The Syntactic Component in a Transformational Grammar[1]

Nahm Sheik Park

(Seoul National University)

1. THE ROLE OF THE SYNTACTIC COMPONENT

The syntactic component is the central component in a transformational grammar and the generative source that plays the role of mediating the pairing of meanings to pronunciations of all and only the sentences of a language by generating an infinite set of abstract strings of formatives (i.e. minimal syntactically functioning units) with their structural descriptions. The structural descriptions of the abstract strings of formatives, as we shall see, consist of hierarchically organized underlying and superficial phrase markers, which serve as inputs to the semantic and phonological components respectively. The semantic and phonological components assign concrete meanings and pronunciations to sentences on the basis of their two different phrase markers given as the output of the syntactic component. The superficial phrase markers are derived from the underlying phrase markers by grammatical transformations to be discussed later in the paper.

The strings are abstract in the sense that they do not necessarily correspond to actual sentences in terms of structure including word order but are hypothetically postulated in most instances. The motivations for the abstractness of the strings will be given implicitly or explicitly in pertinent sections below.

2. THE ORGANIZATION OF THE SYNTATIC COMPONENT

The need to generate the two separate, though interrelated, kinds of phrase markers, i.e. underlying and superfiicial phrase markers, necessitates the subdivision of the syntactic component into 1) the base and 2) the transformational subcomponents, the former producing underlying phrase markers and the latter deriving from the underlying phrase markers their superficial phrase markers by grammatical transformations.

[1] For a most recent and detailed discussion of the syntactic component, see Noam Chomsky, *Aspects of the Theory of Syntax*, Cambridge, Mass. MIT Press, 1965. For background assumptions in transformational grammar, which are presupposed througouht, see my "Some Basic Assumptions in Transformational Linguistics," in *Language Research*, Vol. Ⅳ. No. 1 (1968), Seoul National University, Seoul.

It is the main concern of this paper to examine the formal properties of the rules of the base and transformational subcomponents and their interrelations that will enable the syntactic component to succeed in playing the role discussed in 1. The postulation of the rules of this component as well as the other two components is under the empirical and methodological constraint that they predict correctly and most generally the manner in which native speakers produce and understand the sentences of their language.

3. THE ROLE OF THE BASE (SUBCOMPONENT)

As discussed above, the base serves to generate as its output the underlying structures of sentences, which serve as input to the semantic component. This section will discuss in detail what kind of information the base should provide in underlying structures so that the semantic component can assign semantic interpretations to sentences on the basis of that information in a way that matches the manner in which native speakers utilize such information in understanding the meanings of sentences. An exhaustive discussion of this topic is beyond the scope of this paper, but the following considerations seem to be representative of the kinds of information that native speakers utilize in their semantic interpretations of sentences.[2]

3.1. Native speakers understand sentences not merely as linear concatenations of words but rather as hierarchically structured strings of words.[3] "The boy loves his father", for instance, is understood in terms of a hierarchical structure consisting of the two highest level constituents "the boy" and "loves his father," which in turn consist of "the" + "boy" and "loves "+"his father," and so on down the line. In other words, we understand sentences in terms of hierarchical structures consisting of constituents of internal cohesiveness on various levels. Nobody would understand the meaning of the above sentence in terms of "the+loves," "his+loves," "loves+his," etc., because they do not constitute internally cohesive constituents on any level of the hierarchical structure of the sentence. Then it is correct to say that sentences are more than mere linear sequences of words and are understood as hierarchically structured strings of words.

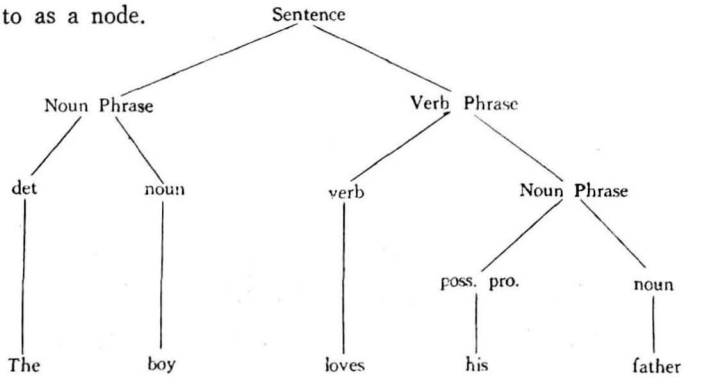
Comparing the above sample sentence with "John has been going out with her," we intuitively feel that "loves his father" and "has been going out with her" are somehow of the same constituent category. We also feel that "the boy," "John," "his father," "her,"

[2] For more discussion of the kinds of information to be included in the base and their formalization, see Noam Chomsky, *Aspects of the Theory of Syntax*, pp. 62—126.

[3] I use "words" here, which may be understood as "morphemes" or "formatives". For convenience sake, I use these three terms interchangeably.

etc., belong to the same category of constituents although they might appear on different hierarchical levels or in different syntactic roles.

Generalizing on these observations, we can represent the hierarchical structure of the sample sentence, in terms of which we understand the sentence, in the following tree diagram with a category label for each constituent, where a constituent name in a tree diagram is often referred to as a node.



Since the hierarchical structure of a sentence representable in a diagram like the above is essential to our understanding of the meaning of the sentence, it must be specified in the underlying structure of that sentence.

Labels given to constituent categories make it possible to determine that certain constituents are of the same, different, related, unrelated categories. Besides, they are necessary for other purposes such as the determination of grammatical relations as we shall see in 4.3. The representation of hierarchical structures without labelled nodes does not make possible the above kind of determination and, therefore, is not fit for the structural analysis of an infinite number of sentences by finite means. This is the main reason for the transformationalists' position that nodes or constituent categories must be labelled. It is clear why the structural analysis without labelled nodes as in the immediate constituent analysis should be rejected, insightful as it may have been. Tagmemics is superior to IC precisely because of its labelling practice.[4]

3.2. The underlying structures of sentences must represent underlying grammatical relations, the motivations for which are evident from the following considerations:

3.2.a. Native speakers of English detect the ambiguity of such sentences as "I know how good meat tastes," "Make her dress fast," "He fed her dog biscuits," "The statistician studied the whole year," "The police were ordered to stop driknking after midnight," "Flying

[4] For inadequacies of the tagmemic labeling practice, see 4.3.

planes can be dangerous," etc., and can assign one of its meanings to each of them as occasion demands. A careful examination shows that the ambiguity of such sentences stems from two or more conflicting grammatical relations in each of them. For one, the two different meanings of "I know how good meat tastes" are due to whether "good" modifies "tastes" or "meat". For this sentence, we should set up two different underlying structures corresponding to the two conflicting grammatical relations commented on if we are to have any systematic basis for the reconstruction of English speakers' ability to assign two different meanings to the sentence.

To generalize, we should posit n-different underlying structures for a sentence which is n-different ways syntactically ambiguous.

3.2.b. Only one underlying structure should be set up for a set of sentences among which syntactic paraphrase relations hold. Syntactic paraphrase relations hold among "John is smarter than I am smart," "John is smarter than I am," and "John is smarter than I," among "John is easy to please," "It is easy to please John," and "To please John is easy," etc. We have to set up only one underlying structure for a paraphrase set of sentences to indicate uniquely and systematically that identical grammatical relations hold and that, therefore, the meaning is the same for all the sentences in the set.

1.3.2.c. Superficial similarities often conceal radically different underlying grammatical relations as in the pairs of sentences "John is easy to please" and "John is eager to please," "The picture was painted by a new student" and "The picture was painted by a new technique," "He was drunk by midnight" and "The beer was drunk by the child," etc. Taking the last pair of sentences as an illustration, we have to show in their underlying structures that what folllows "by" is the logical (i.e. underlying) subject in the second sentence but it is not in the first, since that is how native speakers of English understand the meanings of the two superficially similar sentences.

3.2.d. The preceding arguments 3.2.a.-c. involved the representation of underlying grammatical relations not readily observable from the surface shapes of sentences. However, the specification of underlying grammatical relations is required just as well for sentences, whose surface shapes readily reveal their underlying grammatical relations, if we are to assign two different meanings to such a pair of morphemically identical but order-wise different sentences as "Mary teaches John" and "John teaches Mary".

