XXII. LINGUISTICS

Prof. R. Jakobson
Prof. A. N. Chomsky
Prof. M. Halle
Dr. R. H. Abernathy
Dr. G. H. Matthews

A. THE MORPHOPHONEMICS OF ENGLISH

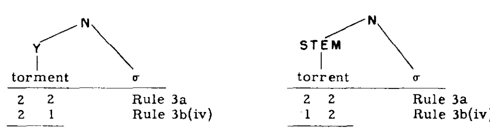
We regard the grammar of a language as a set of rules that account for all well-formed (grammatical) utterances in the language. (The basic features of this type of grammar have been described by Chomsky (1); also see Halle (2).) The grammar must, of course, include an account of the phonetic properties of utterances; that is, a set of rules that contain information on the pronunciation of utterances, their phonetic structure, and so forth. These rules, which we call the "morphophonemic" rules of a language, relate the syntactic representation of an utterance to its phonetic properties. In its syntactic representation an utterance is specified by a string of symbols called "morphemes," which are bracketed together into a hierarchy of nested elements called "immediate constituents" (IC's). The phonetic properties of utterances can be thought as being represented by a very rich ("narrow") phonetic alphabet. The relationship between these two types of representation is expressed in the morphophonemic rules.

When speaking of a phonetic alphabet, linguists have traditionally referred to a set of phonetic properties like voicing, nasalization, palatalization, complexes of which are designated by the symbols of the alphabet. The symbols are, therefore, simply conventional abbreviations standing for the corresponding complexes of properties. In the present work we use the set of phonetic properties that have been proposed by Jakobson et al. (3) under the term "distinctive features." The distinctive features are binary: with regard to them we can ask only whether or not a particular segment possesses a particular feature. It follows that a phonetic representation of an utterance can be regarded as a matrix in which the rows stand for individual features and the columns stand for the segments that constitute the utterance. Since all distinctive features are binary, the matrix will contain only two symbols, plus and minus. (See vowel-system table, page 277.)

Lexical morphemes like <u>dog</u>, <u>slow</u>, <u>bet</u> are represented by distinctive feature matrices. All operations described by the morphophonemic rules are therefore essentially operations on distinctive feature complexes. In the course of applying the rules, feature columns may be added, subtracted, or undergo changes in their constitution. Since lexical morphemes are regarded as distinctive-feature matrices, knowing a particular lexical morpheme is equivalent to having memorized a particular distinctive matrix. It seems reasonable to suppose that memory is at a premium and that morphemes are memorized, therefore, in a form that is most economical of our memory. Hence all possible redundancies in the matrices are eliminated. Specifically, distinctive features that can be inferred from the context (such as, e.g.: in English all vowels are voiced; in initial consonant sequences the first consonant is [s]) are consistently eliminated from the representation and introduced by special (morphophonemic) rules. The rules themselves must also be memorized and must, therefore, be formulated in the most economical fashion; i.e., with minimal distinctive feature composition. This fact plays a deciding role in both the formulation and the ordering of the morphophonemic rules. We should like to stress the ordering in particular because it has very important consequences for linguistic theory, especially for our conception of the differences among dialects and of phonetic change (see Halle (4)).

The rules given below are a part – the most important part – of the morphophonemic rules of General American English. They govern the distribution of the glides [h], [y], and [w], which are shown to be positional variants of the lax vowels [ə̆], [ĭ], and [ŭ], respectively; phonetic alternations in derivational morphology, such as the alternation [ō] ~ [ā] (harmonious~harmo̲nic) or [t] ~ [s] (democrat̲~democrac̲y); and a number of similar processes. Their most important aspect, however, is that they fully predict the distribution of stress (which, therefore, is not distinctive) and the reduction of weakly stressed vowels to [ɨ].

