

Naturalness Arguments in Syntax

Arnold M. Zwicky, Jr.
(University of Illinois, Urbana)

1. Considerable effort has been devoted, explicitly by generative phonologists and implicitly by many other linguists, to the elucidation of the notions natural class and natural rule in phonological theory. Thus, a set of distinctive features is chosen so that the formal simplicity of the specification for a class of segments corresponds to the naturalness of that class (cf. Halle 1961 and Halle 1962); as an extreme example, the class {p, t, k, b, d, g} is much more natural than the class {p, t, l, ç, z, h}--that is, the segments in the former class, but not those in the latter, can be expected to function together in undergoing (or conditioning) morphophonemic processes and to function together in undergoing (or conditioning) phonological change through time--and any system of distinctive features must reflect this fact. Further notational conventions are required to make the formal simplicity of rules correspond to their naturalness. Another extreme example: a rule shifting t to d, p to b, and k to g in position before voiced obstruents is more natural than one shifting t to b, p to g, and k to d in that position, even though both rules refer to the same class of segments. And again: a rule that makes all obstruents voiced before voiced obstruents and voiceless before voiceless obstruents is more natural than one that makes all obstruents voiced before high vowels and voiceless before nasals, even though the classes of segments involved in the two rules are of approximately the same order of naturalness. It is, however, well known that the careful choice of a universal set of distinctive features together with a judicious selection of such notations as parentheses, brackets, subscripts and superscripts, and variables over + and - is far from sufficient for the purposes of distinguishing relatively natural rules from relatively unnatural ones. For example: the triangular vowel system {i, u, a} is quite natural, while the system {ü, ɨ, ʌ} is not at all natural, even though the two systems are isomorphic with respect to the featural contrasts and redundancies they exhibit. Further: a rule dropping word-initial h is quite natural, while a rule dropping word-initial th is much less so, even though the standard practice of generative phonology makes little or no distinction between the two rules. Such clear examples are by no means isolated. They indicate the need for major revisions in phonological theory, revisions which have been undertaken by a number of investigators (e.g., Chomsky and Halle 1968 and David Stampe in recent, still unpublished, work) who propose the revival, in some form, of the Prague School notion of markedness (i.e., relative unnaturalness).

In syntax, the notions of natural class and natural rule are frequently employed in arguments for and against particular analyses, but the appeal to naturalness is rarely made explicit.

In what follows we shall examine several types of arguments in which considerations of naturalness are significant, and we shall show that just as the attempt to achieve naturalness in phonological descriptions indicates the need for stronger (i.e., more restrictive) hypotheses about the content of phonological theory, so the attempt to achieve naturalness in syntactic descriptions indicates the need for a strengthening of syntactic theory.

2. Let us first consider arguments referring to natural classes. One criticism that can be levelled against many syntactic descriptions is that although the description indicates (by means of a list or, equivalently, a syntactic feature) that certain items constitute a class, C (either by having similar distributional restrictions, or by being mentioned together in some transformational rule, or both), no explanation is provided as to why these particular items, and not some other (perfectly arbitrary) collection of forms, should class together. The criticism is especially compelling when C must be referred to at two or more places in the description, or when C is a semantic class, just as the claim that a phonological class P is natural is supported by the recurrence of P in the description of a particular language (and in linguistic descriptions in general), or by the character of P as a phonetic class. In any case, the point of the criticism is that linguistic theory should provide a descriptive device for the specification of C, at least to the extent that it should be simpler to refer to C than to refer to a class C' of the same size as C but composed of randomly chosen items.

