INFORMATION AND CONTROL 30, 307–379 (1976)

A Topological Study of Phrase-Structure Languages*

S.-Y. KURODA

Department of Linguistics, University of California, San Diego,
La Jolla, California 92037

It is proposed that structural equivalence of phrase-structure languages be defined by means of introducing, for each such language, a class of topological structures on the language. More specifically, given a phrase-structure language (either as a set of trees or as a set of strings), we introduce a class of topological spaces associated with finite sets of "phrases." A function from one language to another, where both are equipped with such classes of topological spaces, is said to be structurally continuous, if for any topological space belonging to the first, there is a space belonging to the second such that the function is continuous with respect to these spaces. Then phrase-structure languages, or grammars that generate such languages, may be classified into structurally homeomorphic types in the obvious way. Two different methods of topologizing phrase-structure languages (one dependent on the other) are considered, and it is shown that for the class of context-free languages, one method provides a finer classification of languages (or grammars) than the other. In Part 2 we apply the general theory to a particular subclass of context-free languages, the class of tree language counterparts of regular languages.

INTRODUCTION

Chomsky has defined two distinct notions of equivalence of grammars. Two grammars are said to be weakly equivalent if they generate the same set of strings; two grammars are said to be strongly equivalent if they are weakly equivalent and if they associate the same structural description (or set of structural descriptions, in the case of ambiguity) to each string they generate. (Structural descriptions are given formally as labeled trees; thus grammars are strongly equivalent if they generate the same set of labeled trees.)

* Part of the present work was done at the Laboratoire d'Automatique Documentaire et Linguistique, Centre National de la Recherche Scientifique and at the Research Laboratories of Electronics, Massachusetts Institute of Technology. The results of the work were previously reported in Kuroda (1972, 1973). In fact, the introduction to this work is a adapted translation of the latter.

It is easy to see that if we limit ourselves to context-free grammars, the notion of strong equivalence is trivial. Two context-free grammars are strongly equivalent only if they are identical except for inessential details, such as the existence, in one or the other of the grammars, of rules which are not used in the derivation of terminal strings. This fact, however, does not deprive the notion of strong equivalence of significance, since formal grammars, in general, can be strongly equivalent without being essentially identical. Even in the framework of context-free grammars, there can be essentially distinct grammars that are strongly equivalent, if rules are ordered (as is sometimes done by linguists).

But let us limit ourselves for the moment to context-free grammars, in the usual sense without rule ordering. We might ask if there are ways of classifying weakly equivalent grammars according to the similarity (rather than identity) of the trees they associate with the strings generated. Consider the following grammars:

$$
\begin{aligned}
G_1\colon\ & S \rightarrow PQ \\
& P \rightarrow AP \\
& P \rightarrow a \\
& Q \rightarrow QB \\
& Q \rightarrow b \\
& A \rightarrow a \\
& B \rightarrow b
\end{aligned}
\tag{1}
$$

$$
\begin{aligned}
G_2\colon\ & S \rightarrow P \\
& P \rightarrow AP \\
& P \rightarrow AQ \\
& Q \rightarrow QB \\
& Q \rightarrow b \\
& A \rightarrow a \\
& B \rightarrow b
\end{aligned}
\tag{2}
$$

These two grammars are weakly equivalent; they both generate the set of strings $\{a^m b^n;\ m \geqslant 1, n \geqslant 1\}$. They are not strongly equivalent; the string a^3b^2, for example, is associated by G_1 and G_2 with trees (3a) and (3b), respectively:

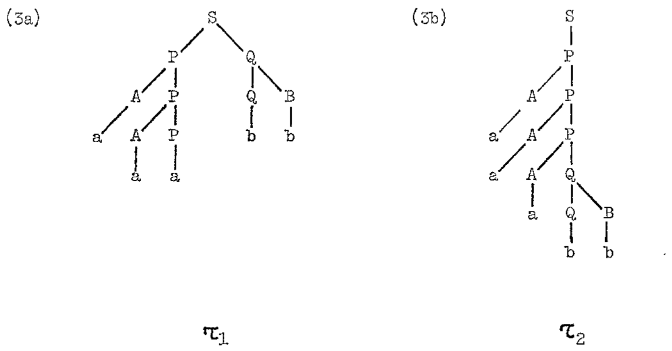


Figure 1

These two trees are obviously quite different. Let us modify the two grammars slightly, as follows:

$$
\begin{aligned}
G_1'\colon\quad & S \to PQ' \\
& Q' \to Q \\
& P \to AP \\
& P \to a \\
& Q \to QB \\
& Q \to b \\
& A \to a \\
& B \to a
\end{aligned}
\tag{4}
$$

$$
\begin{aligned}
G_2'\colon\quad & S \to P \\
& S \to P' \\
& P \to AP \\
& P \to AP' \\
& P' \to AQ \\
& Q \to QB \\
& Q \to b \\
& A \to a \\
& B \to b
\end{aligned}
\tag{5}
$$

G_1' and G_2' associate with a^3b^2 the trees (6a) and (6b), respectively:

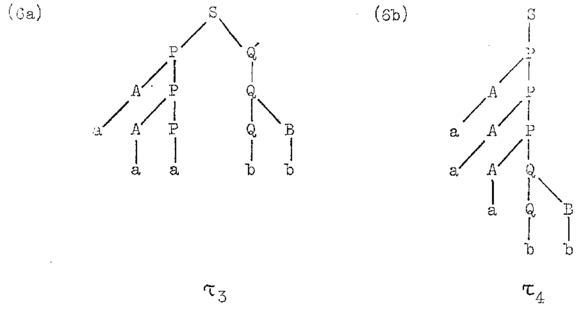


FIGURE 2

It seems reasonable to say that trees τ_1 and τ_3 are similar, as are trees τ_2 and τ_4, while neither τ_2 nor τ_4 is similar to either τ_1 or τ_3. We would like to say, then, that G_1 is similar to G_1' but neither to G_2 nor G_2'. The notion of strong equivalence, however, states simply that none of the four grammars is equivalent to any other. I would like to explore one method whereby a notion of structural similarity less strict and more significant than the notion of strong equivalence can be defined.

This method will involve the introduction of topological structures into constituent structure languages. In this introduction I will include a brief sketch of some of the concepts discussed in Part I of this paper.