Several factors control the distribution of stresses: the quality of the vowel, its phonetic environment, the morphological category to which a particular morpheme belongs, and the immediate constituent structure of the phrase. All of these have long been known to be determining factors. In the present work we describe exactly the manner in which these factors interact. Every morpheme in isolation has its own stress distribution (which is governed by certain morphological and phonetic factors (see Rule 3)). This stress assignment on the morphemes, however, does not remain fixed; it may be modified by the constituent structure of the utterance in which the morpheme is found, where, again, morphological and phonetic factors play a role. (It is noteworthy that we find here the same phonetic factors operating as in the distribution of stress in simple morphemes. Cf. Rules 4 and 9.) The modifications are introduced in a stepwise fashion, successive steps reflecting the influence of successively higher constituents. Note also that the same modifications apply to all constituents regardless of their place in the constituent hierarchy; the same rules are reapplied to each constituent in a repeating cycle until the highest constituent is reached. The final result of such a cyclical reapplication of the same rules reflects to a certain extent the stress distribution of the morphemes as parts of lower constituents. Thus, for instance, the stress difference in the nouns "torment" and "torrent" is due to their different IC structure. As shown below, "torment," being derived from the verb "torment," has a different structure and hence a different stress pattern from "torrent," which is not a derived noun.



Moreover, the cyclical stress rules have the property of assigning primary stress (1) to a particular vowel and at the same time weakening the stress on all other vowels by one degree. (Rule 4.) Hence, the difference in stress between the vowel that carries primary stress and the one that carries weakest stress generally increases as larger and larger constituents are taken into account. But since there is a natural bound on the number of degrees of stress that can be distinguished, there is also a bound on the size of a constituent that can be realistically handled by the rules. We call the largest of these constituents the phonemic phrase and designate it by special markers. We have not yet succeeded in stating exactly the extent of the phonemic phrase.

The reduction of vowels to [ɨ] is a direct consequence of the stress assignment within words. There are very general conditions, involving stress level, vowel quality, and phonetic context that determine vowel reduction.

Comments on the Notational Conventions

1. Parenthesized symbols may or may not be present in the phrase to which the rule applies.
2. Square brackets enclose distinctive feature columns, which in most cases are only partly specified. On some occasions we have replaced a cumbersome column of distinctive features by its equivalent in a phonetic alphabet or in conventional orthography.
3. Symbols enclosed in braces are equivalent alternatives. Thus, the notation $\begin{Bmatrix} a \\ b \end{Bmatrix} \rightarrow c$ in env. $\begin{Bmatrix} -D \\ E- \end{Bmatrix}$ means "rewrite a or b as c if they precede D or follow E."
4. The abbreviation "in env." stands for "in the environment."
5. We operate with the following vowel system:

Distinctive Features	Phonetic Symbols													
	ū	ī	ō	ē	ɔ̄	ā	ǣ	ŭ	ĭ	ŏ	ĕ	ɔ̆	ə̆	æ̆
vocalic	+	+	+	+	+	+	+	+	+	+	+	+	+	+
consonantal	–	–	–	–	–	–	–	–	–	–	–	–	–	–
tense	+	+	+	+	+	+	+	–	–	–	–	–	–	–
compact	–	–	–	–	+	+	+	–	–	–	–	+	+	+
grave	+	–	+	–	+	+	–	+	–	+	–	+	+	–
diffuse	+	+	–	–	0	0	0	+	+	–	–	0	0	0
flat	0	0	0	0	+	–	0	0	0	0	0	+	–	0

6. The symbol # represents word boundary, and the symbol + represents morpheme boundary. The symbol V represents vowels; i.e., segments that are $\begin{bmatrix} \text{vocalic} \\ \text{-consonantal} \end{bmatrix}$. Arabic and Roman numeral superscripts refer to the relative stress levels. The Roman numerals are introduced to avoid excessive weakening of stresses when the main stress

in a word is weakened because of its syntactic environment. C represents nonvowels – segments that are either [-vocalic] or [consonantal]. The subscripts (superscripts) on C refer to the lowest (highest) number of C in a sequence that may be present for the rule to apply. For example, C_1^2 stands for a sequence of either one or two C's.

7. The symbol σ stands for a substantivizing morpheme that occurs in nouns and adjectives, both simple and compound. Phonetically, it is actualized by assignment of major stress to the first vowel of a simple word or to the first main-stressed vowel of a compound word.