3.3. In addition to the above kind of grammatical relations, the underlying phrase marker must contain different kinds of grammatical relations called as selectional restrictions. Let

us begin our discussion by considering the selectional restrictions involving verbs. The selection of a verb depends on 1) whether or not a noun follows it as its object, whether or not it can occur freely with manner adverbials, etc., and 2) whether or not nouns, used as the head nouns of the subjects or objects, contain certain features such as human, physical, animate, etc. The first kind of selectional restrictions will say that intransitive verbs cannot occur with objects while transitive verbs can, and will mark such expressions as "The boy went the book," "The boy keeps" as unacceptable or wrong. The second kind of selectional restrictions will say that verbs like "burn up" requires their object nouns to contain the feature "physical," verbs like "read" require their object nouns to contain the feature "legible" and their subject nouns to contain the feature "human," etc. Rules for this second type of selectional restrictions will mark such expressions as "The boy elapsed," "He surprised the desk," etc. as unacceptable. These selectional restrictions must be represented in underlying structures if we are to distinguish such perfectly well-formed sentences as "He went," "A long time elapsed," etc., which are assigned semantic interpretations straightforwardly by English speakers, from such ill-formed sentences as "He went the book," "The boy elapsed," etc., which are not assigned any semantic interpretations or, if any, by analogy to well-formed sentences.[5]

3.4. Since the meaning of any sentence is derived from the combination of the meanings of the words that constitute the sentence, the underlying structures of sentences should contain a specification of the meanings of their constituent words.

It seems to be the case that the level of words in linguistic structures manifests the largest amount of idiosyncracies, the rest of the linguistic structures being characterized largely by regularities. In conformity with our conception of grammar in which the semantic and phonological components are purely interpretive, it is best to list the idiosyncracies of words in the base. The strongest argument for the inclusion of the idiosyncracies of words in the base is that later rules such as grammatical transformations, phonological rules, and projection rules frequently operate on the basis of these idiosyncracies. The idiosyncracies of words, which are not predictable, include, syntactic markers, meanings, pronunciations, grammatical behaviors such as the fact that certain verbs allow object deletion while other do not, etc.[6]

[5] Selectional restrictions have been noted by grammarians for a long time but have been bypassed, for the most part, in linguistic descriptions. The incorporation of rules taking care of these restrictions has been first attempted in a serious manner by transformationalists like Chomsky.

[6] Many qualifications are called for here. However, their discussion is avoided as they are beyond the scope of this paper. Perfectly predictable portions of meanings, pronunciations, grammatical behaviors are left out of the lexicon.

3.5. The preceding disucssion was largely confined to the necessity for the representation of information that is relevant to the semantic interpretations of sentences. This section will be devoted to the argument that a different sort of information that relates to abstract underlying ordering of elememts must be represented in the underlying structures of sentences. The argument rests on two grounds: 1) only the underlying ordering of an abstract nature enables us to capture the maximum degree of underlying regularities in natural languages, and 2) even where some degree of free ordering is allowed the simplest and the most natural description of a language requires abstract underlying ordering. Supporting the first argument is that the abstract underlying ordering in the rule Aux→tense (modal) (have+en) (be+ing) reduces the seeming superficial irregularities of the English verbal auxiliary system to one underlying regularity. As for the second argument, consider a language, in which sentences of three words A,B,C may have three orderings, i.e. A+B+C, A+C+B, B+C+D. If no underlying ordering is set up, we have to have three rules to predict the three ordering possibilities; if an underlying ordering, say, A+B+C is set up, then we need only two later rules. Furthermore, it seems to be the case that one of the allowed orderings, where the so-called free word order is observed, is preferred rather than the rest for a conceptually simple and natural description of the language. Consider as a concrete example "Call that man up" and "Call up that man." The latter expression is preferred to represent the abstract underlying ordering, because we can say that "call up" is a verb behaving like such verbs as "take" in that both are followed by an object in their underlying structures, the discontinuity in "Call that man up" resulting by the application of a later grammatical rule. If we choose "Call that man up" to represent the underlying ordering, then it amounts to recongnizing discontinuity in English verbs as an underlying regularity, which will complicate English grammar considerably. For similar reasons, "It+S+is+easy" is preferred to "It+is+easy+S," where S dominates a string like "Somebody pleases John," as repesenting the ordering in the underlying structure of the paraphrase set comprising "It is easy to please John," "To please John is easy," and "John is easy to please."

3.6. To recapitulate, the underlying structures of sentences generated by the base must contain information on 1) hierarchically categorized structures, 2) underlying grammatical relations, 3) underlying ordering, 4) selectinonal restrictions, and 5) idiosyncracies of words. Providing this range of information is the role of the base.

4. RULES OF THE BASE

Here we are concerned with the definition of the formal properties of the rules appearing

in the base. Put differently, we want to give answers to the question "What kind of rules are necessary for the provision of information of the sort summarized in 3.6. ?"[7]

4.1. Rules for the Representation of Categorial Structures: The hierarchical organization of the underlying structure of a sentence can be formulated in a series of rules of the form X→Y, where X is a single symbol and Y is a string of symbols.[8] As a concrete example, consider the following set of rules: 1) Sentence→Noun phrase+Verb phrase, 2) Verb phrase→verb+Noun phrase, 3) Noun phrase→det+noun, 4) noun→boy, girl, det→the, verb→hates. If we apply the rules in the order given, we will get the following set of derivational lines as one of the two possibilities.[9]

A) Sentence (given as the initial symbol of any base derivation)

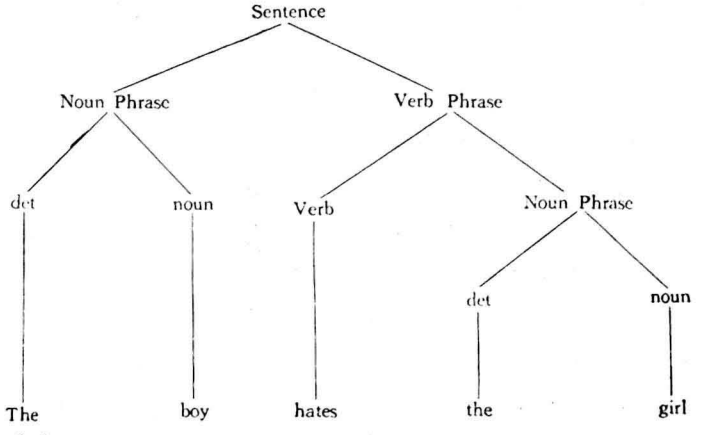
B) Noun phrase+Verb phrase (by rule 1)

C) Noun phrase+verb+Noun phrase (by rule 2)

D) det+noun+verb+det+noun (by rule 3)

E) the+boy+hates+the+girl (by rule 4)

The successive lines of derivation shown above expresses the hierarchical structure of the sentence "The boy hates the girl", which may be made clear in the following labelled tree diagram or bracketting, where S, Np, Vp, n, v, etc., are convenient abbreviations for Sentence, Noun phrase, Verb phrase, noun, verb, etc.



Let us now define several notions that may be useful for our later discussions. The base

[7] For detailed discussions of the formal properties of base rules, see most of the literatures listed in the Bibliography by Chomsky and Postal.

[8] Conditions on the rules of the base are discussed in detail in 4.6—7.

[9] Derivation refers to the process of applying rules and a derivational line is the string of elements resulting from the application of a series of rules as in the base or the transformational subcomponent.

derivation, i.e, the derivation of strings by base rules, starts with the initial symbol #S#. (for sentence) and terminates when all the rules of the base have applied. The final string in a terminated base derivation is called the base terminal string. Symbols appearing in the terminal string are called either terminal or formative symbols; formatives are of two kinds, lexical and grammatical. "boy", "girl", etc. are lexical formatives while "number," "tense," etc. are grammatical formatives. All other symbols appearing in the base derivations are category names such as S, Np, n, v, det, etc. Category symbols are of two kinds: 1) major or phrase category symbols such as S, Np, etc., and 2) non-major or lexical category symbols such as n, v, adj, etc. It is presumed that the category symbols and the rules that interrelate them in the base are substantive and formal universals respectively.

