
"elsewhere" in phonology¹

Paul Kiparsky

Massachusetts Institute of Technology

The purpose of this paper is to investigate the conditions under which there is disjunctive ordering of rules in the phonological component of a grammar. I will present some evidence to show that disjunctive ordering is independent of the use of abbreviatory notations like parentheses or angled brackets, and I will defend a formal condition not involving any abbreviatory notations which seems to come close to being an empirically adequate necessary and sufficient condition on disjunctive ordering.

Let us begin by recapitulating the familiar example of English stress. For verbs like *édit*, *cóvet*, *devélop*, *rélísh*, in which the last syllable has a lax vowel followed by no more than one consonant, there must be a rule like (1) stressing the penult:

(1) $V \rightarrow [1stress] / _ _ C_0 \check{V} C_0 \#$

Elsewhere, the final syllable is stressed. An enumeration of these "elsewhere" environments would be complex and would miss the point that the totality of "elsewhere" environments is exactly the complement of the cases in which (1) applies. Instead of having various rules putting final stress on verbs with a tense last vowel (*refrâin*, *caréén*, *erâse*), verbs ending in two or more consonants (*elâpse*, *corréct*, *respéct*), and monosyllabic verbs (*rín*, *hít*, *sée*), we should be able to have simply a rule like (2), which stresses the final syllable *wherever* (1) is inapplicable:

(2) $V \rightarrow [1stress] / _ _ C_0 \#$

Rule (2) is quite general and subject only to the restriction that it does not apply when (1) applies (lest we get ^{2 1}**edit*, ^{2 1}**develop*, and so on). This relation between (1) and (2) is what Chomsky and Halle (1968) term "disjunctive ordering."

¹ This paper was supported in part by the National Institutes of Health (Grant No. 5 T01 HD 00111-08) and the National Institute of Mental Health (Grant No. 2-P01-MH 13390-06).

The relation of disjunctive ordering is obviously not one which needs an ad hoc specification for each pair of rules between which it holds. For example, it seems reasonable to assume that (1) and (2) must be disjunctively ordered because of their intrinsic nature, not because of some accidental facts about English stress. We may conclude this from the fact that rule pairs of the general type represented by (1), (2), which are very frequent in languages, have always been found to apply disjunctively. We would not expect to encounter a dialect of English with (1) and (2) not ordered

disjunctively, that is, a dialect that has *edit*, *develop*, but otherwise the stress system of regular English. For this reason it makes sense to look for a general principle that will assign disjunctive ordering to rules like (1) and (2).

The ideal result would be the discovery of some sufficient and necessary condition K which would have to hold for rules to be disjunctively ordered. Chomsky and Halle (1968, p. 77) propose that K is abbreviability by means of parentheses (and angled brackets, which can be regarded as discontinuous parentheses). The reason for the disjunctive ordering of (1) and (2), then, would be that they can be collapsed by parentheses, as in (3):

$$(3) V \rightarrow [1\text{stress}] / \text{---} C_0(\check{V}C_0^1)\#$$

However, there is a rather different way of looking at the disjunctive relation between these stress rules: we have here a *general* case (final stress) which is limited by a *special* case (penultimate stress in the environment $\text{---}C_0\check{V}C_0^1\#$). We might say, then, that it is this general/special relationship which constitutes the required condition K which induces disjunctive ordering. We can state the condition more precisely as in (4):

(4) Two adjacent rules of the form

$$\begin{array}{l} A \rightarrow B / P\text{---}Q \\ C \rightarrow D / R\text{---}S \end{array}$$

are disjunctively ordered if and only if:

- (a) the set of strings that fit PAQ is a subset of the set of strings that fit RCS , and
 (b) the structural changes of the two rules are either identical or incompatible²

For example, the two stress rules (1) and (2) are disjunctively ordered because any input subject to (1) is necessarily also subject to (2).

Let us term (4) the "Elsewhere Condition." A version of this condition is consistently employed in Pāṇini's *Aṣṭādhyāyī*.³ And S. Anderson (1969) has also suggested one.⁴ Furthermore, a similar principle is referred to in Koutsoudas, Sanders, and Noll (1971).

² This excludes cases where the two rules have nothing to do with each other, where the ordering must be conjunctive. For example, suppose that obstruents are voiced in the environment $V\text{---}V$ and palatalized in the environment $i\text{---}i$. Clearly, we expect $/iki/ \rightarrow [ig,i]$, with both rules applying. Here the subset relation holds, but the structural changes are neither identical nor incompatible and therefore one rule is still not a "special case" relative to the other.

³ The Elsewhere Condition is explicitly formulated in the *Mahābhāṣya* (for example, ad P. 6.1.89 *etyedhatyūṭhsu*; see Kielhorn (1960, vol. 3, p. 69) and constitutes *Paribhāṣā* LVII of the *Paribhāṣenduṣekhara*: *yena nāprāpte yo vidhir ārabhyate sa tasya bādḥako bhavati* ("a rule which is given [in reference to a particular case or particular cases to which, or to all of which] another [rule] cannot but apply [or in other words, which all already fall under some other rule], supersedes the latter"—Kielhorn's translation).

⁴ However, Pat Wolfe has pointed out to me that there are some problems with the Middle English example which Anderson cites.

