have been resolved so as to permit statements of the sort 'given a certain input, the rule sequence A-B is more natural than the sequence B-A'. We could then, following the procedure in S. Anderson (1969), leave unmarked those sequences which naturally apply and mark only those not predicted by the general principles. The 'cost' of the grammar would then be measured solely by the pairs that apply in marked orders.

The difference, as far as I can tell, between this kind of theory of partial intrinsic ordering and the theory outlined in the body of this paper, which has extrinsic ordering and an associated evaluation metric, is purely conceptual. The same predictions about phono-

logical change, reorderings in casual speech, ordering mistakes among children, and so on are made by both theories, so that there would be no reason to prefer one over the other on empirical grounds. However, Phelps and Brame (1973) have pointed out one kind of difficulty in Anderson's theory of local ordering, namely, that it requires, in certain cases at least, an implausible amount of processing to apply the rules correctly (again on the order of $n!$). Since the theory of partial intrinsic ordering I have outlined here is basically identical to Anderson's local ordering (only the sequencing principles are different), it is probable that any defect in Anderson's theory would afflict my proposal as well. I therefore prefer to stick with the standard theory and extrinsic ordering.

6 I find an inconsistency in the KSN article concerning the terms 'simultaneous' and 'unordered', which I would prefer to call 'random sequential': a rule applies any time its structural analysis is met (see Postal, 1968, 141-3). Such applications are not equivalent, nor does one (simultaneous application) follow from the other (random sequential application) in the way that KSN suggest. This can be demonstrated in various ways, but it is sufficient to point out that Vennemann (MS) argues for a theory that excludes simultaneous ordering (as *ad hoc*) but includes, as its sole ordering convention, random sequential application. Part of the trouble here has been cleared up by the Koutsoudas contribution to this volume.

7 J.Harris (1972, 22) discusses a case from Portuguese in which a stress assignment rule precedes a rule neutralizing certain pairs of unstressed vowels It would seem that Stress Assignment would be ordered before Neutralization INTRINSICALLY in precisely the sense Chomsky (1965,223) originally proposed: solely, by virtue of its formulation, a rule affecting unstressed vowels cannot apply until stress has been assigned. The height distinctions between vowels that are merged by Neutralization are needed to trigger other phonological rules. As Harris points out, given the natural assumption that vowels in a largely stress-predictable language like Portuguese are marked [- stress] in the underlying forms, if rules were allowed to apply any time their structural analysis is met, the rule of Neutralization would incorrectly erase vowel distinctions in underlying representations. In general, given any moderately abstract phonology, it will not be difficult to falsify the flat claim that rules must apply random sequentially.

8 Attention should be directed to the very large number of putative universal principles that have been proposed in theories of no or very little extrinsic ordering: Maximal Utilization (S.Anderson, 1969); Counterbleeding Precedence and Proper Inclusion Precedence (KSN); precedence of obligatory rules over optional rules (Norman, 1973, following Ringen, 1972a) – against this principle see Malone, (1973); Minimization of Opacity (Kisseberth, 1973a); and the principle of Minimum Application (Iverson, in this volume), also discussed in Norman, (1973). With as many principles as this at our disposal, it is fair to ask, what two rules could NOT be sequenced intrinsically? Could the theory ever fail?

9 A common argument given in favor of theories of intrinsic ordering is that they are methodologically preferable because they make more restrictive claims, they force us to look for sequencing principles where we might otherwise not, and so on. A theory of extrinsic ordering, so the argument goes, makes no claim other than that certain sequences of rules work in a language, and hence is uninteresting. This objection would hold only if the theory said: rules are extrinsically ordered, period. But I have argued that there is more to the theory than this, notably that there is an evaluation metric which assigns naturalness measures to rule sequences. The principles that constitute the metric are eminently empirical, of course: the one saying that rules which interact in a paradigm should apply in the order that minimizes allomorphic variation would be immediately falsified if one could produce a reordering that INCREASED the variation; the claim that bleeding orders are always eschewed has been empirically disproved (Kenstowicz and Kisseberth, 1973a; Kiparsky, 1971a, 599-600; and King 1973).