To conclude this section, let us define the notion of the "dominates" or "is a" relation. Given a series of rules: 1) X→Y, 2) Y→A+B, we say that X dominates Y (immediately) and also whatever Y dominates, though indirectly (in this case A+B is dominated by Y immediately and by X indirectly). The "is a" relation is the reverse of the "dominates" relation. If X dominates Y, then Y is an X and whatever is an Y is also an X.[10]

4.2. Convention for the interpretation of Underlying Ordering: If given a categorization rule such as S→Np+Vp discussed in 4.1, we say by an interpretative convention that the rule fixes the abstract underlying ordering from left to right, i.e. Np+Vp in this order. Then given a series of categorization rules, the abstract underlying ordering is automatically determinable by this convention.

4.3. Convention for the interpretation of Underlying Grammatical Relations: The notion "grammatical relation" is a functional notion expressing the role or function of a certain constituent category with respect to the whole or part of a sentence. For instance, the grammatical relation "subject of a sentence" refers to the function of an Np, that is dominated (together with a Vp) by S immediately, with regard to S. Then the functional notion of a grammatical relation should be clearly distinguished from the categorial notion such as that of Np. It is maintained by some linguists that the grammatical relational notion can always be identified with the categorial notion. Notably, the tagmemicists use in this spirit the notation "Clause=Sub: Np, Pred: V, Obj:Np, which identifies the two different notions. Such a tagmemic notation is wrong on two counts: 1) the confusion of two different notions and 2) redundancy of notation as we shall demonstrate in the

[10] The notion defined here is especially important for the interpretative convention which determines grammatical relations to be discussed in 4.3.

following paragraph.[11]

In terms of the "dominates" or "is a" relation within the framework of hierarchically categorized underlying structures (We need further specifiations to have underlying structures.),[12] it is possible to detertmine underlying grammatical relations by the following interpretative convention: given a rule of the form X→A+B, where A and B may dominate further substrings, A and whatever it dominates have the grammatical relation A-X to X, and B and whatever it dominates have the grammatical relation B-X to X. On the basis of this convention, which relies essentially on the configuration of underlying structures, we can define such grammatical relations as the subject of a sentence, the predicate of a sentence, the main verb of a predicate, the object of a predicate, etc., as Np-S to S, Vp-S to S v-Vp to Vp, Np-Vp to Vp, etc. Further, the relation between the subject of a sentence and the main verb of a sentence can be defined as the relation between Np-S to S and v-Vp to Vp, where Vp has the relation Vp-S to S, etc. Since the categorial structure of an underlying structure represented in terms of hierarchically arranged labels of categories makes the prediction of grammatical relations completely automatic, the dual labeling such as Subj:Np is certainly redundant.

4.4. Rules for the Representation of Selectional Restrictions: The rule schema for selectional restrictions should differ from the schema for categorial rules used in 4.1.[13] For 1) categorial rules simply rewrite or expand symbols without regard to contextual environment while selectional rules must take account of the contexts, and 2) categorial rules typically categorize major (i.e. phrase) categories into lower level categories such as in the rule S→Np+Vp, while selectional rules involve the choice of the subcategories of lexical categories as in the rule v→ vi/—#, vt/—Np.

As already noted, there are two kinds of selectional restrictions depending on whether contextual environment is stated in terms of categorial symbols such as noun, det, etc., or of certain features of words such as animate, human, physical, etc. The first kind would have to have rules of the form n→{nl/det___, n2/$\overline{\text{det}}$___}. The second kind would have to have rules of the form v→{vl/___n-animate, v2/___n-physical}, etc.[13] Note that the

[11] Tagmemicists clearly would not like to say that "John" in "John was killed" is the subject of the sentence, which would conflict with their insistence on the identification of slot and filler or categorial and grammatical notion.

[12] I am using the still undefined notion "underlying structure or phrase marker" for the sake of argument and consistency.

[13] Rules of 4.1 and those of 4.4. are called categorial (or phrase structure or rewrite) rules and

second type of rules involve the choice of subcategories of a lexical category in terms of the feature or features in the heads of grammatically related phrases such as the choice of the verb subcategory "terrify," etc., in terms of the feature animate in the head noun of its object Np as in "Terrify the people." Even this kind of relations can be defined in terms of an extension of the convention given in 4.3.

4.5. "Rules" for the Representation of the Idiosyncracies of Words: "Rules" here would just list the unpredictable features of words. The general practice among transformational grammarians is to list the phonological structure of a word on the left hand side of the rule and the rest of the idiosyncracies on the right hand side of the rule of 1) syntactic markers 2) meaning, 3) grammatical peculiarities 4) selectional restrictions. As a concrete example, let us consider how to formulate the rule for the word "still" in the sense of silent. The phonological structure of this word would be given as "C+alveolar stop+high front vowel+lateral."[14] The initial segment is specified merely as a consonant because the phonotactics of English makes it possible to predict on the basis of this specification to predict that it must be /s/ in this position; the second segment is not specified with regard to voicing and aspiration, which are entirely predictable by general rules of English phonotactics if given the information that it is a stop consonant preceded by C in an English consonant cluster. The syntactic marker would just be "adjective." The meaning of the word would be "lack of noise".[15] The grammatical peculiarities would include: No adverbialization by suffixing -ly, No attributive use such as *still night, etc.

The predictable features of words are predicted by redundancy rules. Redundancy rules are of three kinds: 1) syntactic, 2) semantic, and 3) phonological redundancy rules, the latter two of which will not be discussed here mainly for space reasons It is the case in English that verbs that can occur with manner adverbs may occur without them. Thus in the specification of the grammatical peculiarities of the words, we mention that these verbs occur with manner adverbs and then a syntactic redundancy rule of the form "v+manner adv→v" will later predict that the verbs in question can occur without manner adverbs.

Incidentally, selectional restrictions, as we have discussed earlier, involve lexical categories only and consequently they seeem to be lexical peculiarities. In 5., it will be argued that

---

selectional rules respectively.

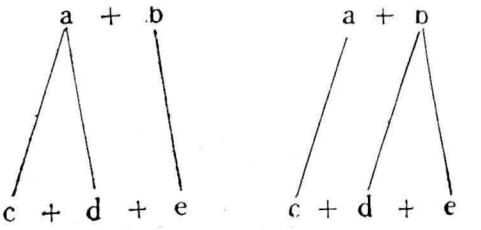
[14] This rule schema will be revised as our discussion proceeds.

[15] Actually binary feature notation is used for the specification of most of the "unpredictable" phonological and semantic features of lexical items. The representations given here are intended as informal abbreviations for feature complexes.

selectional restrictions and lexical idiosycracies should appear in the lexicon.

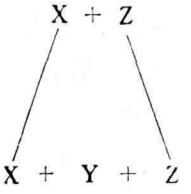
4.6. Conditions on Categorial Rules: Categorial rules should meet several conditions if they are to provide explicitly in the underlying structures of sentences information concerning their underlying grammatical relations, hierarachical structures, and abstract orderings. The following are the five conditions imposed on categorial rules.[16]

Condition 1: X of the rule X→ Y should be a single symbol. The reason for this is that if X is more than one symbol there would be no unique method of assigning "the dominates" or "is a" relation, which is fundamental to the determination of grammatical relations, hierarchical structure, and abstract ordering. Suppose that X=a+b and Y=c+d+e, then there would be no non-arbitrary way of assigning d to a or b as the above diagrams show. A rule of the form Vp+Np→ v+det+n+pronoun cannot tell us just how much is dominated by Vp or Np.



Condition 2: Y of the rule X→ Y is not identical with X. This condition is imposed to avoid the vacuity of rule application. If X=a and Y=a, then a dominates a which apparently does not tell us anything about categorial structure or grammatical relations, for which the rule is designed, as is evident from the diagram: a→a a....

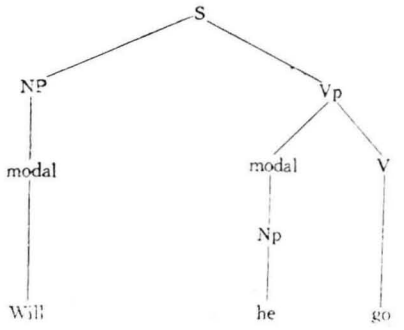
Condition 3: A symbol cannot be rewritten as itself plus something else. If X→ X+Y, where prior to the application of this rule X was already concatenated by a symbol, say, Z on the right, it will be impossible to associate non-arbitrarily the resulting line of derivation to the original one as is shown in the diagram on the left:



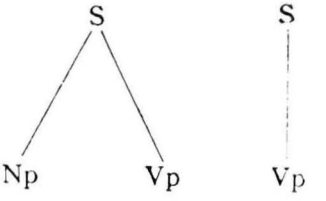
Given the rules 1) S→ Np+Vp and 2) Np→ Np+modifier, the successive lines of derivation will be 1) Np+Vp and 2) Np+modifier+Vp, where the constituency relations between the first and second lines are not uniquely determinable. Hence, the grammatical relation that the modifier bears to the rest of the sentence will not be uniquely determinable.