Let V and W be, respectively, sets of nonterminal and terminal symbols. A V-tree is a tree whose nodes are elements of V; a W-leaved V-tree is a tree whose nonterminal nodes are elements of V and whose terminal nodes are elements of W. A tree language is a set of W-leaved V-trees. Given a context-free grammar G with nonterminal and terminal vocabularies V and W, respectively, the set of S-rooted W-leaved trees generated by G (where S is the initial symbol of G) is a tree language which we denote by $K(G)$. Note that, in general, a tree language isn't necessarily generated by a grammar. Given a W-leaved V-tree τ, we call the sequence of elements of W that consists of the leaves of τ in their left-to-right order the string associated with τ; this string is denoted by $/\tau/$. The set of strings associated with the trees of a tree language K is called the string language associated with K; we denote it by $L(K)$. (When K is generated by a grammar G, we can write $L(G)$ instead of $L(K(G))$, conforming to the notation of the algebraic theory of grammar, which denotes the string language generated by G as $L(G)$.)

Let K be a language of W-leaved V-trees. For each element τ of K, let (τ)

be the V-tree obtained from τ by eliminating all its leaves. For example, for τ_1 and τ_2 given above, (τ_1) and (τ_2) are, respectively:

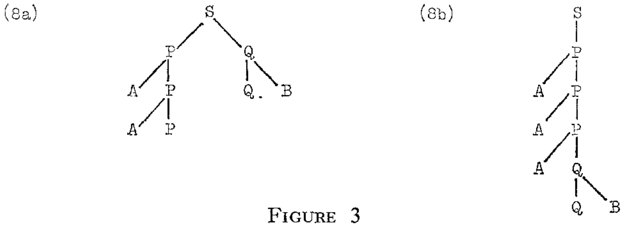


FIGURE 3

Let us define a binary relation $<_0$ on K as follows: $\tau <_0 \sigma$ if and only if (τ) is a subtree of (σ). The relation $<_0$ is a preorder on K, and canonically introduces a topology in K, the topology whose class of open sets is generated by sets $V(\tau) = \{\sigma; \tau <_0 \sigma\}$. We denote this topology by $\mathbf{T}_0(K)$.

Since $\mathbf{T}_0(K)$ represents nothing other than a structure of preorder, it is at this stage superfluous to use the language of topology. But, for one thing, it seems to be convenient to use the language of topology when we eventually deal with a class of preorders defined on K, or when we induce a structure on the string language associated with K from the structure defined on K. For another thing, we shall later deal with a structure involving topologies defined on the basis of these preorders, but not reducible to preorders. We can obtain uniformity in our overall exposition by using the topological language from the outset.

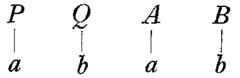
Let us consider the languages $K_1 = K(G_1)$ and $K_2 = K(G_2)$ defined earlier. Denote by $\tau_1^{m,n}$ the element of K_1 such that $|\tau_1^{m,n}| = a^m b^n$, and by $\tau_2^{m,n}$ the element of K_2 such that $|\tau_2^{m,n}| = a^m b^n$. Thus, we have

$$V(\tau_1^{m,n}) = \{\tau_1^{x,y}; x \geqslant m, y \geqslant n\}, \qquad \text{and} \qquad V(\tau_2^{m,n}) = \{\tau_2^{m,y}; y \geqslant n\}.$$

Let f be the one-to-one function from K_1 onto K_2 defined by $f(\tau_1^{m,n}) = \tau_2^{m,n}$. Then, $f(V(\tau_1^{m,n})) \supset V(\tau_2^{m,n})$ but $f(V(\tau_1^{m,n})) \neq V(\tau_2^{m,n})$, that is, the images under f of the smallest neighborhoods of elements of K_1 properly contain the smallest neighborhoods of the corresponding elements of K_2. Thus, f is not continuous, although its inverse is; K_1 and K_2 supplied with the topologies given above are not homeomorphic. It is easy to see, however, that the one-to-one function g from K_1 onto $K_1' = K(G_1')$ defined by $|g(\tau)| = |\tau'|$ is a homeomorphism. K_2 and $K_2' = K(G_2')$ are also homeomorphic by the similarly defined function. Thus, the four weakly equivalent grammars are separated into two classes. In general we can divide each class of weakly equivalent grammars into subclasses of "homeomorphic" grammars.

This method, however, is not sufficient to introduce an interesting classification on grammars that reflects their structural similarity. Suppose we replace, in grammar G_1, the rules $P \rightarrow a$ and $Q \rightarrow b$ by the rules $P \rightarrow A$ and $Q \rightarrow B$. The new grammar, $\bar{G}_1$, is clearly weakly equivalent to G_1. On the other hand, the topology introduced on $\bar{K}_1 = K(\bar{G}_1)$ is easily seen to be the discrete topology. In the same way we can change G_2 to the "discrete" grammar $\bar{G}_2$ by replacing the rules $Q \rightarrow b$ with $Q \rightarrow B$. Since $K(\bar{G}_1)$ and $K(\bar{G}_2)$ are both discrete, they are homeomorphic. The structures given by $\bar{G}_1$ and $\bar{G}_2$ are, however, as different from each other as the structures given by G_1 and G_2.

To obtain a more satisfactory classification, we will introduce a class of topologies, rather than a single topology, on each tree language. Let Π be a finite set of phrases of a tree language K. (If K is a language of W-leaved V-trees, a phrase of K is a W-leaved V-tree which is a branch of a sentence of K. A sentence of K is by definition a phrase of K.) For each element τ of K, let $(\tau)_\Pi$ be the subtree of K which is obtained from τ by pruning all the branches of τ which are in Π. (By "pruning a branch of τ" we mean "eliminating from τ all the nodes of the branch except for its root." For example, if we prune from τ_1 ((3a) above) the phrases



(9)

FIGURE 4

we obtain (τ_1) ((8a) above).) We now define a binary relation $<_\Pi$ on K as follows: $\tau <_\Pi \sigma$ if and only if $(\tau)_\Pi$ is a subtree of $(\sigma)_\Pi$. Again $<_\Pi$ is a preorder on K and trivially introduces a topology, denoted by $\mathbf{T}_\Pi(K)$, or simply $\mathbf{T}_\Pi$. If we put

$$V_\Pi(\tau) = \{\sigma; \tau <_\Pi \sigma\}, \tag{10}$$

$V_\Pi(\tau)$ is the smallest neighborhood of τ in $\mathbf{T}_\Pi(K)$.