So while I basically am inclined to agree in principle with the methodological argument favoring intrinsic ordering, I think it is vastly overrated (along with a lot of other 'philosophy of science' arguments in linguistics). If there are universal principles that sequence rules, they will come to light either in the search for universal principles that sequence rules intrinsically or in the search for the principles in the evaluation metric of the standard theory that measure rule order naturalness. I do not believe that Kiparsky stum-

bled on feeding and bleeding orders or opacity simply because he had adopted the premise that rules are intrinsically ordered in the grammar.

10 I think it is established that Modern Standard German is better described by the derivations in (3) than those in (2). There is very little evidence in Modern German for a synchronic rule lengthening vowels in open syllables or for any other rule that would derive the long vowel in *wohnen* and the like.

11 See Stampe (1969, 453) and Kiparsky (1971a, 627-30).

12 In fact, there is motivation for assuming a shortening rule of some form, at least historically. It is the well known 'mirror image' counterpart of Lengthening in open syllables, and it was operative in the history of German and West Germanic generally: in German, earlier *dāhte* 'I thought' became *dahte* by this rule.

If the 'extra rule' reanalysis is at least not grossly implausible here, I would claim that the parallel reanalyses that KSN give for Kiparsky's Swiss German examples are. The extra rule that KSN argue for in these cases serves 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" (KSN, 1971, fn. 10). As Sommerstein (MS, 6) observes, it is not generally regarded as a sufficient condition for a phonological rule that it express a "true generalization about the language", e.g., that certain Swiss German dialects do not have low front rounded vowels. There have to be alternations. Just because English has no front rounded vowels, is it a phonological rule of English that front rounded vowels become unrounded? Certainly not. The difficulty here is compounded by the fact that the historical use to which KSN put their well-formedness rules – that their loss accounts for the difference between conservative and innovating dialects – is highly suspect since it requires the loss of a rule that is TRANSPARENT in the most obviously basic meaning of the term. Well-formedness rules, in this sense, are NEVER violated on the surface.

13 In the discussion that followed the oral presentation of this paper, it was pointed out that the rule shortening vowels in closed syllables was opaque because of forms like *wīn* 'wine', *nōt* 'stringency', and a large number of similar words. This suggests that the Shortening rule must be formulated differently: vowels become short before consonant clusters, e.g., *dāhte* > *dahte*. In this formulation, the rule is transparent, and the arguments against KSN given in the text still hold.

14 The paradigm is given in phonetic transcription. [j] is orthographic *g* (or *ġ* in some normalizations), and the voiced velar fricative [ɣ] is spelled *g* intervocalically.

15 Cooley (1972, 39-65) gives detailed motivation for the underlying forms and rules that I deal with, as well as evidence that the paradigm (6) existed at an earlier stage of Old English. Let me emphasize, though, that (6) is a reconstructed paradigm, albeit one that seems well secured both internally and comparatively.

16 This example from Old English, simple as it is, offers striking support for the reason why generative phonology often allows one to make more specific predictions about analogical developments than the simple, unqualified proposition in traditional historical linguistics that analogy causes changes in surface forms. The paradigm (6) offers any number of possible and equally likely analogical changes that would reduce the amount of paradigm variation: the [x] of [dæx] could be generalized throughout the singular or, for that matter, the whole paradigm; the [ɣ] of the plural could be extended to the singular; the [x] of [dæx] could be replaced by the [j] from the other singular forms; and there are still other less probable regularizations that could have occurred. Indeed, assuming that the nominative singular in general is the unmarked morphological category in a nominal declension, one would expect it to be the form from which analogical changes emanate, so that the morphological markedness hierarchy would lead us to expect the [x] of [dæx] to be generalized, which is counterfactual. (The idea implicitly appealed to in the preceding sentence is found in a good deal of Kuryłowicz's work.) Of all the plausible predictions about analogical change in the paradigm (6), only ONE actually takes place: precisely the one which follows automatically from a reordering of the two rules. I have nothing against the traditional ideas about analogy; I merely observe that they become immeasurably strengthened and more precise when integrated with grammars of ordered rules.

