VIII. LINGUISTICS

Academic and Research Staff

Prof. R. Jakobson
Prof. M. Halle
Prof. N. A. Chomsky
Prof. S. Bromberger
Prof. J. A. Fodor
Prof. M. F. Garrett
Prof. K. L. Hale
Prof. J. W. Harris
Prof. L. J. Karttunen
Prof. J. J. Katz
Prof. R. P. V. Kiparsky
Prof. W. A. O'Neil
Prof. D. M. Perlmutter
Prof. Krystyna Pomorska
Prof. J. R. Ross
Dr. Milton Baxter
Dr. Carol E. Conrad
Dr. J. Hankamer
Dr. J. T. Jensen
Dr. Frances E. Karttunen
Dr. Martha Laferriere
Dr. J. A. McGilvray
Dr. Marga E. Reis
Dr. D. L. Stampe
A. T. Hakulinen
R. Hurtig

Graduate Students

E. I. Allen
A. D. Andrews
M. H. Aronoff
Sarah J. Bell
Ivonne A. Bordelois
Nomi Erteschik
R. A. Faraci
R. W. Fiengo, Jr.
J. A. Goldsmith
H. W. Gray
F. R. Higgins
D. Kahn
Ellen J. Kaufman
Mary-Louise Kean
W-C. M. Kim
T. J. Klokeid
Judith Kornfeld
Miriam Lemle
J. Lenerz
M. Y. Liberman
Susan E. Martin
G. L. Milsark
R. T. Oehrle
P. B. Platero
A. S. Prince
P. T. Schwenn
Dorothy Siegel
A. W. Sparer
J. R. Vergnaud
Mary Lou Walch
E. S. Williams
E. F. Woisetschlaeger

A. CONSTRAINING ACTION AT A DISTANCE IN PHONOLOGY

National Institutes of Health (Grant 5 TO1 HD00111-08)
National Institute of Mental Health (Grant 2 PO1 MH13390-06)

J. T. Jensen, Margaret T. Stong

Phonological rules are generally written in the form (1), where we have labeled the elements for convenience of reference.[1]

(1) $A \longrightarrow B \;/\; D \_\_\_\_ X$

input or focus	structural change	determinant	intervening material	= focus

In this notation, A, B, D, and X stand for feature matrices comprising a (possibly null) sequence of segments. The claim made by this notation is that each of the feature matrices in the rule is totally arbitrary and has no necessary relation to the other matrices in the rule. We shall demonstrate the falsity of this claim by showing that X, the intervening part of the environment, is not arbitrary, but is predictable on universal grounds from the feature matrices of A and D.[2]

In a series of papers[3-5] we argued for the existence of an "adjacency" constraint in phonology, of the form:

(2) Adjacency Principle I

If the determinant and focus of a phonological rule are not of the same major class $\begin{bmatrix}\alpha\text{vocalic}\\ \beta\text{consonantal}\end{bmatrix}$, the intervening material must be null. If the determinant and focus are of the same major class $\begin{bmatrix}\alpha\text{vocalic}\\ \beta\text{consonantal}\end{bmatrix}$, the intervening material may contain only segments that are not of that class.

The major classes defined by the features vocalic and consonantal are vowels $\begin{bmatrix}+\text{voc}\\ -\text{cons}\end{bmatrix}$, glides $\begin{bmatrix}-\text{voc}\\ -\text{cons}\end{bmatrix}$, liquids $\begin{bmatrix}+\text{voc}\\ +\text{cons}\end{bmatrix}$ and consonants $\begin{bmatrix}-\text{voc}\\ +\text{cons}\end{bmatrix}$. If the determinant is of one major class, e.g., a vowel, and the focus is of another, e.g., a consonant, then the intervening material must be null. If both the determinant and the focus are of the same major class, e.g., vowels, the intervening material may contain only glides, liquids, and consonants.

Adjacency I predicts the intervening environment in a large number of rules, including Latin LIQUID DISSIMILATION, GRASSMANN'S LAW (in both Greek and Sanskrit), Hungarian VOWEL HARMONY, and UNROUNDING. We shall discuss these rules below in terms of the revised adjacency constraint. But there are two cases, Finnish VOWEL HARMONY and Sanskrit N-RETROFLEXION, which do not conform to Adjacency I.[6] Finnish has neutral vowels both at the surface and underlying levels. VOWEL HARMONY, formulated as (3), skips over neutral vowels.[7]

(3) $$\left[\begin{matrix}+\text{syll}\\ \left\{\begin{matrix}[+\text{low}]\\ [+\text{round}]\end{matrix}\right\}\end{matrix}\right] \longrightarrow [\alpha\text{back}] \;/\; \left[\begin{matrix}+\text{syll}\\ \alpha\text{back}\\ \left\{\begin{matrix}[+\text{low}]\\ [+\text{round}]\end{matrix}\right\}\end{matrix}\right] \left(C_o \begin{bmatrix}+\text{syll}\\ -\text{low}\\ -\text{round}\end{bmatrix} \right)_o C_o \_\_\_\_$$

(3) clearly violates Adjacency I, since the major class features do not distinguish neutral from non-neutral vowels. (Hereafter we shall use the revised feature system, in which the major class features are consonantal, syllabic, and sonorant.)