Returning to the "discrete" grammar $\bar{G}_1$ defined above, let $\bar{\tau}_1^{3,2}$ be:

(11)

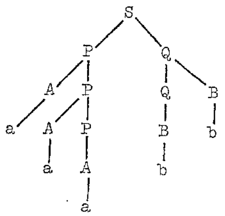
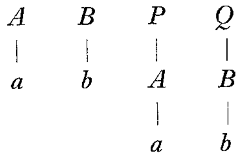


FIGURE 5

Let $\bar{\Pi}$ be the following set of phrases:



(12)

FIGURE 6

Then $(\bar{\tau}_1^{3,2})_{\bar{\Pi}}$ is:

(13)

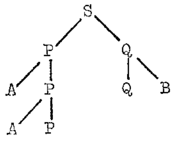


FIGURE 7

and we have

$$V_{\bar{\Pi}}(\bar{\tau}_1^{3,2}) = \{\bar{\tau}_1^{x,y};\ x \geqslant 3, y \geqslant 2\}, \tag{14}$$

where $\bar{\tau}_1^{x,y}$ is the element of $\bar{K}_1$ such that $/\bar{\tau}_1^{x,y}/ = a^x b^y$. If $\bar{f}$ is the one-to-one function from K_1 onto $\bar{K}_1$ defined by $/\bar{f}(\tau)/ = /\tau/$, then $\bar{f}$ is a homeomorphism from $\mathbf{T}_0(K_1)$ to $\mathbf{T}_{\bar{\Pi}}(\bar{K}_1)$.

Since $\mathbf{T}_0(\bar{K}_1)$ is discrete, the function $\bar{f}$ is not continuous from $\mathbf{T}_0(K_1)$ onto $\mathbf{T}_0(\bar{K}_1)$; by pruning the phrases in $\bar{\Pi}$ we obtain a weaker topology, which allows $\bar{f}$ to be continuous. It isn't difficult to see that the function f from K_1 onto K_2 can't be rendered continuous by this method of pruning several phrases.

To develop a general theory along these lines, we consider the class of topologies $\mathbf{T}_\Pi(K)$ on K, where Π is any finite set of phrases of K. Let Π_1 and Π_2 be two finite sets of phrases of K such that Π_1 is a subset of Π_2. $\mathbf{T}_{\Pi_2}$ is not necessarily weaker than $\mathbf{T}_{\Pi_1}$ (in other words, the identity function is not necessarily continuous from $\mathbf{T}_{\Pi_1}$ to $\mathbf{T}_{\Pi_2}$); for if σ and τ are two elements of K, $(\tau)_{\Pi_1} < (\sigma)_{\Pi_1}$ doesn't imply $(\tau)_{\Pi_2} < (\sigma)_{\Pi_2}$. But we can show that there exists an extension Π_3 of Π_2 such that $\mathbf{T}_{\Pi_3}$ is weaker than $\mathbf{T}_{\Pi_1}$ (Part I, Theorem 1). Thus we can say "roughly" that the topology $\mathbf{T}_\Pi$ becomes weaker as the set of phrases Π becomes larger.

Let Π_1 and Π_2 be defined as above. Although $\mathbf{T}_{\Pi_2}$ is not necessarily weaker than $\mathbf{T}_{\Pi_1}$ (that is, although for some elements τ of K, $V_{\Pi_1}(\tau)$ might not be contained in $V_{\Pi_2}(\tau)$), we will show that the elements of the former which are outside the latter must be close to τ in a certain sense. It will follow that if a certain condition is satisfied, for each Π we can define a new topology $\mathbf{T}_\Pi{}^*$ from $\mathbf{T}_\Pi$ so that for topologies of this new type, if Π_2 contains Π_1, then $\mathbf{T}^*_{\Pi_2}$ is weaker than $\mathbf{T}^*_{\Pi_1}$ (Part I, Theorem 4). To define $\mathbf{T}_\Pi{}^*$ we must first introduce the notion of partial sentence. Let us limit ourselves here to context-free grammars. Let G be such a grammar. We say that a node is fertile if it has infinite generative capacity, in the obvious sense. We say that a tree generated by G is a partial sentence if it is a sentence of G or if it contains at least one branch whose terminal nodes are all terminal symbols and whose root is fertile. (For example, tree (15) is a partial sentence of the grammar G_1 above. The branch whose root is the circled P is a branch whose terminal nodes are terminal symbols and whose root is fertile.) Given a partial sentence $\bar{\sigma}$ of G, we call the set of sentences of $K(G)$ containing $\bar{\sigma}$ as a subtree the closure of $\bar{\sigma}$. We then define the topology of partial sentences of G, denoted $\mathbf{T}^*$, as the topology whose class of closed sets is generated by the closures of the partial sentences of G. Finally, we designate by $\mathbf{T}_\Pi{}^*$ the topology generated by $\mathbf{T}_\Pi$ and $\mathbf{T}^*$. We can demonstrate that $\mathbf{T}^*_{\Pi_2}$ is weaker than $\mathbf{T}^*_{\Pi_1}$, if Π_2 contains Π_1 and if Π_1 (thus also Π_2) contains all branches whose roots are not fertile nodes.

(15)

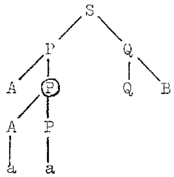


FIGURE 8

Given two tree languages, K and K', and a function f from K to K', we can ask if, for each set of phrases Π of K, there is a set of phrases Π' of K' sufficiently large so that f will be continuous from $\mathbf{T}_\Pi(K)$ to $\mathbf{T}_{\Pi'}(K')$, or from $\mathbf{T}_\Pi{}^*(K)$ to $\mathbf{T}^*_{\Pi'}(K')$. We will say that f is structurally continuous (structurally $*$-continuous) if for each finite set of phrases Π of K there exists a finite set of phrases Π' of K' such that f is continuous from $\mathbf{T}_\Pi(K)$ to $\mathbf{T}_{\Pi'}(K')$ (from $\mathbf{T}_\Pi{}^*(K)$ to $\mathbf{T}^*_{\Pi'}(K')$); f is called a structural homeomorphism ($*$-homeomorphism) if f is bijective and if f and its inverse f^{-1} are structurally continuous (structurally $*$-continuous).