In
conditi
are by
the one
means
be abb
orderc
On this
and Ha

A
the foll
where z

(5) *mu*
mū:
iṣpi
ʔu
ʔūs
iṣke
nik:
yēʔ

We hav

(6) *s -*

But the
'water'
Elsewh
a "spec

A c
not inv
process
Finland
deleted
tion is o
The scc
present
(One m
the poir

(7) *men*
men
men

The
in (8):

(8) (a)

(b) 1

In the English stress example involving rules (1) and (2), Chomsky and Halle's condition and the Elsewhere Condition have the same outcome. But the conditions are by no means equivalent generally. I will try to establish here that there are, on the one hand, cases of disjunctively ordered rules which cannot be abbreviated by means of parentheses or angled brackets and, on the other, cases of rules which must be abbreviated by means of parentheses or angled brackets but are not disjunctively ordered. In both classes of examples, the Elsewhere Condition gives the right result. On this basis I will argue that the Elsewhere Condition is to be preferred over Chomsky and Halle's principle for disjunctive ordering.

A simple type of situation requiring parentheses but not disjunctive ordering is the following. Karok (Bright (1957)) has a rule turning *s* into *š* after a front vowel or *y*, where a consonant may optionally intervene. Examples are given in (5):

- (5) *mu* 'his'
müspuka 'his money'
išpuka 'money'
ʔu 'he'
ʔüskak 'he jumps'
iškak 'jump'
nükšup 'I pointed'
yē pša 'good ones'

išš
 iš

We have, then, the rule in (6):

- (6) $s \rightarrow \text{š} / \left[\begin{array}{l} - \text{back} \\ - \text{consonantal} \end{array} \right] (\text{C}) _$

But the two subrules must apply conjunctively, as is clear from examples like *ʔiššaha* 'water' versus *vässih* 'back'. As opposed to Chomsky and Halle's convention, the Elsewhere Condition correctly predicts conjunctive ordering since one subrule is not a "special case" of the other in the sense of (4).

A common type of situation which calls for the Elsewhere Condition but does not involve parentheses is a disjunctive relation between assimilation and deletion processes. A typical case is the treatment of word-final *-k* in Finnish. In western Finland this segment is assimilated to an initial consonant in the following word and deleted before vowels and pauses (see Itkonen (1964)). In some dialects the assimilation is optional or inapplicable in certain environments (for example, *h*, *f*, or clusters). The scope of deletion is then correspondingly greater. Synchronically, *-k* must be present as a consonant in underlying representations (see, for example, Wiik (1969)). (One might argue that it has been reanalyzed as */h/*, but this is of no importance for the point at issue.) Some examples are given in (7):

- (7) *menek # pois* → *menep pois* 'go away'
menek # alas → *mene alas* 'go down'
menek # → *mene* 'go'

The rule for final */k/* (or */h/*, as the case may be) would now have to be stated as in (8):

- (8) (a) $k \rightarrow C_i / _ \# C_i$
 (b) $k \rightarrow \phi / _ \# \left(\begin{array}{l} V \\ \text{pause} \end{array} \right)$

generalization

In the dialects where Assimilation takes place in a more restricted environment, Deletion has to be modified accordingly. For example, Itkonen (1964, pp. 59-60) notes that many speakers have Deletion rather than Assimilation before clusters (/huomauttaak # professori/ → *huomauttaa professori* 'remarks the professor'). For these speakers the rules would be stated as in (9):

(9) (a) $k \rightarrow C_t / _ \# C_t V$

(b) $k \rightarrow \phi / _ \# \begin{pmatrix} V \\ CC \\ \text{pause} \end{pmatrix}$

Additional complications in both rules would have to be made for dialects which delete instead of assimilating before *h* and/or *f*.

What the rules in (8) and in (9) miss is that their environments are complementary. They represent as accidental the fact that (b) applies wherever (a) does not apply and vice versa; yet, as the dialect variation shows, this is not an accidental fact but a systematic one. The present theory of phonology can only use conjunctive ordering to express this kind of complementarity. But conjunctive ordering, as illustrated in the rule in (10), will not work in this case:

(10) (a) $k \rightarrow C_t / _ \# C_t$

(b) $k \rightarrow \phi / _ \#$

Rule (10) would wrongly delete the final *-k* before an initial *k*. For example, in /menek # kotiin/ 'go home' we would, after (vacuously) applying (10a), get **mene # kotiin* by (10b). Nor can we put *k*-Deletion first and reformulate Assimilation as an Epenthesis rule: *k*-Deletion would wipe out the underlying distinction between words that end in /k/ and words that end in a vowel so that Epenthesis would wrongly apply to the latter.

Now, given the Elsewhere Condition, we can simply write (10) and have it apply correctly, for (10b) will be automatically interpreted as applying only where (10a) does not apply. Since in /menek # kotiin/ (10a) does apply, albeit vacuously, (10b) will not apply to delete the final *k*. Any change in the scope of (10a) will automatically be reflected in the interpretation of (10b).

A more elaborate example of the same general type is consonant sandhi in Diola-Fogny as reported by Sapir (1965). As Sapir notes, the basic situation is this: "Consonant reduction is achieved by eliding the first two adjacent consonants. If the first consonant is nasal it assimilates where possible without eliding (p. 16)." Examples are given in (11). (The palatal nasal, which Sapir transcribes inconsistently, is here written *ɲ* throughout.)