To predict the intervening environment for Finnish VOWEL HARMONY we propose as the revised Adjacency Constraint:

(4) Adjacency Principle II

Check the determinant and focus for shared major class features. If there are none, X must be null. If there are major class feature(s) in common, construct a feature matrix containing all features common to both determinant and focus. The feature matrix so constructed is just the class of segments that may not appear in X; all other segments may occur freely.

To see how this principle operates for Finnish VOWEL HARMONY, check the input and

determinant of (3). Since they share the major class feature [+syll], we begin to construct a matrix with this feature. We see that the determinant and focus also have the feature disjunction $\left\{\begin{matrix}[+\text{low}]\\[+\text{round}]\end{matrix}\right\}$ in common. Now by (4) we know that the class of segments which is excluded from X is the class $\left[\begin{matrix}+\text{syll}\\\left\{\begin{matrix}[+\text{low}]\\[+\text{round}]\end{matrix}\right\}\end{matrix}\right]$, that is, the class of non-neutral vowels. All other segments – consonants, glides, liquids, and neutral vowels – can freely intervene. If we assume (4), Finnish VOWEL HARMONY can be written:

$$(5)\quad \left[\begin{matrix}+\text{syll}\\\left\{\begin{matrix}[+\text{low}]\\[+\text{round}]\end{matrix}\right\}\end{matrix}\right] \longrightarrow [\alpha\text{back}] \;/\; \left[\begin{matrix}+\text{syll}\\\alpha\text{back}\\\left\{\begin{matrix}[+\text{low}]\\[+\text{round}]\end{matrix}\right\}\end{matrix}\right] X \;\_\_\_$$

X need not be specified, but is predicted by Adjacency II.

Notice that the class of segments that cannot intervene in Finnish VOWEL HARMONY is just the class of segments that constitutes the input to the rule. Irwin Howard[8] has used this observation as the basis of a weaker constraint on intervening material, which he calls the Crossover Constraint.

(6) Crossover Constraint

The intervening material cannot contain any instance of the focus.

The Crossover Constraint is a subcase of Adjacency II. The two constraints make the same prediction for Finnish VOWEL HARMONY, but there are many cases in which they have empirically different results. Consider N-RETROFLEXION in Sanskrit. In this rule n is retroflexed to ṇ if it is preceded by r or ṣ in the same word, and followed by a vowel or glide. Retroflexion takes place across vowels, glides, and non-coronal consonants, but is blocked by coronal consonants:

(7)	nara	'man'	rathena	'the chariot' (instr.)
	narāṇam	'of the men'	darśana	'seeing'
	brahmaṇya	'pious'	grasana	'swallowing'

The rule is written

$$(8)\quad \left[\begin{matrix}+\text{cons}\\-\text{syll}\\+\text{nas}\\+\text{cor}\end{matrix}\right] \longrightarrow [-\text{ant}] \;/\; \left[\begin{matrix}+\text{cons}\\+\text{cor}\\-\text{ant}\\-\text{nas}\\+\text{cont}\\-\text{lat}\end{matrix}\right] \left\{\begin{matrix}[-\text{cons}]\\[-\text{cor}]\end{matrix}\right\}_{0} \;\_\_\_\; [-\text{obst}]$$

(8) applies to narāṇam and brahmaṇya, but it does not apply to the forms in the

right-hand column of (7), since here a coronal consonant intervenes between r and n. Adjacency I is too strong to predict the intervening environment of (8), since r is a liquid and n and ṣ are consonants. Adjacency I predicts that there is no intervening material.[9] On the other hand, the Crossover Constraint is too weak, since it says only that n cannot intervene. This is true, but not sufficient, since no other coronal consonant can intervene either. What prediction does Adjacency II make for N-RETROFLEXION? Since the determinant and focus of (8) share the major class feature [+cons], we begin to construct a matrix with this feature. The only point-of-articulation feature shared by the determinant and the focus is [+cor]. Adding this feature to our matrix, we conclude that the class of excluded segments is $\begin{bmatrix}+\text{cons}\\+\text{cor}\end{bmatrix}$, which is the complement to the class of intervening segments specified in (8). Since Adjacency II makes the correct prediction, we can replace the feature specification $\left\{\begin{bmatrix}-\text{cons}\end{bmatrix}\atop\begin{bmatrix}-\text{cor}\end{bmatrix}\right\}_0$ in (8) by X. Adjacency II has just the right amount of power to define the features of X.

Before we decide to replace Adjacency I with Adjacency II, we must first be sure that Adjacency II handles all cases that Adjacency I was designed for. Consider Latin LIQUID DISSIMILATION, stated in the older feature system,[10] using the major class features consonantal and vocalic:

(9) $\begin{bmatrix}+\text{cons}\\+\text{voc}\end{bmatrix} \longrightarrow [-\text{lat}] \;/\; \begin{bmatrix}+\text{cons}\\+\text{voc}\\+\text{lat}\end{bmatrix} \text{X} \;\_\_\_$

(9) is a minor rule that changes l to r in the suffix -ālis if the last liquid in the stem is l.

(10)	mīlitāris	'military'	from	mīles	'soldier'
	lūnāris	'lunar'	from	lūna	'moon'
	lupānāris	'whorish'	from	lupān-ar	'brothel'

LIQUID DISSIMILATION operates across vowels, glides, and nonliquid consonants. In the feature framework of (9), liquids are considered a major class with the features $\begin{bmatrix}+\text{cons}\\+\text{voc}\end{bmatrix}$. Adjacency II predicts that the class of liquids cannot be contained in X, since the features in common between the determinant and focus are [+consonantal] and [+vocalic].