It can be shown, for context-free languages K_1 and K_2, that if f is a structural homeomorphism from K_1 onto K_2, then it is a structural

*-homeomorphism (Part I, Theorem 6). The converse is not true; there exist context-free languages which are structurally *-homeomorphic, but not structurally homeomorphic. (See, for example, Part II, Section 6, Examples 1 and 2.) Thus for context-free languages it is justified to use the terms "strong structural homeomorphism" and "weak structural homeomorphism" instead of "structural homeomorphism" and "structural *-homeomorphism."

We have been speaking about tree languages. Let us consider now the string languages that are associated with tree languages. Topologies on a string language can be canonically induced from an associated tree language. Let K be a tree language and let $L = L(K)$. $\mathbf{T}_{\Pi}(L)$ and $\mathbf{T}_{\Pi}{}^*(L)$ are, by definition, the strongest topologies such that the function $\tau \to |\tau|$ from K to L is continuous from $\mathbf{T}_{\Pi}(K)$ to $\mathbf{T}_{\Pi}(L)$ and from $\mathbf{T}_{\Pi}{}^*(K)$ to $\mathbf{T}_{\Pi}{}^*(L)$. We can define structural continuity and structural homeomorphism for functions from string languages to string languages in the same way as with tree languages.

Let us now assume that the string languages $L(K)$ and $L(K')$, associated with distinct tree languages K and K', are identical, and let L be the string language $L(K) = L(K')$. If the identity function on L is structurally homeomorphic (structurally *-homeomorphic) from $L(K)$ to $L(K')$ (that is, if the identity function on L is structurally homeomorphic (structurally *-homeomorphic) from the structure induced on L by K to the structure induced on L by K'), we say that K and K' define homeomorphic (*-homeomorphic) constituent structures on L. More particularly, suppose that K and K' are generated by context-free grammars G and G'; then we say that G and G' are strongly (weakly) structurally equivalent.

Each weakly equivalent (in Chomsky's sense) class of context-free grammars can be divided into subclasses of weakly structurally equivalent grammars, and each class of weakly structurally equivalent grammars can be divided into subclasses of strongly structurally equivalent grammars. The weakly equivalent grammars (in Chomsky's sense) G_1, G_1', $\bar{G}_1$, G_2, G_2', $\bar{G}_2$ given above separate into two strong structural equivalence classes, one comprising G_1, G_1', and $\bar{G}_1$, and the other comprising G_2, G_2', and $\bar{G}_2$.

The topological method of language and grammar classification which we have just sketched seems to provide a better means than the notion of strong equivalence (in Chomsky's sense) for discussing the structural similarity of languages and grammars. We can, for example, show that for each context-free grammar there is a normal grammar (in Chomsky's sense) to which it is strongly structurally equivalent (Part I, Section 6, Example). The implications of the notion of structural equivalence, however, are still not entirely clear. This sort of classification might be more or less fine than what one expects to find in searching for a classification corresponding to the "intuitive idea" of structural similarity of sentences. Of course, a formal characterization of one's intuitive idea of the structural similarity of sentences might not be a well-justified theoretical goal, but an attempt in this direction might lead to theoretically interesting results, possibly independent of the initial intuitive motivation. The situation here is reminiscent of the diverse mathematical definitions for the notion of curve.

In Part II of this article we shall apply our general theory to a restricted class of context-free languages, regular languages, which are in a certain sense counterparts of regular sets. It will be shown that within structural homeomorphism regular languages can be "normalized" in various ways. Now, a type of question one might pose with respect to structural homeomorphism is this: The condition for structural homeomorphism given in the definition of this concept is not finitary, as it refers to infinitely many spaces. One might wonder whether the criterion for structural homeomorphism can be reduced to a finitary condition. More specifically, one might wonder if regular grammars can be normalized by structural homeomorphism in such a way that among normalized grammars a finitary condition for structural homeomorphism can be stated. The result in Part II, Section 6 belongs to this problem area. It will be shown that there is a subclass of regular languages such that (i) each regular language is structurally homeomorphic to a language in this subclass, and (ii) for two languages, K and K', of this subclass to be weakly homeomorphic it is sufficient that a bijective function and its inverse defined between K and K' be continuous at the "bottom" (in a certain specified sense) of the hierarchies of structural topologies $\{\mathbf{T}_{\Pi}(K)\}$ and $\{\mathbf{T}_{\Pi'}(K')\}$ (Part II, Theorem 6).

I will mention here some questions that might be studied in the conceptual framework of this paper. First, we should clarify the meaning of the proposed classification by characterizing it according to the algebraic properties of languages and grammars. We might examine what sorts of modifications of a grammar (or operations applied to its rules) leave its structure topologically unchanged. Since languages of different types can be structurally equivalent (for example, if L is a regular set generated by a grammar whose rules are of the form $A \rightarrow Ba$, then the linear language L' generated by the grammar whose rules are $A \rightarrow aBa$, for each rule $A \rightarrow Ba$ of the regular grammar, is strongly structurally equivalent to L by the function $x \rightarrow \tilde{x}x$, where $\tilde{x}$ is the mirror image of x), we might ask, for instance, which context-free languages are structurally equivalent to regular languages. Or consider the language $L'' = \{y; y = xx, x \in L\}$. It is well known that this language is context-sensitive. But it can also be generated by a tranformational grammar whose base is the grammar G of L. More precisely, we introduce a transformational rule mapping output τ of G to a tree (16):

(16)

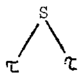


Figure 9

Now, L'' with the topological structure defined by this transformational grammar is structurally equivalent to L with the topological structure defined by G, by the function $xx \rightarrow x$ from L'' to L. We can ask, in general, what types of context-sensitive languages are structurally equivalent to context-free languages.