[16] For discussions of the conditions on categorial rules, see most of the literatures listed in the Bibliography, especially Paul. M. Postal, *Constituent Structure* (The Hague: Mouton & Co, 1965), esp. Chapter 3.

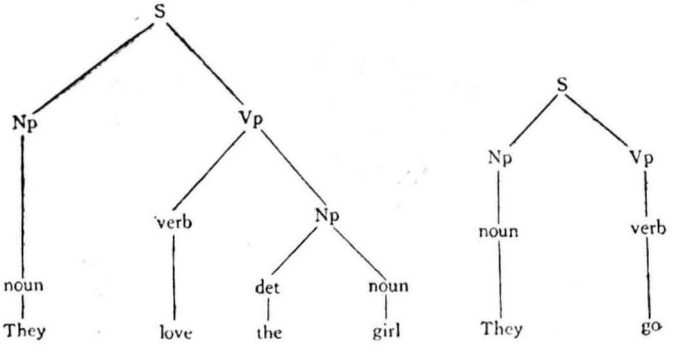
Condition 4: No rule of the form X→ Y, where X=a+b and Y=b+a, is permitted. In other words, no permutation is allowed. The reason for the imposition of this condition is that, we would be making the absurd claim that a becomes b and b becomes a, which will lead to counterintuitive "dominates" or "is a" relations. Consider the following set of rules and the tree digram that follow from them: 1) S→ Np+Vp, 2) Vp→ modal+v, 3) Np→ modal, modal→ Np, 4) Np→ he, Modal→ will, v→ go. No native speaker of English would accept the idea that a modal is an Np and an Np is a modal.



Condition 5: Y of the rule X→ Y must not be null. If X is rewritten as null, X is erased and there is nothing left about X. Hence, there will be no recovery of the constituent structure or grammatical relations involving X. As a concrete example, consider the following set of rules and the resulting tree diagram: 1) S→ Np+Vp, and 2) Np→ ϕ.



4.7. Conditions on Selectional Rules: Our problem here is fixing the contextual environment relevant to selectional rules. First let us consider the rule schema for the representation of selectional rules, which was tentatively formulated as V→ {Vi/___#, Vt/___ Np} in 4.4. We feel that "love the girl" in "They love the girl" and "go" in "They go" are of the same category, i.e. Vp. The verbs "love" and "go" differ in their selectional restrictions in that the former always requires an object Np (in this case "the girl") while the latter does not. Important for our discussion is the fact that both "love the girl" and "go" are Vp's, i.e. dominated by Vp, as the following tree diagrams show:



Thus the contextual environment relevant to the selectional restrictions of verbs of the above kind is fixed by the string of categorial symbols (which includes V(erb) also) internal to Vp, i.e. the string dominated by Vp immediately. Such selectional restrictions are said to strictly subcategorize lexical cateogories (in the above case the lexical category v) and the rules treating these restrictions are called strict subcategorization rules. The subcategorization is strict in the sense that the contextual environment relevant to the subcategorization is strictly internal to an immediately higher level node that dominates it. On the basis of the preceding observations, we can impose a general condition on strict subcategorization rules to the effect that any such rule should be of the form X→ CS/A___ B, where AXB is a Y and furthermore Y dominates AXB immediately.[17] Technically any rule that relies on the structure of a string, i.e. the derivational history of the string is transformational by definition,[18] and all strict subcategorization rules are transformational in this sense. Since the structure of the string that is relevant for a strict subcategorization rule is confined strictly to one particular node in a phrase marker, it is said to be a strictly local transformational rule.

Second let us consider what conditions should be placed on the rule schema for the other kind of selectional restrictions. As noted already, these selectional restrictions are governed not by the category symbols in the contexts themselves but by certain constituent features in the heads of grammatically related phrases. As opposed to the strict subcategorization rules of the above kind, the rules covering this kind of selectional restrictions are called selectional rules. The selectional rules differ from the strict subcategorization rules in that the subcategorization of a lexical item in the case of the former rules is determined by the feature or features of the heads of grammatically related phrases such as the choice of the verb "read" only when the contextual environment contains the feature "human" in the head of the subject Np and the feature "legible" in the head of the object Np, which makes the environment not strictly local, while the environment for the latter type of rules is strictly local in the sense defined. Note that there must be a considerable amount of elements that intervene between the lexical category being subcategorized by selectional rules and the heads of grammatically related phrases that determine the subcategorization as in "The *girl* might have been *read*ing a *book*." Generalizing on these observations, we drop the requirement that

[17] CS (for Complex Symbol) refers to any subcategory of X determined by the context A___B, such as X1, X2, ···, Xn determined by A1___B1, A2___B2, ···, An___Bn.

[18] For the definition of a transformational rule, see 10.

the environment be strictly local and propose the rule schema $X \rightarrow CS / A\alpha \_\_ \beta B$, where A and B refer to the feature sets in terms of which X is subcategorized into CS and the symbols α and β refer to anything intervening between X and A, and X and B respectively. The selectional rules are transformational also but not strictly local. Note that $X \rightarrow CS$ should be given in the grammar only after proper contexts for it is provided by previons rules.[19]

5. THE ORGANIZATION OF THE BASE

We have seen that there are roughly three types of rules that should be incorporated in the base: 1) categorial rules often called phrase structure or rewrite rules, 2) selectional rules, and 3) "rules" for the idiosycracies of words, which we will call lexical entries in later discussions.[20]

As a general rule, lexical entries typically subcategorize the semantic and phonological structures of a word into simultaneous bundles of semantic and phonological features, cross-clasifying words semantically and phonologically. This cross-classifying nature of lexical entries makes them entirely different from the rewriting nature of simple categorial rules, which are designed for a hierarchical derivation of structures. Selectional rules also subcategorize lexical items not in terms of their inherent semantic and phonological structures but in terms of contextual environment. Lexical entries and selectional rules have the same characteristic of involving the cross-classification (subcategorization) of lexical items, their only difference being that in the former the cross-classification is determined by the inherent features of words while in the latter it is determined by contextual features. As we have discussed, selectional rules differ from categorial rules in that the former is transformational while the latter is non-transformational.[21]

To summarize, there are significant differences in essential characteristics between the categorial rules on the one hand and the selectional rules and the lexical entries on the other. The intermixture of these widely different types of rules in the base would complicate the base conceptually. Thus we subdivide the base into two components: the categorial subpart and the lexicon, the former comprising the categorial rules only and the latter comprising the lexical entries where selectional restrictions are also specified according to the schemas discussed in 4.7.

---

[19] For a discussion of this topic, see Chomsky's *Aspects of the Theory of Syntax*.

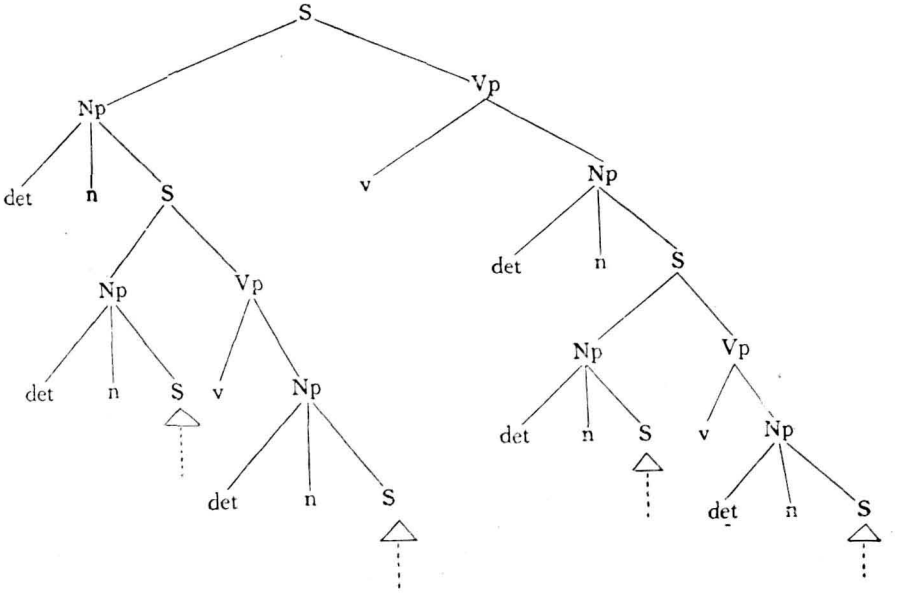
[20] Selectional rule is a cover term for both selectional and strict subcategorization rules.