Reorderings of this type, which proceed from one bleeding order to another, quite frequently have this property, i.e., given only the surface paradigms and not the rules that relate the surface forms of the paradigm to the underlying forms, they precisely pinpoint the one analogical change which actually took place even though a number of others seem equally plausible. In the German case discussed earlier, which was from a bleeding order to a nonbleeding order and not mutual bleeding, the ordinary idea of analogy would predict leveling. This did in fact happen, although traditional analogy, without additional constraints, leaves us in the dark concerning which type of vowel levels the other.

17 Sommerstein (MS, 26, fn. 3) proposes to redefine Proper Inclusion as follows: "a structural description X is defined to properly include a structural description Y just in case every logically possible matrix satisfying X also satisfies Y but not conversely". Since a matrix, e.g., [burx] *burg* ~ *burh* 'city', can satisfy (7") but not (8"), and a matrix, e.g., [dæges] 'of day', can satisfy (8") but not (7"), I am unable to see that Sommerstein's redefinition of Proper Inclusion will serve to sequence the rules in either order.

18 I exclude the Slavic example because it is not clear whether reordering is involved or not. The evidence seems to support another kind of explanation; see S. Anderson (1969) and Cohen (1969).

19 Of course, rule loss is another way of carrying out morphophonemic analogy. In these cases, however, only reordering is (initially) involved.

20 Vennemann (MS, 5) has criticized the KSN discussion of the Alsatian/Low German example on different grounds.

21 See, for example, the criticism of Saporta (1965) in J. Harris (1969a, 550-1).

22 The question in this example is what the underlying forms are in those dialects where *Ding/Dinge* are realized as [diŋ/diŋə] . As Vennemann (1970) has shown so well, there are solid, independent reasons for assuming [ŋ] to be derived from /ng/; and, as mentioned earlier, I am troubled by the cavalier way in which Vennemann (MS) dismisses his earlier work. It would be most interesting to determine whether the evidence of speech errors supports underlying /ng/ in German as Fromkin (1971, 34-5) has shown such evidence does for English.

23 Norman (1973) offers an interesting discussion of historical change in a theory of intrinsic ordering. I wish to comment here only on her treatment of the two dialects of Canadian English. She claims there is no evidence that Dialect A (the [rʌyDr] dialect). is chronologically prior to Dialect B (the [rayDr] dialect). First there is in fact evidence supporting a chronological relationship of just the sort she denies (see King, 1972b). Second, as pointed out to me by Robert K. Szabo, her assumption that speakers of Canadian English incorporated two different lowering rules in their dialects predicts that the two dialects, A and B, would be randomly distributed throughout the English-speaking area of Canada since there is no surface data suggesting one rule instead of the other. However, the situation seems to be one of nonrandom geographical distribution of the two dialects. In the absence of better dialect-geographical and historical data on these problems, the issues cannot be resolved at present.

24 There is a well known dispute about the underlying sources of the endings here: it has been argued, for example, that they should be /Vz/ and /Vd/. Nothing in my discussion hinges on this issue since, whether Schwa-Epenthesis or some kind of vowel deletion rule is preferable, either will interact with Progressive Devoicing in the way described.

25 The kinds of rules meant here have been discussed by Stampe (1969, 1972) and Bailey (1974).

26 The order of Nasal Dropping, which is an optional fast-speech rule, and Flapping, which is obligatory, contradicts the claim made by Norman (1973) that optional rules never precede obligatory rules in phonology (see also Malone, 1973).