More generally, we might ask to what extent grammatical transformations, in Chomsky's sense, preserve the topological structure of a language. We might ask if it's possible to characterize or generalize the notion of grammatical transformation in terms of topology. To these ends, we might have to modify the notion of subtree to adapt it to the operation of deletion, which is essential in the transformational theory of natural languages.

Further, we might pose in a new way the question of how to describe the notion of grammatical ambiguity. Note that two grammars can be structurally equivalent, while the tree languages that they generate are not structurally homeomorphic. As a trivial example, consider two structurally equivalent grammars (which, by definition, generate the same string language L) and a third grammar formed as their "union." (That is, we suppose that the two given grammars do not have any nonterminal symbols in common, and we take the rules of the "union" grammar to be all the rules of the two given grammars plus rules replacing the initial symbol of the new grammar by those of the old grammars.) The "union" grammar also generates the same string language L. Algebraically, the "union" grammar is as ambiguous as the two given grammars considered together. More precisely, for each string in L generated by the three grammars, the degree of ambiguity in the "union" grammar is the sum of the degrees of ambiguity in the two given grammars. But, the "union" grammar is structurally equivalent to each of the two component grammars. This is because in the tree language generated by the "union" grammar, the tree languages generated by the given grammars form, so to speak, two disjoint homeomorphic spaces that doubly cover the string language associated with the three grammars. In this case, the increasing of the degree of algebraic ambiguity doesn't cause any change in the topological structure of the string language; the ambiguity introduced by the "union" grammar is not structurally "essential." In general, however, ambiguity can be "essential." For example, a string can be associated with various trees in such a way that the canonical function $t \to |t|$ is not locally homeomorphic at any of these trees. The topological relation between an ambiguous string and the trees associated with it represents, so to speak, the way in which the branches of the space of tree sentences cover the "surface" string language. Thus, we might be able to study the "geometric" structure of ambiguity.

As a final speculation, note that it might be possible that the topological study of grammars and languages commenced here will permit the generalization of the formal theory of grammars to the case where "languages" are continuous domains. One might axiomatize the properties of classes of structural topologies defined for certain types of grammars. One might then define in the opposite manner a "surface" language (corresponding to a string language) from a topological structure satisfying certain properties. This method might permit us to define a language without requiring a "discrete" grammatical mechanism, and to introduce a "continuous" domain provided with "constituent structure."

I. General Theory

1. *Preliminaries on Trees and Languages*

Let J^* be the free monoid generated by the set of all positive integers, J. The operator and the identity of the monoid will be denoted by $\cdot$ and 0, respectively. For p, q in J^*, we say p *dominates* q, if there is r in J^*, such that $q = p \cdot r$; p is said to *properly* dominate q if $r \neq 0$; p is said to *directly* dominate q if r is in J. We say that p *precedes* q if there is r in J^*, and j and k in J such that $p = r \cdot j$, $q = r \cdot k$ and $j < k$. A finite subset Δ of J^* is an (*unlabeled*) *tree* if it satisfies the following conditions:

1. If q is in Δ and p dominates q, p is in Δ.
2. If q is in Δ and p precedes q, p is in Δ.

An element of a tree Δ is called a *node* of Δ. If Δ is not empty, 0 necessarily belongs to Δ; 0 is called the *root* of Δ. A node of Δ is called *terminal* if it does not properly dominate any node of Δ; otherwise, it is called *nonterminal*. If a tree Δ' is a subset of a tree Δ and if a node q of Δ is in Δ' in case some node p of Δ that precedes q is in Δ', Δ' is called a *rooted subtree* of Δ (or, in case no confusion is likely, simply a subtree of Δ); we write $\Delta' < \Delta$.

For p and q in J^*, we say p is a *sister* of q if p precedes q or q precedes p. If a tree Δ' is a rooted subtree of Δ, a node of Δ which is a sister of a node of

Δ' is in Δ'. In fact, a subset Δ' of a tree Δ is a rooted subtree of Δ if and only if any node of Δ which dominates a node of Δ' and any node of Δ which is a sister of a node of Δ' are in Δ'.

Let Δ' be a nonvoid rooted subtree of Δ and consider a node p of Δ. If p is not in Δ', there exists a node q of Δ' which properly dominates p but does not properly dominate any node of Δ' that dominates p. Then, q is a terminal node of Δ'. For, otherwise, q would directly dominate a node q' of Δ', which, however, does not dominate p. Then there would be a sister of q' that dominates p but is not in Δ', a contradiction. It follows that for each node p of Δ, either it dominates or is properly dominated by some terminal node of Δ'.

Conversely, assume that a subset E of Δ satisfies the following conditions: (1) None of the nodes of E properly dominates any other; (2) for each node p of Δ, either p dominates some node in E or p is properly dominated by some node in E. Then, define Δ' as the set of those nodes of Δ which dominate some node in E. Δ' is a rooted subtree of Δ. Consider first a node q of Δ which dominates some node p in Δ'. p dominates a node in E, and, hence, so does q, too, and q is in Δ'. Secondly, suppose q is a node of Δ which is a sister of some node p in Δ'. If q dominates a node of E, it is in Δ' by the definition of Δ'; if not, it must be properly dominated by a node of E, a contradiction, since that node of E would then have to properly dominate p, and hence, also a node of E that p dominates; but no node of E properly dominates any node of E. Consequently, we have the following.

LEMMA 1. *Let Δ be a nonempty tree. Each subset E of nodes of Δ satisfying conditions* (1) *and* (2) *above determines a nonempty rooted subtree Δ' of Δ, and vice versa. E is the set of terminal nodes of Δ', and Δ' is the set of nodes of Δ that dominate a node of E.*

More generally, given a tree Δ, consider a subset Γ of Δ that satisfies the following conditions:

1. There exists a node p_0 in Γ such that p_0 dominates each node of Γ.

2. If q is in Γ, p dominates q, and p_0 dominates p, then p is in Γ.

3. For each q in Γ and p in Δ, if $q \neq p_0$ and q is a sister of p, then p is in Γ.

Γ is said to be a *branch of Δ at node p_0*, and p_0, the *root of branch Γ*. A branch of Δ at p_0 is said to be *full* if no other branch of Δ at p_0 properly includes it. For each node p of Δ there is one and only one full branch of Δ at p, which is the set of nodes of Δ dominated by p.