LIQUID DISSIMILATION is characterized with the revised feature framework of SPE at greater cost.

(11) $\begin{bmatrix} -\text{syll} \\ +\text{cons} \\ -\text{obst} \\ -\text{nas} \end{bmatrix} \longrightarrow [-\text{lat}] \; / \; \begin{bmatrix} -\text{syll} \\ +\text{cons} \\ -\text{obst} \\ -\text{nas} \\ +\text{lat} \end{bmatrix} X \; \underline{\quad}$

Adjacency II makes the correct prediction for (11), since the features shared by input and determinant are $\begin{bmatrix} -\text{syll} \\ +\text{cons} \\ -\text{obst} \\ -\text{nas} \end{bmatrix}$, which define the class of (nonsyllabic) liquids. This feature system, however, creates a problem for Adjacency II. The feature nasal is not a major class feature, but it is needed here to distinguish liquids from nasals. In a more general form of the rule, [-nas] would not be specified, since only liquids are lateral in Latin. But to make Adjacency II work, [-nas] must be stated in the rule. We might require that the input and determinant be fully specified for major class features, or that they be specified for features higher on the hierarchy than lateral. But neither condition can handle (11), since nasal is not a major class feature, nor is it higher than lateral on a hierarchy, if there is such a hierarchy. Since this was not a problem for (9), the fault must lie in the feature system and not in the Adjacency Principle. The main problem is that the new feature system fails to characterize liquids as a major class, but includes them in a class with nasal consonants. The real question is not the validity of Adjacency II, but the status of liquids as a major class.

We encounter an additional problem if we consider that the most nonredundant form of LIQUID DISSIMILATION is

(12) $[\;] \longrightarrow [-\text{lat}] \; / \; [+\text{lat}] \left\{ \begin{matrix} [-\text{cons}] \\ [-\text{voc}] \end{matrix} \right\}_{o} \; \underline{\quad}$

From (12), Adjacency II wrongly predicts that X is null, since determinant and focus have no major class features in common. This problem will be discussed in the context of GRASSMANN'S LAW.

GRASSMANN'S LAW in Greek conforms to Adjacency II. The rule is stated as

(13) $\begin{bmatrix} -\text{syll} \\ +\text{cons} \\ +\text{obst} \end{bmatrix} \longrightarrow [-\text{asp}] \; / \; \underline{\quad} \; X \begin{bmatrix} -\text{syll} \\ +\text{cons} \\ +\text{obst} \\ +\text{asp} \end{bmatrix}$

By comparing the feature matrices of input and determinant, we conclude that obstruents are excluded from the intervening environment. The rule operates freely across vowels, liquids, nasals, and glides. Adjacency I makes the same prediction with the new feature system. In the old feature system, Adjacency I predicts that the rule applies across vowels, glides, and liquids, which is correct for (13) because nasals do not

appear in Greek between underlying aspirated consonants. But in Sanskrit, nasals can occur in such environments:

(14) bandh- from /b^{h}and^{h}/

To predict X for Sanskrit deaspiration, the rule must be written (in the old feature system) with the major class feature sonorant. In the new feature system, Adjacency II predicts the correct intervening environment from (13) for both Greek and Sanskrit.

In GRASSMANN'S LAW we encounter the same problem that we found in Latin LIQUID DISSIMILATION. The statement of (13) can be made more general by eliminating redundant feature specifications. Since in both Greek and Sanskrit, only stops can be aspirated (voiced stops in Sanskrit, voiceless ones in Greek), (13) can be generalized as

(15) $[\ ] \longrightarrow [-\text{asp}] \ / \ \underline{\quad} \ X \ [+\text{asp}]$

In (15) Adjacency II requires X to be null, since determinant and input share no major class features. (Non-stops are vacuously deaspirated by (15).) But since this prediction is wrong, X must be specified – since it is not predictable:

(16) $$[\ ] \longrightarrow [-\text{asp}] \ / \ \underline{\quad} \left\{ \begin{bmatrix} +\text{cons} \\ -\text{obst} \end{bmatrix} \right\}_{0} [+\text{asp}]$$

There are two factors to consider here. (16) is more highly valued than (13) in the present theory, since it contains fewer feature specifications. On the other hand, many rules, such as Finnish VOWEL HARMONY and Sanskrit N-RETROFLEXION, are considerably simplified if the intervening material is specified as X, its content being predicted by Adjacency II. This presents us with a choice. On the one hand, we could assert that rules are written in the form that requires the fewest feature specifications. In some rules, the intervening environment is specified, like (16), and in others, it is indicated by X, like (5), with the features of X determined by Adjacency II. But this choice forces us to lose the generalization contained in Adjacency II. Rather than functioning as a constraint on phonological rules, the Adjacency Principle would be reduced to a minor clause in the evaluation measure. Moreover, it is generally agreed that the status of feature counting in evaluating phonological descriptions is very much in doubt.[11] To preserve the explanatory power of the Adjacency Principle, we have chosen the second alternative: all phonological rules are written in a form that allows the intervening environment to be predicted by Adjacency II. This theory claims that the intervening environment in phonological rules is not arbitrary, but is related directly to the feature composition of the input and determinant. It selects (13) rather than (16) as the

correct description of GRASSMANN'S LAW in Sanskrit and Greek. Adjacency II requires a condition on the statement of rules: that the input and determinant of a rule be specified for redundant major class features.