- | | | | |
|----------|-----------------------------|----------------------|-------------------------|
| (11) (a) | <i>ni + gam + gam</i> | <i>nigangam</i> | 'I judge' |
| | <i>pan + ji + manj</i> | <i>panjimanj</i> | 'you (pl) will know' |
| | <i>ku + bəɲ + bəɲ</i> | <i>kubəmbəɲ</i> | 'they sent' |
| | <i>na + ti:ɲ + ti:ɲ</i> | <i>nati:nti:ɲ</i> | 'he cut (it) through' |
| (b) | <i>na + mi:n + mi:n</i> | <i>nami:mni:n</i> | 'he cut (with a knife)' |
| (c) | <i>takun + mbi . . .</i> | <i>takumbi . . .</i> | 'he must not . . .' |
| (d) | <i>na + laɲ + laɲ</i> | <i>nalalaɲ</i> | 'he returned' |
| | <i>na + yəkən + yəkən</i> | <i>nayəkəyəkən</i> | 'he tires' |
| | <i>na + waɲ + a:m + waɲ</i> | <i>nawaja:waɲ</i> | 'he cultivated for me' |

(c)

(f)

(g)

In the c

(12) (a)

(b)

By c
as show

(13) (a)

(b)

But t
the basic
statement
the fact th
the arrow
spite of th
the theory
The I
way that c
the origin
Rule as in

(e) <i>let + ku + jaw</i>	<i>lekujaw</i>	'they won't go'
<i>ε + rent + rent</i>	<i>ererent</i>	'it is light'
<i>na + manj + manj</i>	<i>namamanj</i>	'he knows'
(f) <i>kuteb # sinajas</i>	<i>kutesinajas</i>	'they carried the food'
<i>eket # bɔ</i>	<i>ekεbɔ</i>	'death there'
<i>ban # ja</i>	<i>baja</i>	'finish now'
(g) <i>napum # kupilak</i>	<i>napuŋkupilak</i>	'he pushed back the children'
<i>najum # tɔ</i>	<i>najuntɔ</i>	'he stopped there'

In the current framework we can state the two rules as in (12):

(12) (a) Assimilation

$$\left[\begin{array}{c} C \\ +nasal \end{array} \right] \rightarrow [\alpha\text{place}] / \text{---} + \left\{ \begin{array}{l} (\#)[+\text{obstruent}] \\ [+nasal] \end{array} \right\} \begin{array}{l} \text{(i)} \\ \text{(ii)} \end{array}$$

(b) Deletion

$$\left[\begin{array}{c} C \\ +nasal \end{array} \right] \rightarrow \phi / \text{---} \left\{ \begin{array}{l} + \left[\begin{array}{l} -nasal \\ -obstruent \end{array} \right] \\ \#[-obstruent] \\ CC \end{array} \right\} \begin{array}{l} \text{(i)} \\ \text{(ii)} \\ \text{(iii)} \end{array}$$

$$\left[\begin{array}{c} C \\ -nasal \end{array} \right] \rightarrow \phi / \text{---} + (\#)C \quad \text{(iv)}$$

By ordering Assimilation after Deletion, we can simplify Assimilation somewhat, as shown in (13):

(13) (a) Deletion

$$\left[\begin{array}{c} C \\ +nasal \end{array} \right] \rightarrow \phi / \text{---} \left\{ \begin{array}{l} + \left[\begin{array}{l} -nasal \\ -obstruent \end{array} \right] \\ \#[-obstruent] \\ CC \end{array} \right\} \begin{array}{l} \text{(i)} \\ \text{(ii)} \\ \text{(iii)} \end{array}$$

$$\left[\begin{array}{c} C \\ -nasal \end{array} \right] \rightarrow \phi / \text{---} + (\#)C \quad \text{(iv)}$$

(b) Assimilation (revised)

$$\left[\begin{array}{c} C \\ +nasal \end{array} \right] \rightarrow [\alpha\text{place}] / \text{---} (\#)C$$

But this description is still not very enlightening. It completely fails to bring out the basic complementarity between Assimilation and Deletion that Sapir's verbal statement of the rules emphasizes. The formal consequences of this weakness include the fact that each one of the rules must mention the feature of nasality on the left of the arrow and also the fact that Deletion is divided into two separate parts which, in spite of their similar form and function, cannot be collapsed by the notations that the theory of Chomsky and Halle (1968) provides.

The Elsewhere Condition, on the other hand, enables us to write the rules in a way that closely corresponds to Sapir's evidently correct verbal formulation. Leaving the original Assimilation Rule unchanged, we can radically simplify the Deletion Rule as in (14b) so that it essentially deletes consonants "elsewhere":

(14) (a) Assimilation

$$\left[\begin{array}{c} C \\ +nasal \end{array} \right] \rightarrow [xplace] / \text{---} + \left\{ \begin{array}{l} (\#)[+obstruent] \\ [+nasal] \end{array} \right\} \begin{array}{l} \text{(i)} \\ \text{(ii)} \end{array}$$

(b) Deletion (under Elsewhere Condition)