27 Notice that Nasal Dropping bleeds Nasalization and that Nasal Dropping feeds Flapping. It should be observed that the derivation in (13) cannot be obtained by the principle that opacity must be minimized, since Nasalization has been made opaque by Nasal Dropping, and the correct derivation does not follow from the claim that rules apply random sequentially. This is one more example of the difficulty, discussed in note 2, that theories of no extrinsic ordering get into when more than two rules that are needed in a

derivation must be considered. The example which follows (*burnt him*) poses an even greater problem for such theories.

DISCUSSION

GREGORY IVERSON (University of Minnesota): I haven't seen Cooley's dissertation that you cite in your paper, but I'm wondering just how sure we can be that the earlier stage of Old English with *dax* with a voiceless fricative at the end really existed.

ROBERT KING (University of Texas): This is the way a traditional comparative linguist would account for the transition from, say, Ingvaeonic to Old English. Old Saxon, for instance, and Old Low Franconian have this paradigm. The evidence is basically comparative, and to some extent internal.

GREGORY IVERSON: Okay. Another thing: can you show me any morpheme in Old Saxon, or Old English, or anything else that demonstrates the reordering of those rules?

ROBERT KING: You get it through the third person plural.

GREGORY IVERSON: Yes, just with the morpheme /lanþ/. What I'm asking is, you might want to use a reordering argument to account for that particular inflection – the third person inflection – but I'd like to know if there are any OTHER morphemes in the language that would exhibit that particular ordering of the rules – like, say, the word *lamb*, for example. That's underlying /lamβ/ [bilabial fricative], and the ordering of the rules you posit for Ingvaeonic would be such that you first devoice the fricative to a voiceless bilabial fricative and then we delete the nasal; so today we should be saying [læf] instead of [læm]. I can't think of any other morphemes for which that reordering obtains.

ROBERT KING: Okay, I'll concede that, because it's a good point. I hadn't thought about *lamb* or similar words.

GREGORY IVERSON: You can't do a reordering argument based on one morpheme.

ROBERT KING: But what about the German example?

GREGORY IVERSON: That's a better example. I'll concede that one to you . . .

(laughter) . . .

ANDREAS KOUT
German exampl
the innovative d
the rule become
environment otl
for the German
has long vowels
case, then your
you cannot pre
cause if it's opa
be lost right th
reordering is nc

ROBERT KING
older stage or

ANDREAS KOU
the direction c
come short ins
the rule was tr
fore, since you
So if the rule
both sides: yo
any reorderin

ROBERT KINC
follows from
ly opaque rul

ANDREAS KO
for otherwise
say that the
become shor
rule can be l

ROBERT KIN
nition. Right

* The length
at the request
editor.]

(laughter) . . .

ANDREAS KOUTSOUDAS (Indiana University): But I don't concede this German example to him. First you say, in response to this example, that in the innovative dialect, the Lengthening rule is opaque by condition (1). I think the rule becomes opaque by condition (2), since you have lengthening in an environment other that open syllables. Secondly, in the sources that you cite for the German example, it seems also to be the case that Middle High German has long vowels in monosyllabic words, like *den*, *der*, and *dem*. If this is the case, then your Lengthening rule is opaque in the older stage of the language, you cannot predict both the direction of the change and the reordering – because if it's opaque in the older stage, then there's no reason why it couldn't be lost right then. We can then explain the German example by rule loss, and reordering is not involved.

ROBERT KING : Okay, to me it makes no difference whether it's opaque in the older stage or not.*

ANDREAS KOUTSOUDAS: It does make a big difference if you want to predict the direction of leveling. You keep asking, why is it that all vowels don't become short instead of all vowels becoming long? The reason is that in Stage 1, the rule was transparent, and it became opaque by reordering in Stage 2; therefore, since you make opacity a necessary condition for loss, now it can be lost. So if the rule was opaque in Stage 1, it seems to me, you lose the argument on both sides: you cannot explain either the direction of change or claim there is any reordering.