Next we shall examine two rules from Hungarian, VOWEL HARMONY and UNROUNDING, where the intervening environment is correctly predicted by Adjacency I, to show that they conform to Adjacency II. Hungarian has neutral vowels at the surface level, but at a deeper level, when VOWEL HARMONY operates, there are no neutral vowels. VOWEL HARMONY is most simply stated as

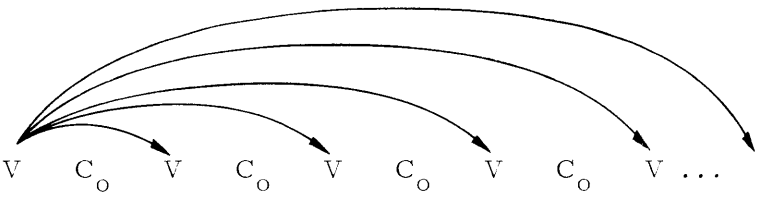
$$(17)\quad [+\text{syll}] \longrightarrow [\alpha\text{back}] \;/\; \begin{bmatrix} +\text{syll} \\ \alpha\text{back} \end{bmatrix} X \;\_\_\_\_ \qquad \text{(L-R Iterative)}$$

The only feature shared by the input and determinant is the major class feature [+syllabic]. Adjacency II makes the correct prediction that syllabic segments cannot be part of X, but that consonants, glides, liquids, and nasals can intervene freely.

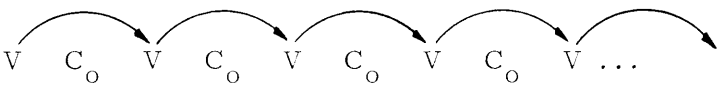
Vowel harmony in Hungarian (and in all other vowel harmony languages that we know of) operates many times in a string, since all vowels in a word agree in backness (except stem vowels in certain foreign words, and neutral vowels in back vowel words, which we regard as a superficial phenomenon). We have argued elsewhere[12, 13] that this multiple application is the result of iterated application, rather than simultaneous application (as claimed in SPE).[14] For example, the iterative rule (17) is written in simultaneous theory as

$$(18)\quad [+\text{syll}] \longrightarrow [\alpha\text{back}] \;/\; \begin{bmatrix} +\text{syll} \\ \alpha\text{back} \end{bmatrix} (\, C_o\ V\,)^*\ C_o \;\_\_\_\_$$

(18) operates by assimilating every vowel in a word to the first vowel of the word. This process is represented schematically as

(19) 

(19) violates the Adjacency Principle (both I and II), since it assimilates a vowel to a vowel across other vowels. It also violates the Crossover Constraint, since it allows the rule to operate across possible foci. The iterative rule (17) operates as in schema (20) and conforms to the Adjacency Principle.

(20) 

The Adjacency Principle predicts the intervening environment for rules that do not iterate, such as Sanskrit N-RETROFLEXION. If we adopt the additional hypothesis of iterative application – which is independently motivated[15] – to supplant simultaneous application, the Adjacency Principle will predict the intervening environment for iterating rules. We have here two independently needed hypotheses that support each other. To preserve the generality made possible by this interacting pair, we must accept them both.

Hungarian UNROUNDING is another iterative rule that is more restricted than VOWEL HARMONY, since only the short vowel ö undergoes the rule. UNROUNDING is written as an iterative rule:

(21) $$\begin{bmatrix} +\text{syll} \\ -\text{high} \\ -\text{low} \\ -\text{back} \\ -\text{long} \end{bmatrix} \longrightarrow [-\text{round}] \; / \; \begin{bmatrix} +\text{syll} \\ -\text{round} \end{bmatrix} X \; \_\_\_\_$$

In (21) only the major class feature [+syll] is shared by input and determinant. Adjacency II correctly predicts that vowels are excluded from X, but that other classes of sounds can intervene freely.

One rather striking confirmation of Adjacency II is a rule in Navajo, STRIDENT ASSIMILATION. In Navajo, strident obstruents occur in two series, anterior and nonanterior:

(22)

+anterior	-anterior
c	$\overset{\vee}{c}$
dz	$d\overset{\vee}{z}$
c'	$\overset{\vee}{c}'$
s	$\overset{\vee}{s}$
z	$\overset{\vee}{z}$

By a regressive rule of STRIDENT ASSIMILATION, strident obstruents within a word agree in anteriority:

(23) a. /$\overset{\vee}{s}$i + má/ ⟶ $\overset{\vee}{s}$imá (no change)
b. /$\overset{\vee}{s}$i + ziiz/ ⟶ siziiz

In paradigm (24), the rule applies across strings of nonstrident segments. In (24), si is the perfective aspect marker, $\overset{\vee}{s}$ is the first person singular marker, ł is a classifier, and ni$\overset{\vee}{s}$ 'to work' is a verb root.