[21] Categorial rules do not require information about derivational history and therefore are non-transformational in the technical sense of the word.

Set up in this way, the base will provide the information of the sort summarized in 3.6. by having the categorial subpart that serves to determine hierarchical structures of phrase markers in terms of which grammatical relations and underlying ordering are determinable by the conventions discussed, and the lexicon that determines the selectional restrictions and other inherent idiosyncracies of words including the meaning and pronunciation.

6. THE RECURSIVE PROPERTY OF GRAMMAR

It is important to realize that the sentences of a language are infinite in number, this infiniteness being demonstrable, among other things, by the fact that there is no longest sentence. We account for this fact by allowing the node #S#, which is the initial symbol of a base derivation, to recur in certain specified positions such as in the string det+noun+S dominated by an Np within the configuration of an underlying phrase marker. Let us demonstrate how the recursive property of grammar given by recurrence of #S# in a phrase marker can account for the fact that there is no longest sentence. Given the set of base rules: 1) S→Np+Vp, Vp→v+Np, Np→ det+n+S, we can derive a string as long as we want it to be, as the following diagram shows, where vertical dotted lines indicate that the derivations in those positions can be infinitely long:[22]



[22] Where the rule Np→ det+n+S underlies such expressions as "the man who went to New York," etc; any internal occurrence of #S# in a phrase marker starts a new base derivation beginning with the initial symbol #S#.

After no more #S# is chosen in the base derivation like the one above and applying all the categorial rules, we get a string called "preterminal string." Every symbol in the preterminal string is expanded into a dummy. Into these dummies are inserted words from the lexicon according to substitution transformations to be discussed below. After the insertion of all the words into the dummy positions, the phrase marker becomes the underlying phrase marker or generalized phrase marker, which contains all information that the base is supposed to provide. The final line of such an underlying phrase marker is the base terminal string discussed briefly in 4.1.

7. INSERTION OF WORDS IN THE BASE TERMINAL STRING

The preterminal string as defined above is a line of category symbols, grammatical and lexical, with their constituent structure in the phrase marker indicated. Every lexical entry in the lexicon has the contextual environment and the part of speech affiliations represented by selectional restrictions and syntactic marker specifications respectively. According to this information associated with each lexical entry, we insert a word, i.e, a lexical entry, into a dummy position which is dominated by each category symbol of the preterminal string, if the selectional restrictions and the part of speech affiliations are satisfied by the structure of the preterminal string. Technically, the rule for the insertion of words into dummy positions is transformational because it relies on the structure of the string. Furthermore, it is a substitution transformation since a word is substituted for a dummy according to the structure of the string, of which the dummy is a part.[23]

8. THE INADEQUACIES OF PHRASE STRUCTURE GRAMMAR FOR LINGUISTIC DESCRIPTION

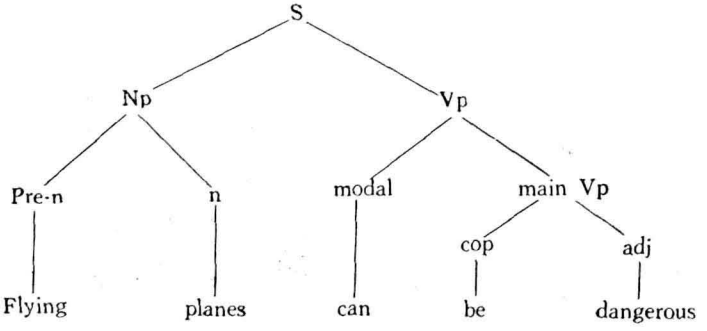
Most linguistic models, which we called taxonomic or descriptive linguistic models, assume either implicitly or explicitly that the system of categorial rules and superficial structure can describe the entire linguistic structure of a language "adequately." Any such model, i.e. grammar, is technically called a phrase structure grammar. Paul M. Postal demonstrates that all taxonomic models of linguistic description developed in the United States after World War II are versions of such a phrase structure grammar and, furthermore, all these models are inadequate in two general ways: 1) they fail to assign correct structural descriptions, i.e. fail in strong generative capacity, and 2) they fail to enumerate all the sentences of a

[23] This way of inserting words is called the substitution format; for an alternative way, called the matching format, see Noam Chomsky, *Aspects of the Theory of Syntax*, pp. 121—122.

language, even when the important requirement that a grammar assign correct structural descriptions to the sentences it generates is dropped, i.e. fail in weak generative capacity.[24] In other words, they fail in both descriptive and observational adequacy, which renders them of little linguistic significance if linguistics is to be more than a mere collection and observation of limited data.

Instead of presenting arguments about the inadequacies of those models, which is beyond the scope of the present paper, we will simply mention some of the serious inadequacies cited by Postal, which concern strong generative capacity only:

1) Phrase structure grammar assigns only one phrase marker, which corresponds to our superficial phrase marker, to most sentences, thus failing to make the empirically well-motivated distinction between underlying and superficial structures or phrase markers of sentences. For example, the IC analysis of the sentence "Flying planes can be dangerous" would roughly take the form of the following one phrase marker.[25]



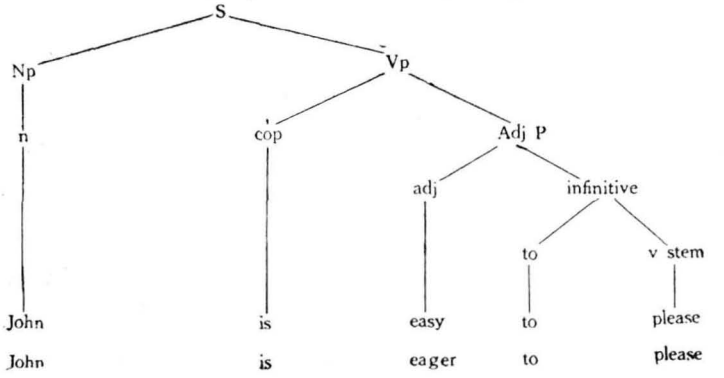
This IC analysis is clearly wrong because it does not indicate the two different syntactic roles of "Flying" here as 1) the noun modifier derived from the expression "The planes that fly," and 2) the nominalized verb derived from "△ fly planes," both derivations being very general syntactic processes in English.[26]

2) Phrase structure grammar assigns identical phrase markers to sentences, which manifest

[24] For this demonstration, see Paul M. Postal, *Constituent Structure: A Study of Contemporary Models of Syntactic Description* (The Hague: Mouton & Co., 1964).

[25] As already mentioned, the IC analysts do not use labels for constituents; I label the constituents here partly because I want to highlight the fact that labeling shows the inadequacy of IC structural descriptions clearly in such cases as when we use the same label "pre-noun" for the two different syntactic roles of "Flying" in the samples IC analysis; although I use only the IC structural descriptions of the sentences here, other taxonomic models can be shown to have the same defects as IC in this regard.

[26] △ refers to an unspecified Np such as "someone"; the need to have an unspecifed Np will be discussed in connection with the discussion of the unique recoverability condition on transformations in 2.11.



different underlying grammatical relations, mainly due to its failure to distinguish underlying from superficial phrase markers. Thus the IC analysis would assign the same structure to "John is easy to please" and "John is eager to please" as shown in the following tree diagram:

That this IC analysis is wrong has already been demonstrated in our preceding discussions.

3) Phrase structure grammar cannot reconstruct intuively felt grammatical relations such as those holding among sentences in a paraphrase set, because it assigns only one phrase marker per sentence. Thus, it would assign different structures (i.e. phrase markers) to such clearly related sentences as "John is easy to please," "It is easy to please John," "To please John is easy," etc.