ROBERT KING: Again, reordering is motivated by the Paradigm Condition; it follows from that. You've got an irregular paradigm. To me whether an initially opaque rule is involved or not is a matter of indifference.

ANDREAS KOUTSOUDAS: To you it's very IMPORTANT that it is transparent, for otherwise you cannot explain the direction of leveling. See, you need to say that the reason the rule did not get lost in Stage 1, and therefore all vowels become short instead of long, is that it was transparent, and no transparent rule can be lost. Isn't that correct?

ROBERT KING: Well, remember, it's opaque by the second branch of the definition. Right? And in general, that's not enough to make a rule get lost. Even

* The length in *den*, *der* and *dem* (if they had length) is due to a separate rule. [Added at the request of Robert King during the preparation of this volume for publication: editor.]

opacity by branch (1) isn't enough to MAKE you lose the rule; it's a necessary condition, not a sufficient one.

ANDREAS KOUTSOUDAS: But then, how do you predict the direction of change? How do you predict the fact that all vowels are long instead of that all vowels are short?

ROBERT KING: Because the Lengthening rule can't be lost, since it's not opaque by branch (1). It's opaque only by branch (2). And that's the one that's crucial, in my opinion, to rule loss.

ANDREAS KOUTSOUDAS: So the branches of the opacity definition matter.

ROBERT KING: Look, Final Devoicing in Modern German is opaque by branch (2) of the definition; and that rule is, in the intuitive sense of transparency, totally transparent. It's transparent by branch (1) of the definition and opaque by branch (2); but there's absolutely no tendency whatsoever to lose that rule in German. Now in Yiddish, that rule was lost, but only after you had lost final schwa, which then brought voiced obstruents into word-final position – which makes the rule opaque by branch (1). That's when you lose it. I don't think that branch (2) has anything to do with rule loss.

ANDREAS KOUTSOUDAS: But neither in the Middle High German nor in the innovative dialect is the rule opaque by condition (1); therefore, you should not be able to lose the Lengthening rule in Modern Standard German.

Okay, I'll move to a second question. This has to do with the directionality of your prediction. There are three possibilities: (1) if paradigm regularity obtains, then reordering has occurred; or (2) if reordering occurs, then paradigm regularity results; or (3) bidirectional: if paradigm regularity, then reordering – and if reordering, then paradigm regularity.

ROBERT KING: Well, I don't know how to state what I feel in terms of these three categories. The prediction I'm making is that if two rules interact in a paradigm and make it irregular, they will reorder. Another of the predictions I'm making is that reordering is a likely change. But the paradigm could also be regularized by rule loss, if the Opacity Condition is met, or by rule Generalization. I'm not saying reordering is more likely than any other kind of change.

ANDREAS KOUTSOUDAS: So you're saying the second? Can I say that you're saying if [you have] reordering, then [there is] paradigm regularity?

ROBERT KING: Well, it conflicts with opacity . . .

ANDREAS KOUTSOUDAS: Suppose you have two grammars – one is an older stage of a language and the other is a later stage of the same language – and we see that there has been a reordering, and we also see that the paradigm has been regularized. Now, which predicts which?

ROBERT KING: Obviously, you look at the first stage, and you say, "Look, I've got two rules that interact in the paradigm and make it irregular." My prediction is that they will reorder.

ANDREAS KOUTSOUDAS: Okay, well, now . . . if there are other things that can happen too – rule loss, for example – then I don't see how you can say that you have just one mechanism and just one principle to predict what you want to predict.

ROBERT KING: No, what I meant was, in this category of cases, a very small category of cases, where two rules are involved with a paradigm and in one order they make it irregular and in the other order they make it regular, the favored ordering is the one that makes it regular. But it's not a unique prediction; a rule can also become lost or generalized.

ANDREAS KOUTSOUDAS: But one of your criticisms against me was that I don't have one mechanism that explains one event.