(24)

singular	1	nsisnis	/na + si + s + 1 + nis/
	2	nsínílnis	
	3	naasnis	
nonsingular	1	nsiilnis	
	2	nsoo±nis	
	3	naasnis	

STRIDENT ASSIMILATION is written as

(25) $\begin{bmatrix} +\text{obst} \\ +\text{strid} \end{bmatrix} \longrightarrow [\alpha\text{anterior}] \; / \; \underline{\quad} \; X \begin{bmatrix} +\text{obst} \\ +\text{strid} \\ \alpha\text{anterior} \end{bmatrix}$

It is clear from Adjacency II that X cannot contain segments that are $\begin{bmatrix} +\text{obst} \\ +\text{strident} \end{bmatrix}$.[16]

We have yet to illustrate the operation of the first clause of Adjacency II. The rule of VELAR SOFTENING in English turns certain k to s in derived forms. So we have electric with final [k], but electricity with [s] before the suffix -ity. But VELAR SOFTENING never applies if any segment intervenes between k and i: thus from elect we have election [il ɛ ks ʌ n], where t turns to s before i, but the k before the t remains [k], so we do not have [il ɛ ss ʌ n]. VELAR SOFTENING can be written as

(26) $\begin{bmatrix} +\text{cons} \\ +\text{obst} \\ +\text{back} \\ -\text{cont} \end{bmatrix} \longrightarrow \begin{bmatrix} +\text{cor} \\ +\text{cont} \\ +\text{strident} \\ +\text{anterior} \end{bmatrix} \; / \; \underline{\quad} \begin{bmatrix} +\text{syll} \\ +\text{high} \\ -\text{back} \end{bmatrix}$

Since input and determinant of (26) have no major class features in common, Adjacency II claims that (26) can have no intervening material. This is the correct claim.

Some rules are even more strictly constrained than Adjacency II predicts. Consider, for example, Hungarian VOICING ASSIMILATION. This rule assimilates an obstruent cluster in voicing to its last member.

(27)

a.	víz	'water'	+	től	'away from'		→	[vi:sto:l]
b.	kút	'well'	+	ban	'in'		→	[ku:dbɔn]
c.	küzd	'fight'	+	t	'past' + em	'1st sg'	→	[küsttɛm]
d.	füst	'smoke'	+	ben	'in'		→	[füzdbɛn]

VOICING ASSIMILATION is a right-to-left iterative rule:

(28) $[+\text{obst}] \longrightarrow [\alpha\text{voice}] \; / \; \underline{\quad} \begin{bmatrix} +\text{obst} \\ \alpha\text{voice} \end{bmatrix}$

Since the input and determinant of (28) share the feature [+obst], Adjacency II predicts that nonobstruents can intervene. But this is wrong, since obstruents do not assimilate

in voicing across vowels. If Adjacency II were strengthened to exclude intervening segments in (28), it would not be able to account for the intervening environment in GRASSMANN'S LAW (13).

David Stampe[17] has suggested that phonological operations are of two types: rules and processes. The majority of cases that we have considered in this report are rules, but English VOWEL NASALIZATION (26) is a process. Processes are automatic operations that are innate rather than learned. Rules are learned and can be reversed. By this definition, Hungarian VOICING ASSIMILATION (and similar rules in languages such as Russian and Turkish) are processes rather than rules, since it is difficult to pronounce sequences of obstruents, some of which are voiced and some voiceless. We suggest that rules are subject to Adjacency II, while processes are subject to a stricter Adjacency Constraint, which we may call the Strict Adjacency Condition. This Condition disallows any intervening material between determinant and focus in processes.

References

1. The terms input, structural change, and intervening material are traditional in generative phonology. We have adopted the terms focus and determinant from I. Howard, "A Directional Theory of Rule Application in Phonology," Ph. D. Thesis, Department of Foreign Literatures and Linguistics, M. I. T., 1972 (unpublished) without, however, adopting his theory of directional rules, for a critique of which see J. T. Jensen and Margaret T. Stong, "Ordering and Directionality of Iterative Rules," a paper to be presented at the Conference on Rule Ordering, Indiana University, Bloomington, Indiana, April 6-7, 1973.
2. We assume that propagating rules are described by an iterative process, not by the infinite schema theory of SPE. Infinite schemata constitute a class of exceptions to the theory developed here, although there are many independent grounds for preferring iteration. See J. T. Jensen and Margaret T. Stong (1973), op. cit.
3. Margaret T. Stong, "On the Concept Adjacency in Phonological Theory," a paper presented at the Second Meeting, New England Linguistic Society, Montréal, P. Q., October 23-24, 1971.
4. J. T. Jensen and Margaret T. Stong, "Two Constraints in Phonology: Evidence from Hungarian," 1971 (unpublished).
5. J. T. Jensen, "Hungarian Phonology and Constraints on Phonological Theory," Dissertation, McGill University, 1972.
6. These rules are discussed in greater detail in J. T. Jensen, ibid.
7. There is an additional problem in Finnish Vowel Harmony that has been pointed out to us by R. P. V. Kiparsky. All mixed vowel roots are regular, in that suffix vowel harmony is conditioned by the last non-neutral vowel of the root. But roots containing only neutral vowels are irregular in the way in which they condition suffix harmony. Polysyllabic all-neutral roots take front suffix harmony. But monosyllabic all-neutral roots take back harmony with most vowel initial suffixes, otherwise front harmony (except for two consonant initial suffixes, moinen and lainen, which take back harmony). This wrinkle could be handled by entering suffixes in the lexicon with the vowel that appears after a monosyllabic all-neutral root, with a special fronting rule to front back vowel suffixes after polysyllabic all-neutral roots. This solution is attractive but has some exceptions. Kiparsky has proposed that polysyllabic all-neutral roots actually condition vowel harmony. He posits a special backing

rule to back vowel initial suffixes after monosyllabic all-neutral roots. Although this solution fails to account for moinen and lainen and also violates Adjacency II, it has fewer exceptions.