A criticism of the type just described is the one advanced in Ross 1967a against the analysis of the English verbal auxiliaries first presented in Chomsky 1957. In this analysis the sequence

$$\text{TNU} \left(\begin{array}{c} \text{have} \\ \text{be} \\ \text{M} \end{array} \right) ,$$

where TNU is the marker of tense and number and where M represents the class of modals, must be mentioned in a number of transformations--the interrogative inversion rule (Why are you happy? Did he call? How should we go? Has she been hurt?), tag question formation (You are happy, aren't you? He called, didn't he? We shouldn't go, should we? She's been hurt, hasn't she?), one or more rules for placement of adverbs (You are certainly happy. He didn't call. We should soon go. She has never been hurt.), one or more rules of truncation (You are happy, and so is Marvin. He didn't call, and neither did Francesca. We should go, and you should, too. She hasn't been hurt, and I haven't, either.), and a quantifier placement rule (You are all happy. We have both been hurt.). Although it might be possible to reduce the number of references to the

sequence, it seems most unlikely that any amount of ingenuity can reduce this number to less than three (interrogation, adverb placement, truncation). Following Ross, we observe that the parts of the sequence do not necessarily form a complete constituent (in fact, sometimes TNU and he are not even immediately dominated by the same node, except in surface structure; this is the case in John is happy.) and that the formula for the sequence is just as simple as the following absurd formula:

$$\text{PREP} \left(\left\{ \begin{array}{l} \text{fish} \\ \text{the} \\ \text{TNU} \end{array} \right\} \right) .$$

That is to say, the usual analysis of the English auxiliary does not treat the sequence as composed of a natural collection of elements, and a new analysis is called for. An interesting point about this conclusion is that it might very well be in conflict with considerations of simplicity, although in the absence of anything but the grossest indication of an evaluation metric for syntax it is difficult to determine whether or not a conflict does in fact exist. At any rate, the analysis which Ross suggests requires at least one new transformation (in exchange for a saving of a few symbols in each of several other rules), and we may conclude that the issue is certainly not clear-cut, especially in view of the fact that the extent of our detailed, systematic knowledge about the syntax of any language is so slight that the consequences of a choice of one partial analysis over another (even when the first appears clearly to be simpler than the second) are incalculable, so that one can never be certain that a simplification at one point in a description is not canceled out by a complication at some other point. In the case at hand, considerations of naturalness presumably outweigh all simplicity arguments except those based on considerable independent evidence.

Another example: the association of English verbs with various types of complements is by no means random, but rather is clearly related to the meaning of the verbs. The assumption of "structuralist" grammar is that this correlation is an accident--an historical accident, perhaps, but nevertheless an accident. To aspire to naturalness in syntactic descriptions is to take such correlations as the paradigm and (in the absence of other evidence) to class deviations from a regular association as exceptions. Thus the English verbs which occur with a marked infinitive in their complements (for example, persuade, want, and plan) are contrasted with the verbs that occur with the present participle (for example, find, imagine, and avoid), despite the considerable formal similarity between the two classes noted in Chomsky 1958 and elaborated in Rosenbaum 1967,

in that the verbs in the former class (We persuaded him to run for the Senate., Eloise wanted to drown the goldfish., Adolph planned to rule the world.) refer to a time preceding the (not necessarily realized) state described by the complement, while the verbs in the latter class (We found Romulus and Remus building a city., Susan imagined Paul wearing a turtleneck sweater., Delilah couldn't avoid cutting Samson's hair.) do not imply such a sequence. Some verbs may have both senses and hence may occur with both types of complements; the result is sometimes a subtle distinction, as in I hate to wash dishes. (my anticipatory loathing is with me always) as opposed to I hate washing dishes. (my loathing accompanies the act, or at least is greatest during the act), sometimes a very clear one, as in She tried to wash her hair in beer. (but didn't succeed) as opposed to She tried washing her hair in beer. (but couldn't stand the resultant stickiness).

To mark (say, by a feature indicating that a transformation may, or must, apply) these verbs occurring with marked-infinitive complements is to provide an account of the facts, but it is not to explain them. An attempt at explanation requires at the very least that the feature uniting the verbs of this class be supplied by linguistic theory. Otherwise, this group of verbs is no more to be expected to function together than is the set { manufacture, infiltrate, crush, amuse, ascertain, vanish, illuminate, pass, furnish, batten, paint, wane }.