A rooted subtree of Δ is a branch of Δ at its root 0; Δ itself is its full branch at its root.

If Γ is a branch of Δ at p, the set Δ' of all q in J^* such that $p \cdot q$ is in Γ is a tree; Δ' is called the *subtree of Δ associated with branch Γ.* The subtree of Δ associated with the full branch at p is denoted by $p\backslash\Delta$. A subtree associated with a branch of Δ at its root is nothing but that branch considered as a rooted subtree of Δ.

We shall define the *height* $|\Delta|$ of a nonempty tree Δ recursively as follows: if $\Delta = \{0\}$, $|\Delta| = 0$; otherwise, $1 \in \Delta$, and we put

$$|\Delta| = \operatorname*{Max}_{j \in J \cap \Delta} |j\backslash\Delta| + 1.$$

If Δ' is a nonempty rooted subtree of Δ, the *height of Δ relative to Δ'* is defined as the maximum of the heights of the subtrees of Δ associated with the full branches of Δ at the terminal nodes of Δ'. For each node p of Δ, the *distance* $d(p, \Delta')$ of p from Δ' is, by definition, zero, if p is in Δ', and, otherwise, $d(p', \Delta') + 1$, where p' is the node of Δ which directly dominates p.

Let Δ_1 and Δ_2 be two trees. Then, $\Delta = \Delta_1 \cup \Delta_2$ and $\Delta' = \Delta_1 \cap \Delta_2$ are also trees. But Δ' may not be a rooted subtree of Δ; for, a sister in Δ of a node of Δ' may not be a node of Δ'. If Δ' is indeed a rooted subtree of Δ, we may say that Δ and Δ' are the *union* tree and the *intersection* tree of Δ_1 and Δ_2, respectively (or, if no confusion is likely, simply the union and the intersection of Δ_1 and Δ_2). Otherwise, the union tree and the intersection tree of Δ_1 and Δ_2 are not defined. Note that if Δ_1 and Δ_2 are both rooted subtrees of a tree Δ_3, then Δ_1 and Δ_2 have their union and intersection, which are also rooted subtrees of Δ_3.

A (*labeled*) *tree T over a set V* (or a (labeled) tree if V is understood) is defined as a function from some (unlabeled) tree Δ into V. The domain Δ of T is denoted by $\Delta(T)$. A pair (p, A), where $p \in \Delta(T)$ and $T(p) = A$, is called a (*labeled*) *node* of T, and A, its *label*; by an extended use of language, p is also said to be a node of T, and A, its label. The set of all trees over V is denoted by $V^{\#}$. An element A of V may be considered as a tree T over V such that $\Delta(T) = \{0\}$ and $T(0) = A$. Then, $V \subset V^{\#}$.

The *height* $|T|$ of a labeled tree T over V is, by definition, the height of its domain. If Δ' is a rooted subtree of $\Delta(T)$, the restriction T' of T to Δ' is also a labeled tree and is called a *rooted subtree* of T; we also write $T' < T$. Clearly, $<$ defines a partial order in $V^{\#}$.

We have an analog of Lemma 1 for labeled trees.

LEMMA 2. *Let T be a nonempty tree over V. Each subset E of nodes of the domain $\Delta(T)$ of T satisfying the conditions* (1) *and* (2) *referred to in Lemma* 1 *determines a nonempty rooted subtree T' of T, and vice versa. E is the set of terminal nodes of T', and $\Delta(T')$ is the set of nodes of T that dominate a node of E.*

Let T be a labeled tree and Δ its domain. If Γ is a branch of Δ at node p, the restriction $T \mid \Gamma$ of T on Γ is called a *branch of T at p_0*. Let Δ' be the subtree of Δ associated with Γ and define a tree T' as follows: For each q in Δ', $T'(q) = T(p \cdot q)$. T' is called the *subtree of T associated with the branch $T \mid \Gamma$*. If Γ is the full branch of Δ at p, T' is denoted by $p\backslash T$. We have $\Delta(p\backslash T) = p\backslash\Delta(T)$. Given a labeled tree T, if its root is labeled with A, if n is the greatest integer belonging to $\Delta(T)$, and if $T_i = i\backslash T$ for each i, $1 \leqslant i \leqslant n$, we write $T = A(T_1, T_2, \ldots, T_n)$.

By an extended use of language, somewhat ambiguously, a branch of the domain of a labeled tree T may also be understood as a branch of T.

T_1 and T_2 being labeled trees, if $\Delta(T_1)$ and $\Delta(T_2)$ have union Δ and intersection Δ', and if $T_1 \mid \Delta' = T_2 \mid \Delta'$, then $T' = T_1 \mid \Delta' = T_2 \mid \Delta'$ is called the *intersection* of T_1 and T_2, and the tree T defined by $T \mid \Delta(T_1) = T_1$, $T \mid \Delta(T_2) = T_2$ is called the union of T_1 and T_2. Note that for any tree T_3, $T_3 < T'$ if and only if $T_3 < T_1$ and $T_3 < T_2$; $T < T_3$ if and only if $T_1 < T_3$ and $T_2 < T_3$.

If the domain of a tree T is the domain of a tree T' less the set of nodes properly dominated by a node p of T', and if T is the restriction of T' to this domain, then T is said to be obtained from T' by *pruning* T' at p. A tree T is said to be obtained from another, T', by *grafting* a third, T'', at a node p of T', if p is a node of T, if the subtree associated with the full branch of T at p is T'' and if T' is obtained from T by pruning T at p. A tree T is said to be obtained from another, T', by *replacing* the full branch of T'at p by a tree T'',if T is obtained by grafting T'' at p on a tree which has been obtained by pruning T' at p. When no confusion is likely, by an extended use of language, we may say that a tree T is obtained from T' by grafting at a node p a *branch* of a third, T'', meaning grafting the subtree of T'' associated with that branch; similarly for "replacing by a *branch*."

The *yield function* η of trees over V is a function from $V^{\#}$, the set of trees over V, to V^*, the set of strings over V, i.e., the free monoid generated by V, which is defined recursively as follows:

If $\Delta(T) = \varnothing$, $\eta(T) = \epsilon$, where ϵ is the identity of V^*; if $\Delta(T) = \{0\}$, $\eta(T) = T(0)$; otherwise, $\eta(T) = \eta(1\backslash T) \cdot \eta(2\backslash T) \cdots \eta(j\backslash T)$, where j is the greatest positive integer in $\Delta(T)$.