TENTATIVE RULES OF ENGLISH MORPHOPHONEMICS

Precycle

1. $[\breve{i}, \breve{u}]+ \rightarrow \text{zero} \begin{Bmatrix} \text{in env. } + - \# \\ \text{except in env. } - \#, \text{ ate, al, ous, ant, ity, ary} \end{Bmatrix}$

$\text{str}+ \rightarrow \text{st}+\text{r} \quad \text{in env. } - \text{al, ous}$

2. $[\text{-cons}] \rightarrow \begin{bmatrix} \text{voc} \\ \text{-cons} \end{bmatrix}$

$$[\breve{ə}, \breve{i}, \breve{u}] \rightarrow [\text{-voc}] \text{ in env. } \left\{ \begin{matrix} \left\{ \begin{matrix} \begin{bmatrix} \text{-voc} \\ \text{cons} \end{bmatrix} \\ \# \end{matrix} \right\} - (+) \begin{bmatrix} \text{voc} \\ \text{-cons} \end{bmatrix} \\ \begin{bmatrix} \text{voc} \\ \text{-cons} \end{bmatrix} \text{ ——} \end{matrix} \right\}$$

3. Inherent stress rule

(a) V receives stress 2

(b) $\begin{bmatrix} 2 \\ \text{V} \end{bmatrix} \rightarrow \begin{bmatrix} 1 \\ \text{V} \end{bmatrix}$ in env.

(i) $\#\ æ\ C_1 - C_0\ \#$ in verbs

(ii) $\#\ C_0 - C_1 \begin{bmatrix} \text{voc} \\ \text{-cons} \\ \text{lax} \\ \text{-comp} \end{bmatrix} \begin{Bmatrix} \text{rn} \\ \text{st} \end{Bmatrix} \#$ in adjectives

(iii) $\begin{bmatrix} \text{——} \\ \text{tense} \end{bmatrix}$ in suffixes: ate, ory, ify, ize, ee, een, eer, etc.

(iv) $$\left\{ \begin{matrix} \#\ C_0 V C_0 \left\{ \begin{matrix} \begin{bmatrix} \text{——} \\ \text{tense} \end{bmatrix} C_0 \\ \begin{bmatrix} \text{——} \\ \text{lax} \end{bmatrix} [\text{cons}] \begin{bmatrix} \text{cons} \\ \text{-voc} \end{bmatrix} \end{matrix} \right\} \\ \text{——}\ C_0 \begin{bmatrix} 2 \\ \text{V} \end{bmatrix} C_0 \end{matrix} \right\} \begin{matrix} \#\ \text{in verb, adjective} \\ [\text{voc}]\ \#\ \text{in noun, stem} \end{matrix}$$

(v) $\# C_0 - \ldots$ in verb, adjective, noun, stem, where ... does not contain $\begin{bmatrix}1\\ \mathrm{V}\end{bmatrix}$

<u>Cycle</u>

4. Main stress rule:

$$\text{(a)}\ \begin{bmatrix}x\\ \mathrm{V}\end{bmatrix} \rightarrow \begin{bmatrix}0\\ \mathrm{V}\end{bmatrix} \text{ in env. } \left\{\begin{array}{l} \left\{\begin{array}{l} - C_0\,(+) \left\{\begin{array}{l} \begin{bmatrix}2\\ \mathrm{V}\end{bmatrix} C_0 + \bar{e}t \\ \begin{bmatrix}\mathrm{V}\\ \mathrm{lax}\end{bmatrix} C_1^1 + y \\ y\text{ɨ}n \\ ik \end{array}\right. \\ \left\{\begin{array}{l} \begin{bmatrix}\overline{}\\ \mathrm{tense}\end{bmatrix} C_0 \\ \begin{bmatrix}\overline{}\\ \mathrm{lax}\end{bmatrix} [\mathrm{cons}] \begin{bmatrix}\mathrm{cons}\\ -\mathrm{voc}\end{bmatrix} \\ - C_0 \begin{bmatrix}\geqslant 2\\ \mathrm{V}\end{bmatrix} C_0 \end{array}\right\} + \left\{\begin{array}{l} \text{al, ent, ous, sis, um, } \left(\left\{\begin{array}{l}\bar{e}t\\ it\end{array}\right\} + \right)\left\{\begin{array}{l} iv \\ \bar{o}ri \\ \bar{a}yz \end{array}\right\} \end{array}\right\} \\ \ldots \begin{bmatrix}x\\ -\end{bmatrix} \ldots B \cdot - \cdot - + \sigma, \text{ where} \ldots \text{ does not contain } \begin{bmatrix}y\\ \mathrm{V}\end{bmatrix},\ y < x \\ \# C_0 - \ldots + \sigma \\ \text{etc.} \end{array}\right. \\ \begin{bmatrix}1\\ \mathrm{V}\end{bmatrix} \left\{\begin{array}{l} C_0 + \\ \ldots \# \ldots \end{array}\right\} \begin{bmatrix}1\\ -\end{bmatrix} \end{array}\right.$$