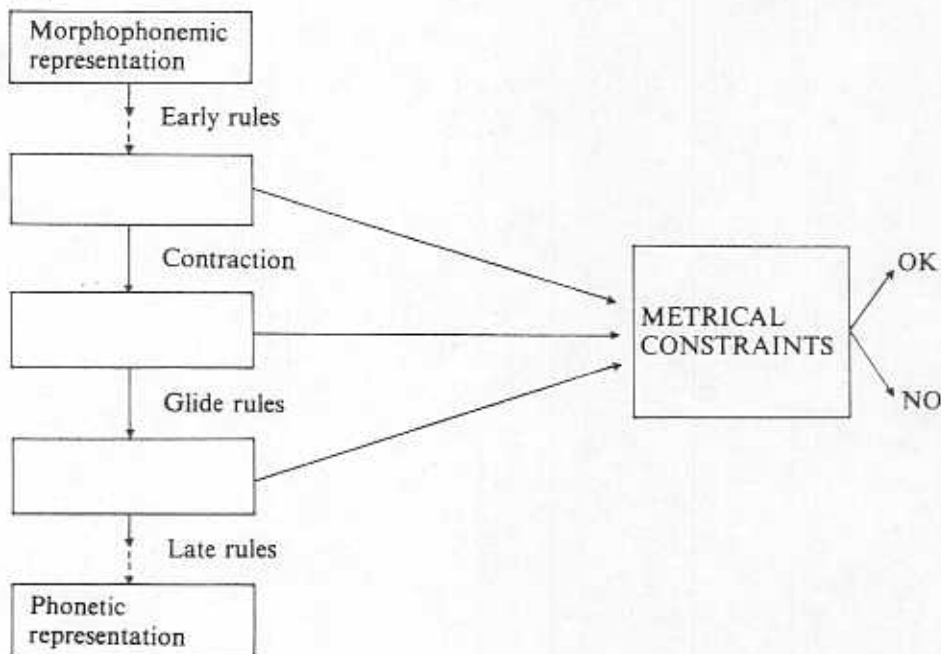
$$C \rightarrow \phi / \text{---} + \left\{ \begin{array}{l} (\#)C \\ CC \end{array} \right\} \begin{array}{l} \text{(i)} \\ \text{(ii)} \end{array}$$

The inputs to both branches (i) and (ii) of the Assimilation Rule are proper subsets of the inputs to branch (i) of the Deletion Rule. Therefore, by the Elsewhere Condition, Assimilation will be disjunctive with branch (i) of Deletion. The rules now say: assimilate, and if assimilation is inapplicable, delete.

A different and somewhat surprising piece of evidence for the Elsewhere Condition, which is of some methodological interest in its own right, comes from metrics. Recent work has shown that there are poetic traditions in which the metrical form of a line is determined not on the basis of its pronunciation, but on the basis of an intermediate representation, or set of representations, in the phonological derivation. In Kiparsky (1972) it is shown that the metrical requirements on a line in the *Rigveda* are allowed to be satisfied either before or after the rules for Vowel Contraction and Glide Formation apply. For example, [ayugdhuam] can be scanned either as four syllables, in accordance with its pronunciation, or as three syllables, in accordance with its morphophonemic representation, which is /a+yug+dhvam/. Conversely, [šacyā] can be scanned either as two syllables, in accordance with its pronunciation, or as three syllables, in accordance with its morphophonemic representation /šaci+ā/.

We show schematically in (15) the linguistic input to the metrical constraints:

(15)



On
be made
must for
tion do
meter. T
intermed
maintain

In I
[y,v] and
other ph
vowels.

Law, a r
language
refer to

(16) GLI
V_ _
| √
| #

In t
vowels in
two diff
-conson
the envir

(17) [+
-
/ūti

ūti

The
reverse c
syllabic i

Give
same ou
between
if and or

(18) [+
-

/ūti
ūti:
ūti

In s
the metr
thereby e

On the basis of evidence presently available, one of the generalizations which can be made about metrical systems of this sort is that the set of input representations must form a consecutive block of rules. Two metrically relevant stages in the derivation do not seem to be separable by a stage which is not a permissible input to the meter. This seems a rather plausible restriction on the psychological availability of intermediate stages in a derivation. Let us now see whether this restriction can be maintained in Sanskrit.

In Kiparsky (1972), it is argued that when the *Rigveda* was composed, glides [y,v] and high vowels [i,ū] were in complementary distribution, but because of various other phenomena underlying glides had to be kept distinct from underlying high vowels. The surface distribution appears to have been in agreement with Sievers' Law, a rule which is widely thought to have operated in the Indo-European protolanguage. Thus we have the distribution of glides and vowels shown in (16); we shall refer to these environments as G and V, respectively:

(16) GLIDES (y,v)	VOWELS (i,u)
V ___ V	___ C
{ \check{V} } (C) ___ V	\check{V} C ₁ ___ V
{#}	CC ___ V

In the present theory of generative phonology, the distribution of glides and vowels in (16) would have to be given by two conjunctively ordered rules. There are two different ways in which this could be done. One way is to have all [+high, -consonantal] segments become syllabic everywhere and then turn nonsyllabic in the environment G of (16). The rules and sample derivations are given in (17):

(17) [+high -consonantal]	→	{ [+syllabic] [-syllabic] in environment G }	(a) (b)
/ūti + ā/	/śaci + ā/	/adug + dhvam/	/ajuṣa + dhvam/
		adugdhuam	ajuṣadhvam (a)
	śacyā		ajuṣadhvam (b)
ūtiā	śacyā	adugdhuam	ajuṣadhvam

The other way to account for the distribution of glides and vowels is simply the reverse of that shown in (17), with everything becoming nonsyllabic first and then syllabic in the environment V of (16).