4) Phrase structure grammar fails to meet the reasonable requirement of simplicity not to mention adequacy in the description of selectional restrictions. This is largely because selectional rules involve, as we have seen, cross-classification, where a category is cross-classified according to the complicated intertwining contextual environments, which cannot be adequately handled by categorial rules, which are designed to take care of simple branching structures only.[27]

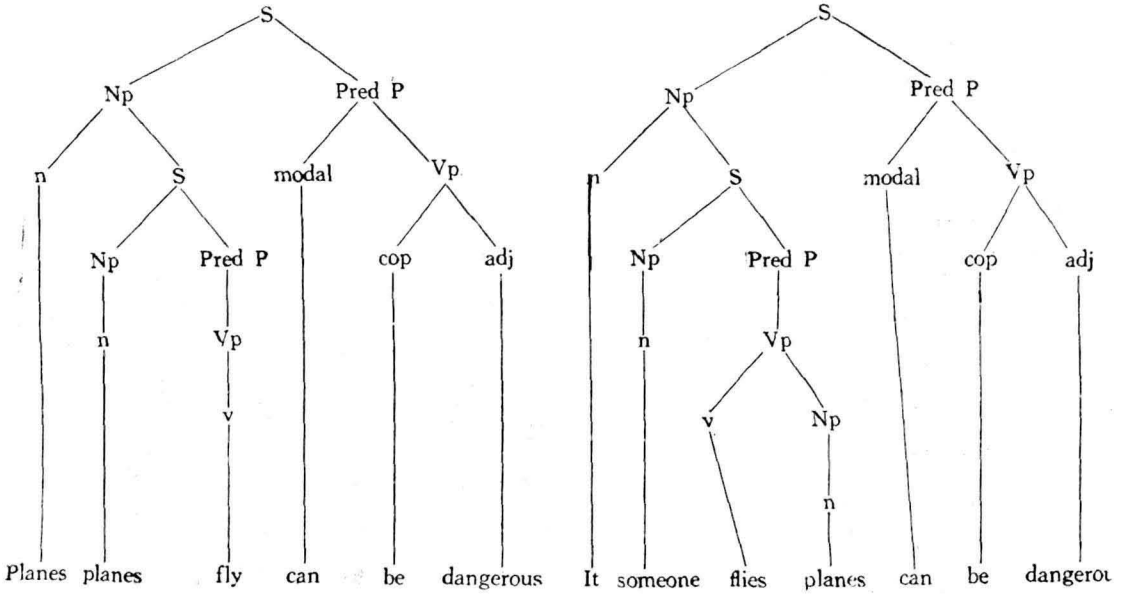
Discussion of these inadequacies is important because 1) these inadequacies have been the major motivation behind the development of the transformational model of linguistic description, and 2) the comparison of taxonomic models, where these inadequacies are inherent, and the transformational model, which overcomes all of the inadequacies, demonstrates the higher level of adequacy achieved by the transformational model.

Inadequacy 4) above can be neatly overcome by the rule schema introduced to formalize selectional restrictions (4.7.) and transformational rules (9). All the inadequcies derive from

[27] For a detailed discussion of the cross-classifying nature of selectional restrictions and lexical meanings versus the hierarchic branching nature of categorial rules, see Noam Chomsky, *Aspects of the Theory of the Theory of Syntax*, pp. 79—106.

the failure on the part of taxonomic models to make the distinction between underlying and superficial structures udner the mistaken assumption that a surface structure of a sentence can show all the structure there is to a sentence. These inadequacies are overcome in transformational grammar by 1) making the superficial and underlying structure distinction, and 2) by treating underlying structures by the base rules, which are so designed as to provide the elementary contents of sentences, i.e. their meanings, and treating superficial structures in the transformational component by means of transformational rules, which treat of various superficial forms of sentences by deriving them from underlying structures. By introducing transformational rules, the details of which will be discussed in the sections to follow in this paper, the transformational model solves most of the problems of phrase structure grammar.

Let us now see how transformational grammar would overcome inadequacy 1) using the sample sentence given, i.e. "Flying planes can be dangerous."[28] The base (sub-)component would assign two different underlying phrase markers to the sentence, which would take roughly the form of the following tree diagrams:[29]



To the first underlying phrase marker would apply a transformation called relative

[28] I omit the discussion of how TG would overcome inadequacies 2) and 3) because our previous discussions of similar cases already show how it is to be done.

[29] As throughout the paper, the phrase markers given are not supposed to be precise; they are all simplified for the sake of convenience.

transformation, which would result in a structure of the form "Planes+that+fly+can+be+dangerous." To this structure resulting from the underlying phrase marker would apply another transformation that suffixes -ing to the verb "fly" and prepose it to "planes" deleting "that," which results in the structure "Flying+planes+can+be+dangerous." To the second underlying phrase marker would apply the so-called -ing nominalization, which will delete "it" simultaneously with the suffixing of the nominalizer -ing to the verb "fly", which will result in the structure Flying+planes+can+be+dangerous.[30]

By recognizing the two different structures in this sentence and by deriving the identical superficial structure from the two different underlying structures, we can clearly bring out the two different syntactic roles of the superficially identical flying, which have to be shown if we are to describe correctly how the native speakers of English interpret and use such expressions as "Flying planes can be dangerous" consistently.

9. THE ROLE OF THE TRANSFORMATIONAL SUBCOMPONENT

Rules in the transformational subcomponent of the syntactic component are so designed as to relate or map underlying phrase markers to their various superficial phrase markers. For example, a transformational rule will map the underlying structure "it+S+is+easy," where S dominates "Mary pleases John," to its various superficial structures such as "it+is+easy+for+Mary+to+please+John," "To+please+John+is+easy+for+Mary", "John+is+easy+for+Mary+to+please," "For+Mary+to+please+John+is+easy."

As we have already discussed, selectional restrictions are already taken care of by the rules of the base so that an underlying phrase marker or structure expresses proper selectional restrictions. Since we apply transformatioal rules to such an underlying phrase marker, seemingly complex selectional restrictions in superficial structures derived by transformational rules are economically fixed in that underlying phrase marker. Suppose now that we do not distinguish underlying and superficial structures as in taxonmic models and that each of the superficial structures originating in a single underlying phrase marker such as "it+S+is+easy" is treated as unrelated to or different from the other. Then we will have to state essentially the same selectional restrictions for each of the sentences, which we know intuitively come from the same underlying source. This will result in unnecessary complications of the grammar of a language. Imagine how complicated it would be to state selectional restrictions separately for "The boy read the book" and its passive "The book

[30] The rationale for "it" as det in the phrase marker is that "to" nominalizations sometimes require it as in "It is dangerous to fly planes."

was read by the boy." Transformational rules have the virtue of taking care of the superficially complex selectional restricions in a very simple manner by applying to an underlying phrase marker fully specified with respect to selectional restrictions, these selectional restrictions being carried over unchanged to the superficial phrase markers which they derive from that underlying phrase marker. This merit of transformational rules is considered one of the rationales for having the transformational subcomponent in the syntactic component. This rationale may be added to the two rationales to be mentioned in the fourth paragraph of the present section.

Transformational rules map not only an underlying phrase marker to one or more superficial phrase markers but also a superficial phrase marker to one or more superficial phrase markers. For example, the underlying phrase marker for "The boy was hit by a car" is "A car+past+hit+the boy+by+passive," which is mapped by one transformation to the superficial structure "The boy+past+be+en+hit+by+a car," which becomes "The boy was hit by a car" after all the rules are applied to it. However, the superficial structure for "Was the boy hit by a car?" is most conveniently derived from the superficial structure "The boy+past+be+en+hit+by+a car" plus the marker Q (for question), rather than directly from "A car+past+hit+the boy+by passive" plus Q.

It is important to note that the rationale for transformational rules is 1) to mediate the pairing between meanings and pronunciations of sentences by relating meaning-bearing underlying structures to sound-bearing superficial structures, which must be distinguished for reasons discussed elsewhere[31] and 2) to reconstruct the intuitively felt formal relations between sentences such as those in a syntactic paraphrase set, etc., all of which are designed to enable the maximally simple but descriptively and/or explanatorily adequate description of the structure of a language, overcoming almost all of the inadequacies of phrase structure grammar in a significant way. To conclude this section, the role of transformational rules is to map phrase markers to phrase markers, the rationale for which is given in the above paragraph.