Let now V and W be two sets. A tree τ over the union of V and W is called *W-leaved V-tree*, if the labels of the terminal nodes of τ are all in W and those of the nonterminal nodes are all in V. The set of all W-leaved

V-trees is denoted by $(V, W)^{\#}$. A *tree language K over* (V, W), or simply a tree language if V and W are understood, is, by definition, a subset of $(V, W)^{\#}$. V and W are called the *nonterminal* and the *terminal vocabularies* of K and their union the *vocabulary* of K.

An element of a tree language K over (V, W) is also called a *sentence* of K. The subtree of sentence τ of K which is associated with the full branch of τ at a node p of τ is called the *phrase of τ at p*. A W-leaved V-tree π is called a *phrase of K* if there exists a sentence of K of which π is a phrase. A tree T is said to *belong to* K if there exists a phrase of K such that $T < \pi$. If τ is a sentence of K, τ "belongs to K" in this sense. But note that "T belongs to K" in general does not mean the same thing as "T is an element of K."

Let U be a set. A *string language over* U is a subset of U^*, the free monoid generated by U. The image of the yield function of a tree language K is a string language over the terminal vocabulary of K. If a string language L is the image of a tree language K by the yield function, L is said to be *the surface string language of* K and K is said to be a *covering tree language of* L.

A *phrase-structure tree language* K is, by definition, a tree language whose terminal and nonterminal vocabularies are disjoint. Then, an element of its terminal (or nonterminal) vocabulary is called a *terminal* (or *nonterminal*) *symbol of* K (or, simply, a terminal or nonterminal of K).

A *phrase-structure string language* is, by definition, a pair of a string language L and a tree language K covering L. When K is understood, by an extended use of language L is also said, somewhat ambiguously, to be a phrase-structure string language. An element of L is called a sentence of the phrase-structure string language (L, K), or simply, of L. An element τ of K is called a *structural description* of sentence $\eta(\tau)$ of L. A sentence of L has at least, and possibly more than, one structural description; if the latter is the case, it is called *ambiguous*.

In what follows, we may simply say "tree language" instead of "phrase-structure tree language," as it is the only kind we shall be concerned with. Also, if no confusion is likely, we may omit the modifiers "tree" and "string" in front of "language."

EXAMPLE. Context-Free Language. A context-free grammar G (or, more specifically, an ϵ-free context-free grammar) is said to be defined if there are given two disjoint finite sets, V and W, a specified element S of V, and a finite number of pairs (A, ω), where A is an element of V and ω is a nonempty string over the union of V and W. The pairs (A, ω) are called rules of G. A tree belonging to G is a nonempty tree T over the union of V and W defined recursively as follows: Let A be the label of the root of T and for each integer j

that is a node of T, let α_j be its label; then T belongs to G if and only if (1) (A, ω) is a rule of G, where $\omega = \alpha_1\alpha_2 \cdots \alpha_k$ and k is the greatest integer belonging to T, and (2) each $j\backslash T$, $1 \leqslant j \leqslant k$, is a tree belonging to G. Note, in particular, if $\Delta(T) = \{0\}$, i.e., if no integer is a node of T, then T belongs to G. (That is, all trees over the union of V and W of height 0, i.e., all elements of this union, belong to G). The set of W-leaved V-trees whose root is S and which belong to G is called the *tree language generated by* G; notation, $K(G)$. The string language which $K(G)$ covers is called the *string language generated by* G; notation, $L(G)$. The phrase-structure string language $(L(G), K(G))$ is called the *phrase-structure string language generated by* G. By an extended use of language, if no confusion is likely, $L(G)$ may also denote this phrase-structure string language.

Most often we are interested in languages that are generated by some specified finitary mechanism, e.g., an ϵ-free context-free phrase-structure grammar. But in what follows we continue to deal with tree languages in the general setting. In particular, we do not assume that the roots of sentences of a tree language are all labeled with the same symbol. If such is the case, as with context-free languages, the languages is said to be *uni-rooted.*

2. *Topologies of Preorder Associated with Pruning Sets*

Let K be a phrase-structure tree language over (V, W) and let Π be a finite set of phrases of K. In the following context, such a set is called a *pruning set of* K. For each sentence τ of K, define $(\tau)_\Pi$ as follows: $(\tau)_\Pi$ is the smallest rooted subtree of τ such that for each of its terminal nodes p, p is a terminal node of τ or the phrase of τ at p is a phrase in Π.

Define a binary relation $<_\Pi$ in K as follows: $\tau <_\Pi \sigma$ if and only if $(\tau)_\Pi < (\sigma)_\Pi$. The relation $<_\Pi$ is reflexive and transitive (but not necessarily antisymmetric); i.e., it is a preorder in K. As a preorder it canonically introduces a topology in K, the topology whose open sets are generated by sets $V_\Pi(\tau) = \{\sigma; \tau <_\Pi \sigma\}$. This topology is denoted by $\mathbf{T}_\Pi(K)$, or when K is understood, by $\mathbf{T}_\Pi$.

The topology $\mathbf{T}_\Pi$ may also be described in the following way. For each tree T over the union of V and W, put $\overline{V}_\Pi(T) = \{\sigma; \sigma \in K, T < (\sigma)_\Pi\}$. (We have, then, $V_\Pi(\tau) = \overline{V}_\Pi((\tau)_\Pi)$ for each τ in K. Note also that for each τ in K, $\overline{V}_\Pi(\tau)$ is empty or the singleton set $\{\tau\}$.) The class of $\overline{V}_\Pi(T)$, where T ranges over $(V \cup W)^{\#}$, constitutes a basis of the class of open sets of $\mathbf{T}_\Pi$.