The example is not an isolated one. Indeed, virtually every word class that has received attention in the literature on transformational grammar is characterized by a high degree of semantic coherence. A particularly interesting class, that of verbs occurring with sentential that-complements whose truth is presupposed by the speaker (for example, regret, forget, and grasp, as opposed to suggest, think, and suppose) is discussed in some detail in Kiparsky and Kiparsky 1967. In many instances the semantic coherence of word classes can be strikingly demonstrated by the consideration of nonsense words whose meaning is stipulated to be within the range of some known class of English words. It is possible to make extensive inferences about the syntactic behavior of these words solely on the basis of their meaning, and often one's intuitions about the "expected" versus the "unexpected" properties of the words are quite clear. Thus, C. LeRoy Baker has pointed out to me that given the information that the nonsense word griffle means "move smoothly and quickly in an elliptical orbit," we can infer that the verb griffle can occur not only in such sentences as The beachball griffled with unusual speed. and Many planets griffle about Alpha Centauri., but also in such causative constructions as My paternal uncle can griffle soccer balls with astonishing skill. It is not a criticism of such examples to suggest that they illustrate a

human ability (or propensity) for analogy; one must have a basis to analogize upon, and such a basis is precisely what is lacking in a treatment of English syntax that supposes that the verbs of pure motion (e.g., move, rotate, sail, roll, shift, and drop) constitute a class merely by virtue of their distribution, just as this basis is precisely what is lacking in a treatment of English phonology that supposes that the English grave stops (p, b, m, k, g, ŋ) constitute a class merely by virtue of their distribution.

3. A second type of argument concentrates not upon the fact that certain distinct forms function together, but rather upon the fact that certain distinct functions are accomplished by the same form. In this way, descriptions may be criticized because they fail to explain why a given word or a given morphological category (such as a case or tense) should represent a particular combination of uses or senses.

For example, McCawley 1967b criticizes the treatment of English "empty it" in Rosenbaum 1967 on the grounds that Rosenbaum's analysis (itself an improvement over the earlier analysis in Lees 1960) of such sentences as It seemed to me that she was stark naked, fails to provide any explanation for the occurrence of the neuter pronoun it, rather than (for example) the pro-verb do or the neuter noun mizzenmast, as the representative of the extraposed clause. McCawley points out, quite correctly, that an explanation for it is available only on the basis of the (or a) rule of English definite pronominalization, but that the required "backwards pronominalization" rule simply cannot apply, given present knowledge about the nature of the conditions on pronominalization (Ross 1967b, chapter 5, and Langacker 1967). McCawley suggests instead a radical revision of syntactic theory along lines suggested in Bach 1967. However, it is argued in Kiparsky and Kiparsky 1967 that McCawley's attempt to unify the "empty" it of extraposition with the "factive" it of I hated it that Kermit kept contradicting Margo, is ill-advised, on the grounds that the latter (but not the former) it blocks the formation of relative clauses and questions (What did it seem to Hera that Zeus had accomplished? as opposed to *What did Hera hate it that Zeus had accomplished?); see the discussion of constraints on reordering transformations in Ross 1967b, chapter 4.

The force of the Kiparskys' criticism is that however desirable it might be to press for one sort of natural analysis of English, in which several closely related uses of the pronoun it are explained as resulting from the application of the same rule, there are nevertheless arguments against this proposal--arguments which, moreover, are weighty principally because they directly involve efforts to construct a consistent description embodying significant generalizations about English.

Criticisms might also be made of standard treatments of the Latin genitive case, or the English "possessive," because the natural unity of the inalienable possessive (my arm), alienable possessive (my chair), subjective genitive (my invention), objective genitive (my trial), and partitive genitive (half of me), among other uses, is not captured. Typical discussions of English transformational grammar call for the generation of a possessive marker at half a dozen or more places (see, for example, Lees 1960); the adjunction of the possessive marker in all of these instances is, however, no simpler than would be the adjunction of [Ep] in one situation, [āwd] in another, [I] in a third, [š] in a fourth, and so on--or the generation of the noun plural suffix in one case, the past tense suffix in another, the agentive suffix in a third, the repetitive prefix re- in a fourth, and so on. In this case, as in the previous example, there is some evidence (see Chomsky 1967) to indicate that any attempt to unite certain of the senses of the English possessive, or the Latin genitive, is bound to fail, on the grounds that the rules introducing the appropriate case marking must be separated from each other. Similar remarks apply to the problem of providing a uniform explanation for all, or even many, of the uses of such words as as, so, but, any, which, to, or, or even, or even want, keep, tell, or see.