10. THE DEFINITION OF A GRAMMATICAL TRANSFORMATIONS

A grammatical transformation, as has been previously indicated, is a rule of grammar that appears in the syntactic component, which maps phrase markers to phrase markers. In order to distinguish it from phonological transformations in the transformational cycle

[31] See my "Some Basic Assumptions in Transformational Lingnistics."

of phonology of the phonological component and also from the substitution transformations in the base, we call it a GRAMMATICAL transformation.[32]

A grammatical transformation as such is defined by 1) its domain, i.e. the class of phrase markers to which it applies, which is called "structural description" or equivalently "analysis," 2) its effect, i.e. the changes it brings about in the structure of the resulting phrase marker, which is called "structural change," and 3) the formal operations that effect the changes from the original phrase marker to the resulting phrase marker.

10.1. The Domain of a (Grammatical) Transformation:[33] To apply transformational rules, we have to know the derivational history, i.e. the hierarchical constituent structure(s) of the string or the class of strings to which they apply. If this were not the case, given the string "The old man who was here was sick" plus the instruction to make it into a question by preposing "was" to the sentence-initial position, we would get "Was the old man who here was sick," which is clearly wrong, just as well as "Was the old man who was here sick," which is correct. Only in terms of the derivational history represented in the phrase marker can we pinpoint that the "was" to be preposed is the second one and not the first.

Furthermore, a transformational rule applies not just to one specific string but to a class of strings characterized by similar structures of phrase markers. For instance, the passivization transformation in English applies not just to the string "A car + hit + the boy + Y + by + passive" but to any string having the underlying structure or phrase marker representable as "Np + tense + vt + Np + Y + by + passive."

Thus the domain or range of the applicability of a transformation is defined by a class of phrase markers having certain specified structures. We define the domain of a transformation formally in the following manner. Suppose that we have a phrase marker X with the terminal string T, where T is segmentable into T1, T2 $\cdots$, Tn in such a way that each Ti is traceable back to a node labelled Ai in X such that T1, T2, $\cdots$, Tn is traceable back to A1, A2, $\cdots$, An in X without remainder." If such is the case, we say that the terminal string T is analyzable as (T1, T2, $\cdots$, Tn; A1, A2, $\cdots$, An) with respect to X." In terms of the notion "analyzability" as defined above, we can define the domain of a transformation GT[33] by the convention: a terminal string T is in the domain of a transformation GT if T

[32] For a detailed discussion of the definition of a grammatical transformation, see Noam Chomsky, "The Notion Rule of Grammar," in J.A. Fodor and J.J. Katz (eds.), *The Structure of Language* (Englewood Cliffs, N.J.: Prentice Hall, Inc.; 1964).

[33] Grammatical transformation may be referred to as transformation without qualification. GT is the usual abbreviation for grammatical transformation.

is analyzable as (T1, T2, ⋯, Tn; A1, A2⋯, An) with respect to the phrase marker X, where (T1, T2, ⋯, Tn) is called a proper analysis of T and (A1, A2, ⋯, An), the structural description of T.[34]

10.2. The Formal Operations that Effect Changes in the Phrase Markers: A passive transformation applied to any string with the structural description "Np1+tense+vt+Np2+Y+by+passive" will change the phrase marker to a new one with the structural description "Np2+tense+be+en+vt+Y+by+Np1."

The changes effected from the underlying to the superficial phrase marker by the passive transformation are: 1) the substitution of Np1 and Np2, 2) the erasure of passive, 3) the adjunction of be+en, etc. These formal operations that underlie a grammatical transformation are called elementary transformations. We do not want these elementary transformations to change phrase markers arbitrarily. For instance, we do not want them to effect a reflection in the resulting phrase marker such as A1, ⋯, An→ An,⋯, A1, because in no natural language this kind of change is allowed, e.g., "Killed was boy the," which is a reflection of the normal English sentence "The boy was killed" would be rejected as totally meaningless. For some more discussion on the conditions that must be imposed on elementary transformations, see 11. in this paper.

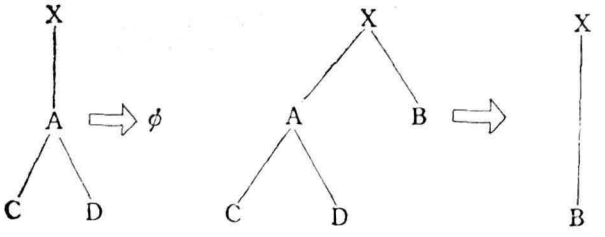
10.3. The Structural Change: We have discussed that a transformation applies to a class of phrase markers with the same structural description, i.e. the same constituent structure defined in terms of the notion "analyzability." Furthermore, we have argued that a transformational rule might apply to a phrase marker that has resulted from a previous application of a transformational rule (see 9). Then we should require that the result of applying a transformation to a phrase marker be a new phrase marker with the new constituent structure explicitly specified if further transformations are to apply to the result.

We refer to the phrase marker that results from the application of a transformational rule as "a derived phrase marker or structure." If no more transformational rules are applied to this derived phrase marker (which is passed onto the phonological component for phonological interpretation), we call it "the final derived phrase marker." The final derived phrase marker is what we call the superificial structure or phrase marker of a sentence.

There are some distinct conventions associated with such elementary transformations as

[34] The formal definition of the domain of a transformation given here is adapted from Noam Chomsky, "The Notion Rule of Grammar," in J.A. Fodor and J.J. Katz (eds.), *The Structure of Language*.

erasure, substitution, and adjunction, which govern the assignment of derived phrase markers to the result of applying transformations. For one, the convention associated with the erasure elementary transformation requires that if a node A is to be erased whatever A dominates be also erased and, furthermore, if A is dominated by a higher leverl symbol X without remainder then X be also erased so that there will no longer be any constituent nodes X, A and whatever is dominated by them. The following diagrams are designed to show this:



Several conventions associated with the other elementary operations such as the one given in the preceding paragraph will "automatically" assign derived phrase markers to the phrase markers that result from the application of a transformational rule to the original phrase marker.[35]

11. SOME CONDITIONS ON TRANSFORMATIONS

In this section, we will confine ourselves to transformations having as one of its elementary transformations an erasure or equivalently deletion transformation and some constraints that must be imposed on them. It is an empirical fact of English and other languages[36] that deletion is allowed only when the deleted element is uniquely recoverable. For instance, the "you" of imperatives is deletable because it is recoverable uniquely after the deletion, but the subjects of other sentences are not deletable because there is no non-arbitrary way of recovering them once they are deleted.

The unique recoverability requirement on deletion operations seems to be extremely well-motivated, because it accounts for some important facts of language. Note that there are

[35] It seems that the derived structure assignment algorithm is far from automatic with many problems still unsolved; the MITer group is working on a possible mechanical algorithm called "mother-daughter" system. I owe this information to Professor Robert P. Stockwell through personal communication.

[36] For a similar condition on the deletion operation in a non-English language, see my M.A. thesis "A Fragmentary Syntactic Component of Korean," University of Hawaii, 1968.

three conditions on deletion to satisfy the recoverability requirement and also due to other facts of language: deletion is allowed only when 1) the deleted element is explicity specified in the structural description such as the "you" and "will" of imperatives, 2) the deleted element is non-distinct from some other element in the phrase marker so that even after its deletion it is uniquely recoverable, and 3) the pro-form, i.e. the universal representative, of a category such as "someone" for human nouns, etc., is the deleted element. Condition one helps reconstruct the fact that native speakers of English associate the so-called understood subject and tense of imperatives with "you" and "will" respectively. Condition two helps to reconstruct the fact that native speakers of English would reject the string "the boy+the boys are bright" as part of the source phrase marker from which the relative transformation derives the string "I met the boy who is bright" or "I met the bright boy," while accepting as part of the source phrase marker the string "the boy+the boy is bright." Condition three helps reconstruct the fact that native speakers of English understand the subject of "The missing officer was found dead in his car" to be some unspecified person such as someone rather than a specific person such as John, the housewife, etc.[37]

Aside from helping an adequate reconstruction of native speakers' intuition about sentences, the conditions on transformations as well as other rules of grammar tighten the formal properties of grammar so that the evaluation measure can operate without too much difficulty in choosing the most valued of the several alternative grammars of a language.