ROBERT KING: Well, this is the dispute here, I think: you take one of the many cases of reordering in the literature – mine, Kiparsky's, and so on – like the Alsatian Low German thing, for instance. Now for me, it's very simple: you reorder, and the underlying representation remains the same, and later maybe the system gets reanalyzed. But what you do in KSN is to have a number of different explanations which depend on the accidental facts at hand for each case. In the two Swiss German cases, for example, you have the business of no front rounded vowels. In the Alsatian case, you have an intermediate stage where fricatives become stops intervocalically – which is counterfactual by the way – and so on. You have always a different explanation for dialects that seem to differ in minimal ways.

ANDREAS KOUTSOUDAS: You're saying that all these things are cases of paradigm regularization, but I have more than one way to explain it. Right? Firstly, I don't see any *a priori* reason why all cases of paradigm regularity should be explained in the same way. Secondly, I can't see, unless you can give me the direction of your prediction, how you have one single way of explaining these cases either. You would have to claim that every time you get paradigm regularity, reordering has occurred; and if I understand you correctly, that isn't what you're saying at all.

ROBERT KING: Okay, let me ask you a question. Why in Swiss German – in those two Swiss German cases – do you lose that well-formedness rule? Why don't you lose another rule?

ANDREAS KOUTSOUDAS: Well, now, that's a different question. Now what you're asking me to predict is the direction of change. That I cannot do. What I'm pointing out with your High German example is that you cannot do it either.

GERALD SANDERS: (University of Minnesota): First of all, I'd like to say something favorable. I think that your paper, of all the things I've seen, comes closest to stating the issues clearly between what you call theories of intrinsic and theories of extrinsic order. To talk about the question of historical argumentation in linguistics is really the subject of another conference in itself. But it's certainly true that historical evidence is relevant, crucial, and important for evaluating claims about the nature of language – the theory of grammar. However, this is possible only if we have empirical hypotheses – if we have hypotheses that we can subject to empirical tests. I would question whether this is the case for any known account of language change so far, including, as I understand it from what you've said today, your own, although there are some things in your paper that I think could be interpreted as proposals of explanatory laws. And I'm certain that that's what Andreas [Koutsoudas] was trying to determine – that is, whether there is a theory there that has some empirical hypotheses – some law-like statements that could be subject to falsification.

Now your second point concerns the limited nature of the data and the rule sets involved. I agree with you entirely on this; but one possible response is, "You too!". All of the classical arguments for extrinsic ordering, going back to Chomsky's early papers and so on, are usually two-rule arguments. Now, some claims have been made for the necessity of ordering to the depth of 37 or 87 or whatever it is – in Chomsky's MA Thesis (1951), for example; but I haven't seen that thesis, and I haven't seen any other cases where really large sets of rules are dealt with on either side of this issue. So this certainly is a defect in the quality of the arguments on both sides. In KSN we do take seven rules from Mohawk; that's not very many either. But how many do you want?

The point that you mentioned concerning the distinction between principles that work and principles that are correct – again we're all on the same side on this. That is, we're concerned with determining things about what are the principles governing the organization of knowledge and so on – on the part of people who are speakers of human languages. And from this point of view, I think that the crucial tests of intrinsic vs. extrinsic ordering depends on tests of the viability of alternative rules. That is, where you have competing

accounts, one that makes use of extrinsic ordering constraints and one that doesn't, the crucial test is: is it possible to determine INDEPENDENTLY of the assumption or nonaussumption of extrinsic ordering that one of the sets of rules is correct? Now, from this point of view, we do have a sort of direct confrontation on the Schaffhausen and on the German example here, where you have one theory that does depend on the assumption of extrinsic ordering constraints and we have an alternative account that doesn't, but has some differences in the rules. The question, then, is, which rule set, independently of extrinsic ordering, is the most highly valued? But as I comment in my paper for this conference, that is often a very difficult thing to try to determine. We obviously have to go get a lot of facts. We have to look at EVERYTHING about the language to determine what are the right rules.