8. I. Howard, op. cit.

9. According to the major class features consonantal and vocalic which we used in formulating the condition. By using the new major class features, r and n are [-obst] and s is [+obst], so the same problem would arise.

10. J. T. Jensen, op. cit., p. 69.

11. One argument against feature counting as an evaluation measure may be found in J. D. McCawley, The Phonological Component of a Grammar of Japanese (Mouton and Co., The Hague, 1968), pp. 50-52.

12. J. T. Jensen, op. cit.

13. J. T. Jensen and Margaret T. Stong (1973), op. cit.

14. N. Chomsky and M. Halle, The Sound Pattern of English (Harper and Row Publishers, Inc., New York, 1968).

15. J. T. Jensen and Margaret T. Stong (1973), op. cit.

16. STRIDENT ASSIMILATION reverses direction in certain cases. Underlying /ha + si + $\overset{v}{s}$ + tih/ is surface hasistih, not *$\overset{v\ v}{\text{hasistih}}$. Apparently, this reversal is conditioned by a tendency toward paradigmatic regularity. The perfective marker si has the initial consonant s in all other forms of the paradigm. Regressive application of STRIDENT ASSIMILATION would change s to $\overset{v}{s}$ in the first person only. To make the perfective marker consonant the same in all persons, STRIDENT ASSIMILATION reverses direction and applies progressively. We are indebted to K. Hale for these facts about Navajo and for discussion of this analysis.

17. D. Stampe, "How I Spent My Summer Vacation," Ph. D. Thesis, University of Chicago, 1972.

B. FOR-TO COMPLEMENTS AND IF-CLAUSES

National Institutes of Health (Grant 5 TO1 HD00111-08)
National Institute of Mental Health (Grant 2 PO1 MH13390-06)

A. D. Carstairs

It is well known that complementizers do not entirely lack semantic import, as Rosenbaum[1] thought. Joan Bresnan[2] cites the contrasting sentences:

(1) It's funny, I remembered Fred to be handsome, but he wasn't really.

(2) *It's funny, I remembered that Fred was handsome, but he wasn't really.

And Kiparsky and Kiparsky[3] have shown how gerundive (as opposed to infinitive) complements go, in general, with factive predicates. But both Rosenbaum and the Kiparskys seem to assume that it is the matrix sentence (or, more precisely, the predicate of the matrix sentence) that determines the choice of complementizer, the content of the embedded sentence being irrelevant. In this report, I suggest that this is not so, and that some for-to or infinitival complements obey the same restrictions on content as

do some if-clauses. Furthermore, in some circumstances a gerundive (or POSS-ing) complement has a nonfactive interpretation even after a normally factive predicate, and here the same correlation with if-clauses holds.

Consider (3)

(3) I hate it that John is more popular than me.

Hate is a factive verb, so (3) presupposes that John is more popular than me. Contrast (3) with

(4a) I hate it for John to be more popular than me.
(4b) I hate John to be more popular than me.

(4a) and (4b) are, for me, synonymous, and in fact all sentences with for-to complements which I shall mention in this report correspond to synonymous sentences with infinitive complements. But (4a) and (4b) are not synonymous with (3). Consider now what happens when we change the complement slightly in (3) and (4).

(5) I hate it that John is older than me.
(6) *I hate (it for) John to be older than me.

(6) is considerably less acceptable than (4). Why?

Here are some more pairs of sentences that illustrate the same contrast as that between (4) and (6).

(7a) Mary hates John to weigh more than twelve stone.
(7b) *Mary hates John to be over six foot.
(8a) John hates Mary to have long hair.
(8b) *John hates Mary to have a long nose.
(9a) Fred hates Penelope to be ill.
(9b) *Fred hates Penelope to be dead.
(10a) Bill hates the weather to be bad in Boston.
(10b) *Bill hates the climate to be bad in Boston.
(11a) George hates his wife to be learning a foreign language.
(11b) *George hates his wife to know Russian.

What is the crucial difference between, for example, having long hair and having a long nose? The answer is that a person's hair-length can vary but his nose-length normally cannot. (4) and the a sentences in (7-11) have complements whose truth-value can fluctuate over time; in (6) and the b sentences, the complement describes a relatively permanent state of affairs. I propose the term 'iterability' to denote this characteristic of the complements of the a sentences. We can say then that for-to complements of hate must be iterable. Other verbs in this class are like, love, can't bear, and some with which only a for-to (not a bare infinitive) complement is permitted: dislike, loathe,

resent. Prefer belongs here too, though it differs from the others in that it can take a subjunctive that-clause (e.g., 'I prefer that my coffee be black'), as Howard Lasnik has pointed out to me. There are also subject-complement verbs that require an iterable for-to, such as embarrass, bother, but not, e.g., benefit. I shall ignore these mysteries here and continue to concentrate on the one verb hate.

Notice that that-complements of hate are subject to no iterability requirement. (5) is good even though (6) is not, and substituting a that-clause for the infinitive in any of the sentences of (7-11) always yields an acceptable result, e.g., (contrast (11)):

(12a) George hates it that his wife is learning a foreign language.
(12b) George hates it that his wife knows Russian.