$$\text{(b)}\ \begin{bmatrix}x\\ \mathrm{V}\end{bmatrix} \rightarrow \begin{bmatrix}x+1\\ \mathrm{V}\end{bmatrix} \text{ in env. } \ldots - \ldots, \text{ where } \left\{\begin{array}{l} \ldots \text{ contains } \begin{bmatrix}0\\ \mathrm{V}\end{bmatrix} \\ x = 0 \end{array}\right\}$$

5. [ē, ō] → [æ̆, ɔ̆], etc., in context ...

6. $$\begin{array}{l} \begin{bmatrix}2\\ \breve{\text{ɔ}}\end{bmatrix} \rightarrow \begin{bmatrix}2\\ \bar{o}\end{bmatrix} \text{ in env. } - C_0\ \# \\ \begin{bmatrix}1\\ \breve{\text{ɔ}}\end{bmatrix} \rightarrow \begin{bmatrix}1\\ \bar{o}\end{bmatrix} \text{ in env. } - C_0\ (+)i \end{array}$$

7. $\bar{e}t \rightarrow \bar{u}t$ in env. $[\mathrm{cons}]\ w + -$ and $\left\{\begin{array}{l} \text{t drops before ible, ant} \\ \bar{e} \text{ drops before V} \end{array}\right\}$

8. $\mathrm{V} \rightarrow \begin{bmatrix}1\\ \mathrm{V}\end{bmatrix}$ in env. $\# C_0 - C_0\, \mathrm{V} C_0 \begin{bmatrix}1\\ \mathrm{V}\end{bmatrix}$

9. $$\begin{bmatrix}2\\ \mathrm{V}\end{bmatrix} \rightarrow \begin{bmatrix}1\\ \mathrm{V}\end{bmatrix} \text{ in env. } \# C_0 \left\{\begin{array}{l} \left\{\begin{array}{l} \text{(a)} \ldots \begin{bmatrix}1\\ \mathrm{V}\end{bmatrix} C_0 - \\ \text{(b)} \begin{bmatrix}\overline{}\\ \mathrm{lax}\end{bmatrix} \end{array}\right\} \left(C_1^2 +\right)(s) \\ \text{(c)} \ldots \begin{bmatrix}1\\ \mathrm{V}\end{bmatrix} C_0 - \end{array}\right\} [\mathrm{cons}] \left\{\begin{array}{l} \begin{bmatrix}x\\ \mathrm{voc}\end{bmatrix} \\ (+)\ [-\mathrm{cons}] \end{array} \ldots \# \right\}$$

where ... does not contain #; x=1 in case (b), (c) and $x \geq 2$ in case (a); and, furthermore,

(i) $\begin{bmatrix} 2 \\ e \end{bmatrix}[\text{cons}] + y \rightarrow \begin{bmatrix} I \\ e \end{bmatrix} + y$

(ii) $3 \rightarrow I$

(iii) $\begin{bmatrix} 1 \\ \breve{ɔ} \end{bmatrix} \rightarrow \begin{bmatrix} 1 \\ \bar{a} \end{bmatrix}$

(Note connection between 9a and the rules of stress assignment: Precycle rule (3b(iv)) and Cycle rule (4a), second part.)

10. Delete innermost parentheses (i.e., erase constituent structure in the just analyzed substring).

Post-cycle

11. $t \rightarrow s$ before $\begin{cases} \text{yɨn} \\ [\text{i, iv}] \text{ in } [\text{ent, krat, mat, mit, sert, } \ldots] \\ [\text{i}] \text{ in } [\bar{\text{e}}\text{t, abort, excrete, abstract, } \ldots] \end{cases}$

12. $\breve{u} \rightarrow \bar{u}$ except in monosyllables and $- C\ \#$

13. $y \rightarrow \begin{bmatrix} I \\ \breve{i} \end{bmatrix}$ in env. C ——

14. $\begin{bmatrix} \text{I} \\ \text{lax} \end{bmatrix} \rightarrow \begin{bmatrix} \text{II} \\ \text{lax} \end{bmatrix}$ in env. —#