Given the Elsewhere Condition, we have the same two rules and, of course, the same output. The only difference is that, because of the disjunctive relationship between the two rules, each form is derived in a single step. That is, rule (18b) applies if and only if rule (18a) does not apply:

(18) [+high -consonantal]	→	{ [+syllabic] [-syllabic] in environment G }	(a) (b)
/ūti + ā/	/śaciā + ā/	/adug + dhvam/	/ajuṣa + dhvam/
ūtiā	śacyā	adugdhuam	ajuṣadhvam (a)/(b)
ūtiā	śacyā	adugdhuam	ajuṣadhvam

In spite of the fact that the rules and the output are the same in the two theories, the metrical reflection of these rules allows us to investigate the actual derivations, thereby enabling us to determine which of the theories is correct.

We know that the metrical range includes a level before the application of the glide rules (since, for example, [šacyā] can be scanned as šaciā, its form before the glide rules apply) and also a level after the application of these rules (since, for example, [adugdhuam] can be scanned as adugdhuam, its form after the glide rules apply). But the derivation required in the conjunctive solution contradicts the hypothesis that the metrically relevant stages form a consecutive block in the derivation. Since the theory of Chomsky and Halle (1968) forces us to have conjunctively ordered rules, we have an intermediate stage, that is, when only the first rule has applied, which *ex hypothesi* should also be in the metrical range. But it cannot be there. The critical type of case is represented by *ajušadhvam*, with an underlying and phonetic glide which is also regularly a metrical glide. As shown in (17), under the present theory our rules send this form through the "fake" intermediate stage *ajušadhvam* which cannot serve as input to the metrics. And we encounter the same type of problem if we reformulate the rules taking glide formation as the general case.

In a theory that includes the Elsewhere Condition, on the other hand, the application of these rules is disjunctive, thereby eliminating the "fake" intermediate stage in the derivation which causes the problem in the conjunctive mode of application. We have, then, additional evidence in favor of the Elsewhere Condition.

A noteworthy feature of this example is that the rejected solutions do not "miss a generalization" of any kind insofar as the formal regularities of the system are concerned; both conjunctive and disjunctive ordering are fully adequate to "capture the regularities" inherent in the data. This simply shows that the choice between alternative theories in linguistics turns on considerations other than language-internal "descriptive adequacy" and "explanatory adequacy." Conjunctive ordering is wrong here not because it fails to "capture" a generalization but because it represents the generalization in the wrong way, namely, by requiring a stage in phonological derivations which, as the metrical evidence shows, has no psychological reality.

Consider now expressions like C_0 which are used to abbreviate an infinite set of strings. How should rule schemata containing such expressions be applied? The question was given a somewhat tentative answer by Chomsky and Halle (1968, pp. 313-314): they proposed that each rule in the infinite set of rules given by the schema applies simultaneously. For example, the rule (19) is a schema which stands for the rules in (20):

(19) $C \rightarrow \phi / _ C_0 \#$

(20) (a) $C \rightarrow \phi / _ \#$
 (b) $C \rightarrow \phi / _ C \#$
 (c) $C \rightarrow \phi / _ CC \#$
 (d) $C \rightarrow \phi / _ CCC \#$
 ⋮

These rules would apply simultaneously to a string of the form $XVCCC\#$ to give the output $XV\#$. On the other hand, the rule in (21) would drop only the last C in such a string since Chomsky and Halle assume that rules do not apply to their output:

(21) $C \rightarrow \phi / _ \#$

Such
ern Paiut
mora of

(22) V -

And
I think—
rather, r
applicati
rule "eat
"right-lin
example,
operating

(23) C -

The same
Southern
operating

(24) V -

Let
of rules r
rules app
required,
have the

(25) (a) -

(b) 1

The
reduced v

(26) V -

Itkor
there is a
rule in (2

(27) (a) -

t

(b) 1

Rule
immediat
Itkonen t

(28) V -

° Cho
which are f
not stressed
consequenc

Such rule schemata also figure in many prosodic rules. A rule such as the Southern Paiute Alternating Stress Rule, which stresses the second, fourth, sixth, and so on mora of each word, would be formulated as in (22):⁵

$$(22) V \rightarrow [+stress] / \# C_0 V (C_0 V C_0 V)^* C_0 _$$

Anderson (1969), Johnson (1970), and Howard (1972) have argued—convincingly, I think—that simultaneous application of rules is not the appropriate convention; rather, rules apply iteratively to their own output. Johnson shows that this iterative application must be constrained in a left-to-right or right-to-left direction so that a rule “eats” its way rightward or leftward in a word; he terms this the “left-linear” or “right-linear” mode of application. This eliminates many of the schemata. For example, rule (19), which converts $XVCCC\#$ to $XV\#$ can now be stated as rule (23), operating leftward:

$$(23) C \rightarrow \phi / _ \#$$

The same rule operating rightward drops just the last C, giving $XVCC\#$. And the Southern Paiute Alternating Stress Rule (22) can now be stated simply as rule (24), operating from left to right:

$$(24) V \rightarrow [+stress] / \left[\begin{array}{c} V \\ -stress \end{array} \right] C_0 _$$