A further comment on iterability: Whether we consider that a complement describes an iterable state of affairs or not depends on our beliefs about the world, and the answer may not be clear-cut. (8b) may be acceptable if we are prepared to assume that Mary undergoes plastic surgery to change the length of her nose periodically. Inasmuch as this is less common than hair-dyeing, (8b) is less acceptable than (13), which still sounds stranger than (8a):

(13) John hates (it for) Mary to have red hair.

What other complementizers or conjunctions require that the proposition they introduce be iterable? One such is whenever. Compare (11) with (14):

(14a) George hates it whenever his wife is learning a foreign language.
(14b) *George hates it whenever his wife knows Russian.

(14a) is, for me, synonymous with (11a). The same applies if we have when instead of whenever. So when and whenever qualify for consideration if we are wondering about how to generate sentences with iterable for-to complements. There is, however, a third candidate: if. After simple present hate, if-clauses must be iterable too.

(15a) I hate it if John is more popular than me.
(15b) *I hate it if John is older than me.
(16a) Mary hates it if John weighs more than twelve stone.
(16b) *Mary hates it if John is over six foot.
(17a) John hates it if Mary has long hair.
(17b) *John hates it if Mary has a long nose.
(18a) Fred hates it if Penelope is ill.
(18b) *Fred hates it if Penelope is dead.
(19a) Bill hates it if the weather is bad in Boston.
(19b) *Bill hates it if the climate is bad in Boston.
(20a) George hates it if his wife is learning a foreign language.
(20b) *George hates it if his wife knows Russian.

To determine which, if any, of the three candidates (whenever, when and if) enjoys a special relationship with for-to after hate, we need to find an environment where two of the three yield unacceptable sentences whereas both the third and for-to are acceptable. There is at least one such environment, namely in counterfactual sentences. Consider the following examples, and contrast them with (11a) and (11b):

(21a) George would hate (it for) his wife to be learning a foreign language.
(21b) George would hate (it for) his wife to know Russian.

When the matrix sentence with hate is counterfactual, it tolerates not only an iterable for-to complement (as in (21a)) but also a noniterable for-to complement (as in (21b)). This can be tested by inserting would in (4a), (6), and (7-11). What happens with if, whenever, and when? The crucial examples are:

(22a) George would hate it if his wife were learning a foreign language.
(22b) George would hate it if his wife knew Russian.
(23a) *George would hate it when(ever) his wife was/were learning a foreign language.
(23b) *George would hate it when(ever) his wife knew Russian.

These data support linking for-to complements of hate with if-clauses rather than with when-clauses or whenever-clauses.[4]

There are two further pieces of evidence that tend to confirm this link, although unfortunately speakers' intuitions differ a lot in this area. First, consider sentences differing from (4a), (6), and (7-11) only in that will is inserted, e.g.,

(24a) George will hate (it for) his wife to be learning a foreign language.
(24b) George will hate (it for) his wife to know Russian.

Some speakers are unhappy with these. For me, they are fine, and the meaning is roughly: "George will hate it if it turns out that his wife . . . ." Notice that a noniterable complement (as in (24b)) is acceptable. But, substituting whenever for for-to in (24b), we get a starred sentence – deviant in just the same way as (14b):

(25) *George will hate it whenever his wife knows Russian.

If and when, on the other hand, both yield good results.

(26) George will hate it if his wife knows Russian.
(27) George will hate it when his wife knows Russian.

Interestingly, though, unlike the sentences with if and when after simple present hate, (26) and (27) are not synonymous. (27) presupposes that George's wife will at some time know Russian, whereas (26) does not. Rather, for me at least, (26) means something like "... if it turns out that his wife knows Russian"; it therefore agrees in meaning with (24b), and provides further support for the link between for-to and if. Unfortunately, however, the correlation in acceptability and meaning after will hate is not so

exact as it is after present-tense and counterfactual hate. Consider the following examples:

(28) *Mike will hate it when Sue gets married, so he's sending her to a convent.

(29) Mike will hate it if Sue gets married, so he's sending her to a convent.

(30) ?Mike will hate (it for) Sue to get married, so he's sending her to a convent.

In (28), a consequence of the so-clause (viz. that Mary will stay unmarried) contradicts a presupposition of the when-clause, so the sentence is bad. In (29), the if-clause has no such presupposition, and no contradiction arises. If the parallelism between for-to and if after hate were complete, (30) would be just as good as (29). But most people agree that it is not. For some speakers, (30) seems synonymous with (28), and is unacceptable for the same reason. For me, however, (30) seems about as odd as

(31) ?Mike will hate it if Sue turns out to get married.

Contrast this with (32), which corresponds to the perfectly acceptable (24a).

(32) George will hate it if his wife turns out to be learning a foreign language.

I leave the problem unsolved here.

The third piece of evidence on the if/for-to connection involves a prima facie counterexample to the iterability requirement. As we have seen, (6) breaks the requirement and is unacceptable:

(6) *I hate (it for) John to be older than me.

But if we substitute a definite description for the proper name John, the sentence becomes good, even when the predicate older than me is retained.

(33) I hate (it for) my friends to be older than me.