15. $\begin{bmatrix} \text{voc} \\ \text{-cons} \\ \text{lax} \end{bmatrix} \rightarrow \begin{bmatrix} \text{voc} \\ \text{-cons} \\ \text{tense} \end{bmatrix}$ in env. $\left\{ \begin{matrix} \text{—}\ \# \\ \text{—}(+) \begin{bmatrix} \text{voc} \\ \text{-cons} \end{bmatrix} \end{matrix} \right\}$

16. $\begin{bmatrix} \text{-voc} \\ \text{-cons} \end{bmatrix} \rightarrow y$ in context $[-\text{ voc}] - (+)\ [\breve{u}, \bar{u}]$

17. $\bar{u} \rightarrow y\bar{u}$ in non-monosyllables in env. $\left\{ \begin{matrix} [\text{cons}]\ (+) \\ \# \end{matrix} \right\}$ ——

18. $y\bar{u} \rightarrow \bar{u}$ in env. $V \left\{ \begin{matrix} \text{dental} \\ \text{r, l} \end{matrix} \right\}$ ——

19. dental $\rightarrow$ palatal in env. $- \left\{ \begin{matrix} + [\breve{i}, \bar{i}, y] \\ y \end{matrix} \right\} V$

20. $\begin{cases} y \rightarrow \phi \text{ in env. palatal } (+) \text{ ——} \\ i \rightarrow \phi \text{ in env. palatal } + - + \left\{ \begin{matrix} \text{al} \\ \text{ous} \\ \text{ent} \end{matrix} \right\} \end{cases}$

21. $[t^y, d^y] \rightarrow [\check{c}, \check{j}]$

22. $\begin{bmatrix} \mathrm{I} \\ \mathrm{V} \end{bmatrix} \rightarrow$ ɨ in env. – [cons]

23. $\begin{bmatrix} \mathrm{x} \\ \mathrm{V} \end{bmatrix} \rightarrow \begin{bmatrix} \mathrm{I+x-I} \\ \mathrm{V} \end{bmatrix}$ in env. — . . . $\begin{bmatrix} \mathrm{y} \\ \mathrm{V} \end{bmatrix}$, where $y < x$ and . . . does not contain #

24. $\begin{bmatrix} \mathrm{x} \\ \mathrm{V} \end{bmatrix} \rightarrow \begin{bmatrix} \mathrm{x+1} \\ \mathrm{V} \end{bmatrix}$ in env. $\begin{bmatrix} 1 \\ \mathrm{V} \end{bmatrix}$ . . . –

25. I → 1+ maximum of (3, maximum value in phrase); ɨ reduced by 1.

N. Chomsky, M. Halle

References

1. N. Chomsky, Syntactic Structures (Mouton and Company, The Hague, 1957).

2. M. Halle, Questions of Linguistics, Nuovo cimento, Supplement No. 2, 13, 494-517 (1959); The Sound Patterns of Russian (Mouton and Company, The Hague, 1959).

3. R. Jakobson, C. G. Fant, and M. Halle, Preliminaries to Speech Analysis, Technical Report No. 13, Acoustics Laboratory, M.I.T. (1952); R. Jakobson and M. Halle, Fundamentals of Language (Mouton and Company, The Hague, 1956).

4. M. Halle, On the role of simplicity in linguistic description, Proceedings of Symposium on the Structure of Language and Its Mathematical Aspects, New York, April 1960 (to be published).

B. SYNTACTIC CHANGE IN CROW AND HIDATSA

Certain facts about the grammars of Crow and Hidatsa (two languages of the Siouan family) are most economically explained in terms of a grammar of a protolanguage and "innovations" applied to this grammar. Hidatsa has a class of verb stems V_{gh} that contain, among others, the stems ahú (bring), asā́ni (steal), kū́ci (take away), and k$\overset{+}{\mathrm{v}}$u (give). Crow has a comparable verb class V_{gc} containing the cognate stems ṓ (bring), asā́ne (steal), and kū́ce (take away); but it does not contain the stem kū (give), even though this stem does occur in the Crow language.