Let us now consider the question of disjunctive versus conjunctive application of rules represented by schemata in the cases where they cannot be reformulated as rules applying to their own output. It is clear that disjunctive ordering is sometimes required. According to Itkonen (1966, p. 156), certain dialects of Eastern Cheremis have the rule for word accentuation in (25):

- (25) (a) The accent falls on the syllable containing the last full vowel of the word
 (b) If the word has only reduced vowels, the accent is usually on the first syllable

The corresponding formal rule can be stated as (26), where \check{V} represents a reduced vowel:

$$(26) V \rightarrow [+stress] / _ [C_0 \check{V} C_0]_0 \#$$

Itkonen (1966) also notes that the Eastern Permyak dialects of Komi, in which there is a similar division of vowels into light and heavy, have stress governed by the rule in (27):

- (27) (a) The accent normally falls on the syllable containing the first heavy vowel of the word
 (b) If the word has only light vowels, the accent is on the last syllable

Rule (27) is exactly the mirror image of the Eastern Cheremis rule (25), as is immediately obvious from the formal version in (28), where \check{V} stands for what Itkonen terms a “light” vowel:

$$(28) V \rightarrow [+stress] / \# [C_0 \check{V} C_0]_0 _$$

⁵ Chomsky and Halle’s formulation of this rule (1968, p. 347) incorporates a few complications which are probably best taken care of by separate rules. For example, the fact that final morae are not stressed presumably does not require a special restriction on the Alternating Stress Rule but is a consequence of the devoicing of final morae.

Now note that the success of these rules depends on the *disjunctive* application of the infinite set of cases represented by the environment $[C_0\check{V}C_0]_0$. We require that the *longest* applicable expansion, and *only* this, should apply to each input. Conjunctive application in Cheremis (rule (25)) would stress not only the last full vowel but all the reduced vowels to its right; and words containing only reduced vowels would be stressed not only on the first syllable but on every syllable. As this would be a very peculiar sort of stress situation, we might look for a general principle that would predict the disjunctive ordering of rules like (26) and (28).

We cannot say that disjunctive ordering is associated with *all* schemata that stand for infinite sets of rules. One example in which the ordering appears to be conjunctive is the Umlaut Rule in its original form in the old Germanic languages. The rule, stated in (29), fronts vowels and diphthongs which are followed by *i* or *j* with only consonants intervening:

$$(29) \quad [-\text{consonantal}] \rightarrow [-\text{back}] / \text{ ______ } [-\text{syllabic}]_0 \left[\begin{array}{l} +\text{high} \\ -\text{back} \\ -\text{consonantal} \end{array} \right]$$

I will assume that the raising phenomena in short vowels that tend to be associated with umlauting ($\text{æ} > e$, $e > i$) are due to a separate process (or several separate processes) which, even if collapsible with the Umlaut Rule, will not affect the point to be made.

In Old Norse the rule applies as shown in (30) (Noreen (1923, pp. 57 ff.)):

- (30) $aC_0i > \text{æ}C_0i$ **saljan* > ONorw *sælia*
 (> OIcel *eCi*) > OIcel *selia* 'sell'
 (NB the special case **ai*
 > ONorw *æi* > OIcel *ei*)
- $oC_0i > \text{ø}C_0i$ **lātiz* > *lætr* 'let (2sg)'
 $oC_0i > \text{ö}C_0i$ **norþire* > *nørþre* 'more northerly'
 $oC_0i > \text{ö}C_0i$ **blōtiz* > *blōtr* 'sacrifice to someone (2sg)'
 $uC_0i > \text{ü}C_0i$ **fulljan* > *fylla* 'fill'
 $\bar{u}C_0i > \bar{ü}C_0i$ **lūkiz* > *lýkr* 'close (2sg)'

Before C_0u, w , *a* turns to ɔ ($aCu > \text{ɔ}Cu$, $au > \text{ɔ}u$). If *i* or *j* also follows, the result is $\bar{\text{ɔ}}$.

Now consider the underlying back diphthong /au/. We get $au > \text{ɔ}u$ by *u*-umlaut and, where *i, j* follows, *i*-umlaut to $\bar{\text{ö}}$: [dr̥ɔwm̥r], *droumr*, *draumr* 'dream', [dr̥öwm̥i], *drøy mi* 'dream'. That is, we get umlauting of *both* components of the diphthong. In Old High German, too, back diphthongs have both components fronted ($ou > \bar{\text{ö}}$). The umlauted diphthong is written $\bar{\text{öu}}$, with the second part then unrounded to give $\bar{\text{öi}}$, which is still the pronunciation in some dialects (for example, Baltic German). Standard German turns $\bar{\text{öi}}$ to *oi*.