12. ORDERING OF RULES IN THE SYNTACTIC COMPONENT

First let us consider the ordering of the rules in the categorial subpart of the base. As an illustration, suppose we have three rules: 1) S→Np+Vp, 2) Vp→v+Np, and 3) Np→det+n. If given in this order, rule 3 applies only once to any constituent that is labelled Np; if this order is violated and rule 3 precedes rule 2, for instance, then when we reach the Np in Vp we have to come back to the rule that rewrites Np as det+n, which will complicate grammatical descriptions considerably. Furthermore, an ordering of categorial rules such as is given in the second line of this paragraph seems to capture the underlying process of sentence construction in a natural manner. On the basis of these observations, which seem to be valid for all languages thus far investigated, we require that the rules in the categorial subpart be ordered linearly. Then let us consider the ordering of rules in the lexicon. As we have discussed, every word has assoicated with it selectional restrictions

[37] For a detailed discussion of the unique recoverability requirement in general, see Noam Chomsky, *Aspects of the Theory of Syntax*, pp. 137–147.

and its part of speech affiliations, which serve as the domain of a substitution transformation that inserts the word into a dummy position in the underlying phrase marker. Since any word can be inserted properly into the phrase marker if it is in the domain of the substitution transformation, the lexicon may be an unordered list of words specified in the manner suggested.[38] To recapitulate, the base rules are linearly ordered in the categorial subpart and unordered in the lexicon. Where we have occurrences of #S# inside a phrase marker the set of base rules are ordered cyclically in the sense that the same set of base rules apply to each #S# in the phrase marker.

Now consider the ordering of transformational rules. Since we have given no discussion of the transformational component of English or any other language in this paper, it is difficult to argue for or against a certain proposal concerning the ordering of transformational rules. However, since we have argued that a transformation may apply to the result of the application of some previous grammatical transformation(s) (see the second paragraph of 9.) and furthermore the optimal grammatical description of language require that some transformation(s) apply to a string only after some other transformation(s) have applied (e.g. the necessity to apply a passivization transformation to a string so that the question transformation can derive the passive question from the passive declarative rather than from the active declarative directly), the transformational rules must also be linearly ordered, such as in the order: 1) passivization T, 2) Question T, etc. As a matter of fact, the transformational component in a grammar consists of just a series of linearly ordered grammatical transformations.[39] The preceding remarks cover the ordering of transformations for a base derivation in which no recursive #S# has been inserted. However, a base derivation may contain any number of recursive #S#'s, in which case transformational rules from the same series of linearly ordered transformations apply first to the most deeply embedded #S#, then to the next most deeply embedded #S#, and so on until the highest sentence #S# constituent node is reached. In other words, the linearly ordered transformations apply cyclically from bottom up.

13. SYNTACTIC WELL-FORMEDNESS VERSUS ILL-FORMEDNESS

Violations of the rules of the syntactic component result in widely varying degrees of

[38] Recall that a substitution transformation inserts a word into a dummy in the phrase marker when the lexical entry of the word satisfies the contextual restrictions of the dummy in the phrase marker.

[39] For a discussion of the composition of the transformational component and especially the ordering of rules there, see C.J. Fillmore, "The Position of embedding transformations in a Grammar," *Word*, vol. 19, no. 2, 1963.

deviance from syntactic well-formedness in the resulting sentences; strict adherence to all the syntactic rules Produces syntactically well-formed sentences. In this way, transformational syntax reconstructs native speakers' ability to accept certain sentences as perfectly grammatical, slightly ungrammatical, totally ungrammatical, etc. This is a natural consequence of our conception of grammar as a device to generate all and only the sentences of a language, where the qualification "only" excludes ungrammatical sentences from being generated by the grammar. The significant thing here is that the notion "ungrammatical" and/or the degrees of deviance can be defined in a formal manner as cases of violations of certain types of rules.

As an illustration, let us consider the origin and possibly degrees of ill-formedness of the following sentences: 1) Sincerity may virtue John, 2) Sincerity may surprise, 3) Sincerity may hate John, 4) Sincerity may hates John, etc. Sentences 1, 2, 3, 4, are the results of the violation of four different types of rules, i.e. 1) categorial rule, 2) strict subcategorization rule, 3) selectional subcategorization rule, 4) transformational rule (that blocks the addition of tense-marking -s by the presence of the modal "may"), respectively. Some tentative proposals have been made to relate the degrees of deviance from well-formedness or equivalently grammaticality to the differences in the types of rules violated.[40] For example, the violation of a selectional subcategorization rule such as in "Sincerity may hate John" results in a less ill-formed sentence than the violation of a strict subcategorization rule such as in "Sincerity may surprise," etc. It is interesting to observe that not all ill-formed sentences are "bad" for communicative purposes. Reduced forms of sentences such as "Yes", "Did you?" are ill-formed semi-sentences, but they are effective means of communication. In poetry, metaphorical effect is often gained by violating some selectional (subcategorization) rules as in Chomsky's classical "Colorless green ideas sleep furiously."

Suppose now that all base rules have applied to produce an underlying or equivalently generalized phrase marker. Since we have allowed the recursive #S# to recur inside a base derivation, base rules sometimes cannot take account of some relevant contexts. Recall that the recoverability requirement on deletion has to block a transformational rule from mapping the source string the boy#the man is sick#to "the boy who is sick," but the base rules as given in the earlier part of this paper are such that it generates "the boy#the man is sick" just as as well as the desired "the boy #the boy is sick#." This means that

[40] For discussions of the degrees of grammaticalness within this framework, see Noam Chomsky, "Degrees of Grammaticalness", Paul Ziff, "On Understanding 'Understanding Utterances'", Jerold J. Katz, "Semi-sentences", in J.A. Fodor and J.J. Katz (eds), *The Structure of Language*.

not all generalized phrase markers can underlie well-formed sentences due to the inability of base rules to take account of certain centexts and that only those underlying phrase markers which satisfy the unique recoverability requirement, in this case the nondistinctness condition (see 9.), underlie well-formed surface structures. Non-distinctness condition on this particular type of erasure transformation serves a filtering function that allows only certain underlying phrase markers to be mapped onto well-formed surface structures while barring others. We can establish a convention to the effect that when underlying phrase markers satisfy the non-distinctness condition and are mapped onto superficial phrase markers, the sentence boundary markers # # are erased but otherwise not as in the "The boy # the boy is sick # is bright"→ "The boy who is sick is bright" but "The boy # the boys are sick # is bright "→ "The boy # the boys are sick # is bright". Thus after all the transformational rules have applied, the presence of any internal occurrence of # would indicate that the superfical structure is not well-formed. Note that # marks for other syntactic boundaries such as word boundaries are not erased because they serve a very useful purpose for the operation of the rules in the phonological component.

To summarize, a syntactically well-formed sentence results only when we have applied 1) categorial rules, 2) selectional rules of both kinds, and 3) transformational rules without violating conditions on their application; differing degrees of syntactic ill-formedness results from the violation of some or all of the different types of syntactic rules.

Selected Bibliography

Chomsky, Noam. *Aspects of the Theory of Syntax*. Cambridge, Mass: MIT Press, 1965.

Chomsky, Noam and Morris Halle. *The Sound Pattern of English*. New York: Harper & Row, 1968.

Fodor, J. A. and J. J. Katz (eds.). *The Structure of Language: Readings in the Philosophy of Language*. Englewood Cliffs, N. J., 1964.

Gleason, Henry A. *Interoduction to Descriptive Linguistics*. New York: Holt, Rinehart & Winston, 1961.

Katz, J. J. *The Philosophy of Language*. New York: Harper and Row, 1966.

Longacre, Robert E. *Grammar Discovery Procedure*. The Hague: Mouton & Co., 1964.

Park, Nahm Sheik. "Some Basic Assumptions in Transformational Linguistics" in *Language Research*, Vol. Ⅳ. No. 1. Available from the Microfilm Library of the Center for Applied Linguistics, Washington, D. C.

Postal, Paul M. *Constituent Structure: A Study of Contemporary Models of Linguistic Descriptions.* The Hague: Mouton & Co,. 1964.

Wells, Rulon S. "Immeidate Constituents" in Martin Joos (ed.) *Readings in Linguistics.* New York: American Council of Learned Societies, 1963.