In the grammar of Hidatsa, there is an optional grammatical transformation, H1, that adds an indirect object to sentences that contain a stem of the class V_{gh}:

$$\begin{array}{lllll} \mathrm{H1} & \mathrm{N} - \mathrm{Af} - \mathrm{V_{gh}} & & 1 = 6 & \\ & 1 \quad\ 2 \quad\ 3 & & & \\ & \mathrm{N} - \mathrm{Af} - \mathrm{N} - \overset{+}{\mathrm{v}}\mathrm{e} & & & \\ & 4 \quad\ 5 \quad\ 6 \quad\ 7 & \longrightarrow & 1, 2, 4, 3 & \end{array}$$

This transformation states that the source of the indirect objects is in the subjects of the verb $\overset{+}{\mathrm{v}}$e (own). Thus, from the sentences ū́tapi Af ahú (the ball was brought) and nā́ka Af ū́tapi $\overset{+}{\mathrm{v}}$e (the child owns a ball) there can be constructed the sentence ū́tapi Af nā́ka ahú (the ball was brought to the child). Similarly, ī́kipi Af macḗ k$\overset{+}{\mathrm{v}}$u (the pipe was given to the man).

Corresponding to this one rule in Hidatsa, there are two optional transformations, C1 and C2, in Crow. The first of these, adds an indirect object to sentences containing the verb stem kū:

$$\begin{array}{llll} \text{C1} & \mathrm{N - Af - k\bar{u}} & & 1 = 5 \\ & 1 \quad\ 2 \quad\ 3 & & \\ & \mathrm{N + Af - N - \bar{e}} & & \\ & \ \ 4 \qquad\ 5 \quad 6 & \longrightarrow & 1, 2, 4, 3 \end{array}$$

Transformation C1 is similar to H1, but instead of being applicable to V_{gc} stems, it applies only to the one Crow verb which is not a V_{gc} stem but which is cognate to a V_{gh} stem in Hidatsa. Another difference between C1 and H1 is that the subject agreement morpheme Af of the verb ē is also carried into the transform. By C1, the sentences ū́sape Af kū (the toy was given) and nā́ke Af ū́sape ē (the child owns a toy) are transformed into ū́sape Af nā́ke Af kū (the toy was given to the child). Similarly, we have ī́cipe Af macḗ Af kū (the pipe was given to the man).

The other Crow transformation, C2, adds a V_{gc} stem to sentences that have already been given an indirect object by C1

$$\begin{array}{llll} \text{C2} & \mathrm{N + Af - N - Af - k\bar{u}} & & 1 = 5 \\ & \ \ 1 \qquad\ \ 2 \quad 3 \quad\ 4 & & \\ & \mathrm{N + Af - V_{gc}} & & \\ & \ \ 5 \qquad\ 6 & \longrightarrow & 1, 2, 6, 3, 4 \end{array}$$

Thus, from the sentences ū́sape Af nā́ke Af kū (the toy was given to the child) and ū́sape Af ṓ (the toy was brought) the sentence ū́sape Af nā́ke ṓ Af kū (the toy was brought to the child) can be constructed.

It is obvious that there is a relationship between the membership of the verb classes V_{gh} and V_{gc} and the transformations H1, C1, and C2, but a description of the relationship that accounts for both the similarities and the differences between the two languages is not immediately apparent. However, by postulating a protolanguage that contains a stem class V_{gp} and the two transformations P1 and P2, we can account for the similarities between these classes and transformations in Crow and Hidatsa. Stem class V_{gp} contains ahú (bring), asā́ni (steal), kū́ci (take away), and k$\overset{+}{\mathrm{v}}$u (give). P1 is like C1, but note that constituent 3 is a V_{gp} stem.

$$\begin{array}{llll} \text{P1} & \mathrm{N - Af - k\overset{+}{v}u} & & 1 = 5 \\ & 1 \quad\ 2 \quad\ \ 3 & & \\ & \mathrm{N + Af - N - \overset{+}{v}e} & & \\ & \ \ 4 \qquad\ 5 \quad 6 & \longrightarrow & 1, 2, 4, 3 \end{array}$$

This produces such sentences as ū́tapi Af nā́ka Af k$\overset{+}{\text{v}}$u and ī́kipi Af macḗ Af k$\overset{+}{\text{v}}$u.