Now (29) will front both components only if it is interpreted as a *conjunctively* ordered set of rules. Given the input [dr̥ɔwm̥i], we want both applicable subrules of (29) to take effect, that is, the case (31a) of (29), to give [dr̥öwm̥i], and the case (31b), to give [dr̥öwm̥i] = *drøy mi*:

$$(31) \quad (a) \quad [-\text{consonantal}] \rightarrow [-\text{back}] / \text{ ______ } [-\text{syllabic}] [-\text{syllabic}] \left[\begin{array}{l} +\text{high} \\ -\text{back} \\ -\text{consonantal} \end{array} \right]$$

(b)

W
stress e
at least
see this
syllable
rule (2)

(32)

S

Ne
necessa
between
lishes d
No

(33):

(33) [-

Unl
to all the
For exam
not (a)
[dr̥öwm̥i]
then, ma
Of c
are foun
case of i

$$(b) [-\text{consonantal}] \rightarrow [-\text{back}] / \text{---} \left[\begin{array}{l} +\text{high} \\ -\text{back} \\ -\text{consonantal} \end{array} \right]$$

Why are the rule schemata interpreted disjunctively in the Cheremis and Komi stress examples but conjunctively in the Germanic umlaut examples? In these cases at least, the correct interpretation follows directly from the Elsewhere Condition. To see this, consider the relationship of the subrules in each case. Let *S* stand for any syllable and *R* a syllable with a reduced vowel. Then the subrules of the Cheremis rule (25)–(26) are as shown in (32):

$$(32) \quad S \rightarrow [+ \text{stress}] / \left\{ \begin{array}{l} \text{---} \# \\ \text{---} R\# \\ \text{---} RR\# \\ \text{---} RRR\# \\ \vdots \end{array} \right\} \begin{array}{l} (a) \\ (b) \\ (c) \\ (d) \\ \vdots \end{array}$$

Now note that an input subject to any one rule in the infinite list in (32) will necessarily also be subject to all the earlier rules. That is, a subset relationship holds between the inputs of each pair of rules. The Elsewhere Condition therefore establishes disjunctive ordering within the entire set of rules.

Now compare the umlaut example. The subrules of rule (29) are as shown in (33):

$$(33) [-\text{consonantal}] \rightarrow [-\text{back}] /$$

$$\left\{ \begin{array}{l} \text{---} \left[\begin{array}{l} +\text{high} \\ -\text{back} \\ -\text{consonantal} \end{array} \right] \\ \text{---} [-\text{syllabic}] \left[\begin{array}{l} +\text{high} \\ -\text{back} \\ -\text{consonantal} \end{array} \right] \\ \text{---} [-\text{syllabic}] [-\text{syllabic}] \left[\begin{array}{l} +\text{high} \\ -\text{back} \\ -\text{consonantal} \end{array} \right] \\ \text{---} [-\text{syllabic}] [-\text{syllabic}] [-\text{syllabic}] \left[\begin{array}{l} +\text{high} \\ -\text{back} \\ -\text{consonantal} \end{array} \right] \\ \vdots \end{array} \right\} \begin{array}{l} (a) \\ (b) \\ (c) \\ (d) \\ \vdots \end{array}$$

Unlike the case in (32), an input subject to a rule in (33) is not necessarily subject to all the earlier rules; no subset relationship holds between the inputs of the subrules. For example, [namnian] > [nemnian] (Old Icelandic *nefna* 'name') is subject to case (c) but not (a) or (b). Hence the Elsewhere Condition is inapplicable, and a word like [dr̥ɔwm̥i] will be subject to cases (b) and (c), giving *dr̥ɔymi*. The Elsewhere Condition, then, makes exactly the correct predictions in these two cases.

Of course one cannot put too much stock in this argument until more examples are found. I must also say that the applicability of the Elsewhere Condition in the case of infinite sets of rules is intuitively less clear to me than where a simple pair of

rules is involved. Nevertheless, the examples with infinite sets of rules do appear to be consistent with what is predicted by the Elsewhere Condition.

Let us now briefly consider some ways in which the Elsewhere Condition might be generalized. The first possibility is that this condition might be applicable not just to rules which are adjacent in the ordering, but also to rules which are separated by other rules. Another possibility is that a subset relationship in the external context—namely, $P _ Q, R _ S$ in (4)—rather than the whole structural analysis— PAQ, RCS —might suffice to establish disjunctivity.

The possibility of disjunction between nonadjacent rules on the basis of the Elsewhere Condition is suggested, for example, by Halle's analyses of Slavic accent (Halle (1971)). In Halle's system, a number of early accent rules place accent marks on stems and endings of words, depending on morphological features of these stems and endings. Close to the end of the phonology, another accent rule (the "circumflex" rule) puts an initial accent on any word which has not been accented by the earlier rules. As Halle remarks (p. 14, note 8), the circumflex rule is *disjunctive* with respect to earlier accent rules. However, there can be no question of getting this disjunctivity from any parentheses or angled brackets since *other rules intervene* between the early accent placement rules and the circumflex rule. Therefore, in the present theory we must add to the circumflex rule an unsightly annotation which says that the rule only applies to words which as yet carry no accent. In general, such annotated rules must be regarded as indications that something is wrong with the theory. In this case such a conclusion might be supported by the following consideration. Let us imagine a hypothetical Slavic language which differs from the actual languages only in that its circumflex rule lacks the annotation. We might expect this to be the normal case. However, not only is such a Slavic language unattested, but it seems reasonable to assume that it could not exist at all.

The subset relation holds between the early accent rules, which apply to words whose morphemes belong to certain specified categories, and the circumflex rule, which (as Halle states it) applies to any word at all. If, then, we allow the Elsewhere Condition to hold for nonadjacent rules, we obtain the desired interpretation. This again restricts the class of possible grammars.