Valid examples of unity of senses or uses can perhaps be appreciated only in the context of an invalid example of unity. Thus, compare, on the one hand, the natural unity of the genitive of alienable possession and the subjective genitive appearing in nominalizations with, on the other hand, the lack of natural cohesiveness in the senses of the verb decline, according to which one can be said to decline either a regular noun or a dry martini.

4. A third type of appeal to naturalness does not necessarily make reference to classes of items which have a single function or to classes of functions which are manifested as a single item, but rather refers to properties which belong to rules themselves (instead of merely to the classes mentioned by rules or created by them). Thus, the formation of relative clauses and the formation of interrogatives have long been considered to be associated with each other in some special fashion because they share certain properties: they both move a specified noun phrase over a potentially infinite amount of intervening material and to the left end of a clause, they both create "wh-words," and they both are subject to a number of special restrictions (although one noun phrase in a series of conjoined noun phrases may be pronominalized--Frederick said Petula had seen Herbert and him--such a noun phrase cannot be relativized or questioned--*Who did Frederick say Petula had seen Herbert and?, *I slapped the man that Frederick said Petula had seen and him; for further details see Ross 1967b). That is, it has been assumed that the association of relatives and interrogatives is natural and that the similarities in the form

of the required movement transformation and in certain aspects of the products of the transformation are to be taken as evidence for this natural association.

However, arguments have recently been advanced (in Ross 1967b) that relative clause formation and interrogative formation must be considered as two rules, having different structural descriptions and possibly even ordered at different points in the grammar of English. Ross has argued further that the fact that the two rules are restricted in the same ways can be explained as resulting from the fact that all rules which move specified constituents over variables are so constrained. In other words, the similarity in restrictions is deducible from a partial similarity in structural descriptions and structural changes. The other property uniting relative clauses and questions, their characteristic "wh-words," remains unexplained under the re-analysis suggested by Ross.

Another example: the rule creating do so as a substitute for verb phrases has some properties in common with definite pronominalization (compare You may scream, if you wish to do so., *You may do so, if you wish to scream., If you wish to scream, you may do so., and If you wish to do so, you may scream. with Jerry criticized the woman who interviewed him., *He criticized the woman who interviewed Jerry., The woman who interviewed Jerry was criticized by him., The woman who interviewed him was criticized by Jerry.), but some indication is given in Bouton 1968 that the two rules cannot possibly be considered to be the same rule, but must rather be ordered at different points in the grammar, with at least the passivization rule intervening between the do-so rule and the rule of pronominalization, at least in dialects in which The girl was shot in Boston last night by the same man who had done so earlier in New York. is fully grammatical. As a result, no explanation is provided for the fact that the two rules are subject to the same constraints--having digested Bouton's examples, we are unable to express the natural relation between the rules.

5. For a final set of instances in which rule naturalness is referred to, we return to phonology. A phonological rule in which the conditioning segment is always three syllables away from the affected segment is much less natural than a rule in which the conditioning segment is two syllables away, or one syllable away, or an indefinite number of syllables away (i.e., anywhere within the same word as the affected segment). Also, a rule which affects every other syllable in a word is more natural than one which affects every third, or fourth, syllable.

A syntactic analogue is the (presumable) restriction that the structural description of a transformational rule may mention constituents in two structures one of which is directly embedded within the other or one of which is embedded indefinitely far down within the other, but not in two structures one

of which is always two (or three or four) levels below the other. This distinction in rule naturalness is not captured in current theories of transformational grammar, although many other similar restrictions have been incorporated into theory by means of notational conventions.