One final point. There's the question of evaluation, the use of the evaluation metric, and judgments of grammars and language change in terms of change from less highly valued to more highly valued grammars, as opposed to a system in which certain things are absolutely excluded. Now, I think the problem with this is related to the problem of explanation and the problem of testable hypotheses. If there is a principle that says there is a tendency for rules to reorder in such a way as to achieve paradigm regularization, there is clearly no way to falsify that principle stated that way. Tendency statements, unless accompanied by some statistical specification, cannot be falsified.

ROBERT KING: I think tendency statements are the strongest we can make, because languages do put up with irregular paradigms. So you're saying that such statements can never be falsified. But, the theory that is proposed in KSN (1971), in spite of the fact that it should be easy to falsify, in fact cannot be falsified either, because you can always reanalyze. If your rules don't have to express or account for alternations, there's no limit to what you can do in the way of reanalysis; and in effect, that removes KSN from the domain of a falsifiable theory.

GERALD SANDERS: Well, I think there's a distinction between what is falsifiable in principle but is in practice very difficult to falsify, and what is not falsifiable even IN PRINCIPLE. The principle of Obligatory Precedence, for example, IS falsifiable in principle. But a tendency statement is not falsifiable IN PRINCIPLE. That is, there's nothing you can do. We can't PILE UP cases, let's say, where you get paradigm regularization by principles other than Rule Reordering, or cases where you can account for language change by some principle other than Rule Reordering. That is, we can't pile these cases up if all that's being asserted is that there's a TENDENCY for languages to change.

ROBERT KING: I understand exactly what you mean, and I'm sympathetic with it; but the theory of historical linguistics is in part a series of statements

of, basically, tendencies. If you rule that out, then of course you rule out *a priori* any way of ever getting a theory of historical linguistics.

GERALD SANDERS: No, I agree with you entirely, and I'm in extreme sympathy with the task of the historian in linguistics, an in any other field. And I don't want to in any sense minimize the importance of making tendency statements. What I'm questioning, though, is the USE of historical evidence to support or to falsify a theory which is simply purporting to account for the nature of natural languages. I think in order for this connection to come through, there have to be some historical principles which are stated as absolutes, too.

ROBERT KING: I think we might have one in the prediction that a rule can't be lost unless it's opaque by branch (1) of the definition of opacity.

GERALD SANDERS: One final remark concerning the German case. Again, this is an instance of supporting a hypothesis by one example; that is, it works for this one verb: there is paradigm regularization by Rule Reordering, and loss of the opaque rule in the second stage. However, I'm sure that you would want to grant that there are many other hypotheses that would be consistent with this. These other hypotheses might be very general also. I'll mention one very simple one, which is probably false, but works fine for this case, and is just as neat in this case as the principle of Rule Reordering. For an example, let's take the principle that paradigms regularize to the infinitive . . .

ROBERT KING: I see where you're going, but that and other principles are going to fail rapidly, given all the cases of reordering that have been claimed. The only one that survives is Paradigm Regularity.

GERALD SANDERS: Are you saying that the principle of generalizing to the infinitive is false? Or are you saying it's not sufficient?

ROBERT KING: I'm saying it doesn't work sometimes.

PAUL KIPARSKY (Massachusetts Institute of Technology): Just on the question of empirical testability of theories of change: you can say, for a certain grammar, "these are the possible changes that can happen and these are the changes that can't happen"; and such statements can be falsified by citing cases of change which are not allowed for. What we're talking about in these cases is exact predictions about direction of one-way changes, and any reverse change would be a falsification of it. A theory with extrinsic ordering and some principles that determine unmarked ordering makes predictions about the direction of change. A theory which has no extrinsic ordering but reformulates the

rules does NOT make predictions of this sort about the direction of change. And I think King is correct in arguing that this counts against the theory without extrinsic ordering.