Note, however, that my friends here cannot refer to specific individuals; if we try to impose that reading the sentence is no better than (6). We must interpret my friends attributively, i.e., as roughly equivalent to 'whoever are my friends at any time'. And the set of my friends in this sense can sometimes contain and sometimes lack members who are older than me, so the iterability requirement is met. The for-to phrase is a transparent context, in Quine's sense,[5] allowing the definite description to have wide scope (cf. J. D. Fodor[6]). The question now is: Are if, when or whenever clauses in corresponding sentences equally transparent? Consider the following:

(34) I hate it if my friends are older than me.

(35) ?I hate it when my friends are older than me.

(36) ??I hate it whenever my friends are older than me.

Some speakers find these all equally good. Some dislike (36) while accepting the other two. But even speakers who dislike both (35) and (36) still consider (33) quite acceptable, just as (34) is. So this evidence provides further slight support for linking these

instances of for-to with if rather than when or whenever.

Assuming, then, that there is a special relationship between certain for-to complements and certain if-clauses, how is this relationship to be expressed in the grammar? For generative semanticists (or 'natural grammarians') the explanation must be syntactic. For interpretivists, there is the alternative possibility of positing special rules of semantic interpretation. I will not take a position on this here, but instead will mention a further mysterious fact, not easily accounted for by either theory. As we have remarked, hate is regularly factive when it has a that-complement:

(3) I hate it that John is more popular than me.

With POSS-ing instead of that, the truth of the complement is still presupposed.

(37) I hate John's being more popular than me.

(3) and (37) go together, then, in contrast to (4a):

(4a) I hate it for John to be more popular than me.

((4a) seems to presuppose that John has been more popular than me at some time, but not that he is more popular than me now.) That POSS-ing complements are typically factive was noticed by the Kiparskys, and the parallelism between that and POSS-ing is preserved in pairs such as the following:

(38) I would hate it that John is more popular than me if I weren't richer than him.

(39) I would hate John's being more popular than me if I weren't richer than him.

The presupposition stays even though the complement is embedded in a counterfactual conditional. Strange things happen, however, if we lop off the if-clause.

(40) *I would hate it that John is more popular than me.

(41) I would hate John's being more popular than me.

(41), but not (40), seems to me just as good as (42), which in turn means the same as (43):

(42) I would hate (it for) John to be more popular than me.

(43) I would hate it if John were more popular than me.

In (42), the for-to phrase seems to function semantically as the protasis of a conditional, whether or not it is derived syntactically from the same source as the if-clause in (43). In (40), we cannot force any such interpretation on the that-clause -- perhaps because this would conflict with the regular factive presupposition -- and the sentence seems, at best, unfinished. But in (41) the POSS-ing complement no longer aligns itself with the that-complement in preserving its factive presupposition. Rather, what seems to happen is that, after hate, a POSS-ing complement is synonymous with a that-complement in environments where a that-complement is permitted, as in (37) and (39). But where a that-complement is not permitted, as in (41) (cf. (40)), the POSS-ing complement is

synonymous with a for-to complement.

If this is a true generalization, it will clearly be impossible to capture it without some mechanism beyond traditional syntactic transformations. For suppose we have a rule optionally changing if to POSS-ing so as to derive (41) from (43) and thereby account for their synonymy; some very strange restrictions will have to be placed on it to avoid the possibility of deriving (37) from (15a) and thereby predicting a nonexistent nonfactive reading for (37). Such restrictions will have to block the application of the rule in just those environments where a that-clause can, as it happens, be substituted for the if-clause so as to yield an acceptable sentence. But this possibility of substitution is clearly not accidental. If we adopt a syntactic approach, we shall need some new mechanism such as Lakoff's transderivational constraints.[7] Alternatively, we could generate if and POSS-ing in all of these sentences by independent syntactic routes, and posit a rule of interpretation stating in effect: Interpret a POSS-ing complement as you would a that-complement if possible, but, failing that, interpret it as a for-to complement. But the theories of transderivational constraints and of interpretation rules are both at such an embryonic stage that I find it hard to decide which approach is less objectionable.

References and Footnotes

1. P. S. Rosenbaum, "The Grammar of English Predicate Complement Constructions," Ph. D. Thesis, Department of Modern Languages, M. I. T., June 1965.
2. Joan W. Bresnan, "Some Properties of Complementizers," Department of Foreign Literatures and Linguistics, M. I. T., 1969 (unpublished).
3. Carol and P. Kiparsky, "Fact," in M. Bierwisch and K. Heidolph (Eds.), Progress in Linguistics (Mouton and Co., The Hague, 1970).

4. Howard Lasnik has pointed out to me that even with matrix verbs outside the hate class, counterfactual environments are more hospitable to for-to complements than noncounterfactual environments:

   (i) ?*For John to be here is amazing.
   (ii) For John to be here would be amazing.

   So the contrast between (11b) and (21b) may be a special case of a more general phenomenon. How seriously this affects my argument will depend on our judgments about examples such as

   (iii) ??If John is here, it's amazing.

   If (i) and (iii) are parallel in oddness, it may be that the if/for-to connection is more general than I suggest. But my intuitions in this area are very weak.

5. W. V. O. Quine, Word and Object (The M. I. T. Press, Cambridge, Mass., and John Wiley and Sons, Inc., New York, 1960).
6. Janet D. Fodor, "The Linguistic Description of Opaque Contexts," Ph. D. Thesis, Department of Foreign Literatures and Linguistics, M. I. T., June 1970.
7. G. Lakoff, "Some Thoughts on Transderivational Constraints," n. d. (unpublished).