On the whole the present theory of syntax (roughly that of Chomsky 1965) is not particularly suited to the natural expression of natural classes and rules. The obvious devices for capturing naturalness lead typically to inconsistent analyses. Thus, one simple way to provide an explanation for the fact that a given item *W* appears naturally in several different uses is to postulate that *W* occurs in the underlying structures for the constructions with which these uses are associated. This is the tack taken by McCawley in his discussion of it mentioned above, and also in his more recent treatment of respectively, in which he maintains that he would "regard the addition to linguistic theory of ad-hoc definitions of functions by listing their values a much cheaper price to pay than the alternative of treating respective and respectively as unrelated despite their semantic and morphological identity and their complementary distribution" (McCawley 1967a), but such an approach seems to lead inevitably to contradictions in rule ordering, as we observed above. The type of example mentioned in section 3 is particularly difficult to analyze consistently and appears to require new principles of syntactic theory for its explanation.

An adequate account of natural syntactic classes (section 2) requires more than the assignment of a feature to all members of the class, this assignment being in general the equivalent of a list rather than an explanation for the membership of the class. Two lines of approach suggest themselves. First, it may be possible to press the correlation of semantic and syntactic classes to the conclusion that a syntactic class is no more than a semantic class minus or plus small lists of exceptions. In other words, some natural syntactic classes might be referable to semantic classes in the same way that some phonological classes might be referable to phonetic classes. Second, if a syntactic feature is found (on independent grounds) to play some significant role within linguistic theory, then the assignment by rule of this feature to certain items is not arbitrary or unnatural. Hence, if it is possible to characterize a "rule of pronominalization" as one which assigns the feature PRO, and if there are universal constraints on rules of pronominalization, then the appearance of the same set of pronouns as the products of several different rules (all "rules of pronominalization") would no longer be treated as entirely accidental but would be, to some extent, explained.

BIBLIOGRAPHY

- Bach, Emmon. 1967. Nouns and nounphrases. Unpublished ms.
- Bouton, Lawrence. 1968. Do-so revisited. Unpublished ms.
- Chomsky, Noam. 1957. Syntactic structures. The Hague, Mouton.
- Chomsky, Noam. 1958. A transformational approach to syntax. Proceedings of the third Texas conference on problems of linguistic analysis in English, ed. by Archibald A. Hill, pp. 124-58. Austin, Tex., the Univ. of Texas.
- Chomsky, Noam. 1965. Aspects of the theory of syntax. Cambridge, Mass., MIT Press.
- Chomsky, Noam. 1967. Remarks on nominalizations. Unpublished ms.
- Chomsky, Noam, and Halle, Morris. 1968. The sound pattern of English. New York, Harper and Row.
- Halle, Morris. 1961. On the role of simplicity in linguistic descriptions. Structure of language and its mathematical aspects, ed. by Roman Jakobson, pp. 89-94. (Proceedings of Symposia in Applied Mathematics, 12) Providence, R.I., American Mathematical Society.
- Halle, Morris. 1962. Phonology in generative grammar. Word 18, 54-72.
- Kiparsky, Paul, and Kiparsky, Carol. 1967. Fact. Unpublished ms.
- Langacker, Ronald. 1966. On pronominalization and the chain of command. Unpublished ms.
- Lees, Robert B. 1960. The grammar of English nominalizations. (Indiana University, Research Center in Anthropology, Folklore, and Linguistics, Publication 12) Bloomington, Ind.
- McCawley, James. 1967a. The respective downfalls of deep structure and autonomous syntax. Paper delivered before LSA.
- McCawley, James. 1967b. Where do NPs come from? Unpublished ms.
- Rosenbaum, Peter. 1967. The grammar of English predicate complement constructions. Cambridge, Mass., MIT Press.
- Ross, John Robert. 1967a. Auxiliaries as main verbs. Unpublished ms.
- Ross, John Robert. 1967b. Constraints on variables in syntax. Ph.D. dissertation, MIT.