Transformation P2 is similar to C2, except that it is obligatory.

$$
\begin{array}{llllll}
\text{P2} & \text{N + Af} & - \text{ N} & - \text{ Af} & - \text{ k}\overset{+}{\text{v}}\text{u} & 1 = 5 \\
 & 1 & 2 & 3 & 4 & \\
 & \text{N + Af} & - \text{ V}_{\text{gp}} & & & \\
 & 5 & 6 & \longrightarrow 1, 2, 6, 3, 4 & &
\end{array}
$$

This rule produces sentences like ū́tapi Af nā́ka ahú Af k$\overset{+}{\text{v}}$u and, since k$\overset{+}{\text{v}}$u is a V_{gp} stem in the protolanguage, also ī́kipi Af macḗ k$\overset{+}{\text{v}}$u Af k$\overset{+}{\text{v}}$u.

The differences between Crow and Hidatsa can now be explained by postulating two innovations applied to the grammar of the protolanguage. One of these is that the obligatory transformation PC1 is added to the grammar.

$$
\begin{array}{llllll}
\text{PC1} & \text{X} & - \text{ k}\overset{+}{\text{v}}\text{u} & - \text{ Af + k}\overset{+}{\text{v}}\text{u} & - \text{ Y} & \\
 & 1 & 2 & 3 & 4 & \longrightarrow 1, 3, 4
\end{array}
$$

This accounts for the facts cited above concerning the grammar of Crow. Transformation PC1 is actually a particular example of a more general rule that reduces certain sequences of two identical morphemes by removing the first. By PC1, protolanguage sentences like ī́kipi Af macḗ k$\overset{+}{\text{v}}$u Af k$\overset{+}{\text{v}}$u are, in Crow, changed into sentences like ī́kipi Af macḗ Af k$\overset{+}{\text{v}}$u. (PC1, of course, explains only the syntactic relationships of the forms of this Crow sentence: Other innovations will explain the changes in the phonological structure of the forms.) With the addition of PC1 to the grammar of the protolanguage, a restructuring of the grammar took place; transformations P1, P2, and PC1 are replaced by the equivalent set of rules, C1 and C2, on the basis of simplicity. Thus, we can think of the addition of PC1 as an explanation for the two separate, but obviously related, differences between the grammar of the protolanguage and that of Crow. That is, V_{gp} contains the stem k$\overset{+}{\text{v}}$u; V_{gc} does not contain the reflex of this stem; and P2 is obligatory, whereas C2 is optional.

Similarly, through a single innovation applied to the protolanguage, the facts cited above about the grammar of Hidatsa can be explained. This innovation is that the obligatory transformation PH1 is added to the grammar.

$$
\begin{array}{llllll}
\text{PH1} & \text{X} & - \text{ V}_{\text{gp}} & - \text{ Af + k}\overset{+}{\text{v}}\text{u} & - \text{ Y} & \\
 & 1 & 2 & 3 & 4 & \longrightarrow 1, 2, 4
\end{array}
$$

This rule changes all of those sentences of the protolanguage that are produced by P1 and P2 by removing the sequence of subject agreement morpheme plus k$\overset{+}{\text{v}}$u when this follows a V_{gp} stem. By this rule protolanguage sentences like ū́tapi Af nā́ka ahú Af k$\overset{+}{\text{v}}$u and ī́kipi Af macḗ k$\overset{+}{\text{v}}$u Af k$\overset{+}{\text{v}}$u are changed into ū́tapi Af nā́ka ahú and ī́kipi Af macḗ k$\overset{+}{\text{v}}$u. As in the case of Crow, the addition of this transformation to the grammar of the protolanguage and considerations of simplicity bring about a restructuring of the grammar. Transformation H1 is equivalent to the three, P1, P2, and PH1. Thus, by a single innovation, – the addition of PH1 to the grammar of the protolanguage, – the several related differences between the protolanguage and Hidatsa are explained: Hidatsa has no transformation that is a reflex of P2; whereas P1 applies only to $k\overset{+}{v}u$, H1 applies to all members of V_{gh}; and P1 preserves the subject agreement morpheme of the indirect object, whereas H1 does not.

I call the kind of linguistic change exemplified by these innovations "syntactic change." A syntactic innovation adds a transformation to the grammar of a language; then, in some cases, such as those discussed here, the principle of simplicity requires a restructuring of the grammar. Such a change can be thought of as a historical event in that the original grammar no longer exists after the innovation and restructuring have taken place.

G. H. Matthews