A second problematic extension of the Elsewhere Condition is indicated by an example that involves external sandhi in Sanskrit. A final dental (including *-s*) assimilates obligatorily in place of articulation to a following coronal stop. In addition, *-s* assimilates optionally to any following segment (in practice, to any following voiceless segment—I assume that *s* before voiced segments has been eliminated by earlier rules). Elsewhere (when a pause follows, or when the optional assimilation rule has not been applied), *s* turns to *h*. I illustrate in (34):

- (34) $t \# t \rightarrow t \# t$
 $t \# c \rightarrow c \# c$
 $s \# t \rightarrow s \# t$
 $s \# c \rightarrow s \# c$
 $t \# p = t \# p$
 $t \# s = t \# s$
 $s \# p \rightarrow \varnothing \# p$ or $h \# p$
 $s \# k \rightarrow x \# k$ or $h \# k$
 $s \# s \rightarrow s \# s$ or $h \# s$
 $s \# s \rightarrow s \# s$ or $h \# s$
 $s \# \rightarrow h \#$

We have,

(35) [+cc

(36) $\left[\begin{array}{l} +c \\ +c \end{array} \right.$

The present

(37) $\left[\begin{array}{l} +c \\ +c \end{array} \right.$

This rule prevents the

(38) $\left[\begin{array}{l} +c \\ +c \end{array} \right.$

But this is Rule (35)

It is not clear that between sufficient inputs. Of $R _$!

At the where Co examples theory.

Anderson,

F

Bright, W.

I

Chomsky,

Halle, M.

Howard,

F

Itkonen, E

Itkonen, T

Johnson,

E

Kielhorn,

Kielhorn,

e

We have, therefore, the Assimilation Rules in (35) and (36):

$$(35) [+coronal] \rightarrow [\alpha\text{place}] / ___ \# \begin{bmatrix} \alpha\text{place} \\ +\text{coronal} \\ -\text{continuant} \end{bmatrix} \quad \text{Obligatory}$$

$$(36) \begin{bmatrix} +\text{coronal} \\ +\text{continuant} \end{bmatrix} \rightarrow [\alpha\text{place}] / ___ \# [\alpha\text{place}] \quad \text{Optional}$$

The present theory does not allow us to write the "elsewhere" case simply as (37):

$$(37) \begin{bmatrix} +\text{coronal} \\ +\text{continuant} \end{bmatrix} \rightarrow \text{h} / ___ \#$$

This rule would apply incorrectly to *devas* # *tiṣṭhati* to produce **devaḥ tiṣṭhati*. To prevent this, we must restrict rule (37) as in (38):

$$(38) \begin{bmatrix} +\text{coronal} \\ +\text{continuant} \end{bmatrix} \rightarrow \text{h} / ___ \# \left(\begin{array}{l} [-\text{coronal}] \\ [+continuant] \\ \text{pause} \end{array} \right)$$

But this restriction merely serves to exclude the case in which the first Assimilation Rule (35) is obligatory, namely, the environment before coronal stops.

It would be desirable to be able to extend disjunctivity to cover a relation like that between rules (35) and (37). The Elsewhere Condition, as formulated in (4), is not sufficient to do this because the requisite subset relation does not hold between the inputs. One possibility is that (4) be extended to the case where $P ___ Q$ is a subset of $R ___ S$, that is, where the external context only is taken into consideration.

At this point such extensions of (4) are hardly more than speculation. The Elsewhere Condition itself, however, seems to work right in a sufficiently wide range of examples to be taken as a serious candidate for a general principle of phonological theory.

REFERENCES

- Anderson, S. (1969), *West Scandinavian Vowel Systems and the Ordering of Phonological Rules*, unpublished Ph.D. dissertation, M.I.T.
- Bright, W. (1957), *The Karok Language*, University of California Publications in Linguistics, 13.
- Chomsky, N., and M. Halle (1968), *The Sound Pattern of English*, New York: Harper & Row.
- Halle, M. (1971), "Remarks on Slavic accentology," *Linguistic Inquiry*, 2, 1-20.
- Howard, I. (1972), *A Directional Theory of Rule Application in Phonology*, unpublished Ph.D. dissertation, M.I.T.
- Itkonen, E. (1966), *Kieli ja sen tutkimus*, Helsinki: WSOY.
- Itkonen, T. (1964), *Proto-Finnic Final Consonants*, Helsinki.
- Johnson, C. D. (1970), *Formal Aspects of Phonological Description*, POLA Reports, 11, Berkeley.
- Kielhorn, F. (1880), *The Vyākaraṇa-Mahābhāṣya of Patañjali*, I-III, Poona.
- Kielhorn, F. (1960²), *The Paribhāṣendusekhara of Nāgojibhaṭṭa*, 2nd ed., K. V. Abhyankar, ed., Poona: Bhandarkar Institute.

- Kiparsky, P. (1972), "Metrics and Morphophonemics in the *Rigveda*," in M. Brame, ed., *Contributions to Generative Phonology*, Austin: University of Texas Press.
- Koutsoudas, A., J. Sanders, and G. Noll (1971), "On the Application of Phonological Rules," reproduced by the Indiana University Linguistics Club.
- Noreen, A. (1923), *Altisländische Grammatik*, Halle: Niemeyer.
- Sapir, D. (1965), *A Grammar of Diola-Fogny*, Cambridge: University Press.
- Wiik, K. (1969), *Suomen kielen morfofonemiikkaa*, Turun yliopiston fonetiikan laitoksen julkaisuja.

The unique characteristics of linguistic ability. They clearly do not communicate before must language.

Many physical and particular occipital c

¹ This Yale University M.I.T., as well as which have