For if $\overline{V}_\Pi(T_1)$ and $\overline{V}_\Pi(T_2)$ are not disjoint, for some sentence τ of K, both T_1 and T_2 are rooted subtrees of τ, hence they have union, call it T. Note that for any σ in K, $T < (\sigma)_\Pi$ if and only if $T_1 < (\sigma)_\Pi$ and $T_2 < (\sigma)_\Pi$. Thus

$\bar{V}_\Pi(T) = \bar{V}_\Pi(T_1) \cap \bar{V}_\Pi(T_2)$. It follows that the class of $\bar{V}_\Pi(T)$ is a basis of a topology of K. Since $V_\Pi(\tau) = \bar{V}_\Pi((\tau)_\Pi)$, $V_\Pi(\tau)$ is an open set in this topology. Conversely, $\bar{V}_\Pi(T) = \cup\, V_\Pi(\tau)$, where the union is over τ in $\bar{V}_\Pi(T)$, hence $\bar{V}_\Pi(T)$ is open in $\mathbf{T}_\Pi$. Consequently, the two topologies are identical.

Let K and K' be two languages and let Π and Π' be pruning sets of K and K', and let f be a function from K to K'. Then, f is continuous from $\mathbf{T}_\Pi$ to $\mathbf{T}_{\Pi'}$ if and only if f is an isotone, i.e., if and only if for any pair σ and τ of K such that $(\tau)_\Pi < (\sigma)_\Pi$, we have $(f(\tau))_{\Pi'} < (f(\sigma))_{\Pi'}$. This is a reflex of the fact that the topology $\mathbf{T}_\Pi$ is nothing more than a preordered set.

EXAMPLE. For any phrase-structure language K, $\mathbf{T}_\Pi$ is the discrete topology on K if Π is void. More generally, if no phrase in Π is of height more than 0, then $\mathbf{T}_\Pi$ is discrete.

An *extension* of a pruning set Π of a language K is a pruning set of K which includes Π. As Π becomes larger, $(\tau)_\Pi$ becomes smaller, but $\mathbf{T}_\Pi$ does not necessarily become weaker as Π extends. Let Π be a pruning set of K and let Π' be an extension of Π. Let σ and τ be two sentences of K such that $(\tau)_\Pi < (\sigma)_\Pi$. Then we have $(\tau)_{\Pi'} < (\tau)_\Pi$ and $(\sigma)_{\Pi'} < (\sigma)_\Pi$, but from these it does not necessarily follow that $(\tau)_{\Pi'} < (\sigma)_{\Pi'}$. That is, the identity map of K may not be continuous from $\mathbf{T}_\Pi$ to $\mathbf{T}_{\Pi'}$. But we have the following.

THEOREM 1. *Let K be a phrase-structure tree language and let Π be a pruning set of K, and let Π' be an extension of Π. Then there exists an extension Π'' of Π' such that the identity map of K is continuous from $\mathbf{T}_\Pi$ to $\mathbf{T}_{\Pi''}$.*

To prove this theorem, define Π'' as the set of all phrases π'' of K satisfying the following condition: There exists a phrase π' in Π' such that $(\pi'')_\Pi < (\pi')_\Pi$. Π'' is a finite set, because $(\pi'')_\Pi$ is a rooted subtree of π'; for each π' in Π', there are only a finite number of rooted subtrees of π', and for each rooted subtree T of π', there are only a finite number of phrases π'' such that $(\pi'')_\Pi = T$. Note also that Π' is included in Π''. Hence, Π'' is an extension of Π'.

Assume, now, for two sentences σ and τ of K, we have $(\tau)_\Pi < (\sigma)_\Pi$. If $(\tau)_{\Pi''}$ were not a rooted subtree of $(\sigma)_{\Pi''}$, there would exist a node of $(\tau)_{\Pi''}$, call it p, which is not a node of $(\sigma)_{\Pi''}$. Since $(\tau)_{\Pi''} < (\tau)_\Pi$ and $(\tau)_\Pi < (\sigma)_\Pi$, p is a node of $(\sigma)_\Pi$, and also of σ. By Lemma 2 of Section 1, there is a terminal node q of $(\sigma)_{\Pi''}$ that properly dominates p in σ. The phrase π'' of σ at q is a phrase in Π''. Since p is a node of $(\tau)_{\Pi''}$, and q dominates p, q is a node of $(\tau)_{\Pi''}$, and, hence, also of $(\tau)_\Pi$ and of τ. Let π be the phrase of τ at q. Then, we have $q\backslash(\tau)_\Pi = (\pi)_\Pi$ and $q\backslash(\sigma)_\Pi = (\pi'')_\Pi$. Since $(\tau)_\Pi < (\sigma)_\Pi$, we have

$(\pi)_\Pi < (\pi'')_\Pi$. From the fact that π'' is in Π'', there exists a phrase π' in Π' such that $(\pi'')_\Pi < (\pi')_\Pi$. It follows that $(\pi)_\Pi < (\pi')_\Pi$ and π is in Π''. But π is a phrase of τ at q, and q properly dominates p, which is a node of $(\tau)_{\Pi''}$, a contradiction. Hence, $(\tau)_{\Pi''} < (\sigma)_{\Pi''}$ must hold. The identity of K is continuous from $\mathbf{T}_\Pi$ to $\mathbf{T}_{\Pi''}$.

Let us call a collection $\mathbf{P}$ of pruning sets of a language K a *pruning system* of K. Given a pruning system $\mathbf{P}$, we can associate it with a system of topologies $\mathbf{T}_\Pi$, where Π ranges over the pruning sets of $\mathbf{P}$. This system of topologies may be considered to be equipped with the partial order induced by the partial order of set-theoretical "inclusion" defined on $\mathbf{P}$, and with a collection of maps from $\mathbf{T}_\Pi$ to $\mathbf{T}_{\Pi'}$ for pairs Π, Π' such that $\Pi \subset \Pi'$, namely the collection of the maps induced by the identity map of K. The system of topologies so conceived is called the *phrase-structure* (*or*, *simply*, *structure*) *topological system of K associated with* $\mathbf{P}$. If this system is an inductive system of topologies (i.e., if the identity map is continuous for each pair Π, Π' such that $\Pi \subset \Pi'$, and if for each pair Π, Π', there is an upper bound) $\mathbf{P}$ is called an *inductive pruning system*.

3. *Partial and Prepartial Sentences*

As indicated above, the class of all pruning sets, in general, is not an inductive pruning system. That is, the phrase-structure topological system associated with the class of all pruning sets is not an inductive system of topologies. But on the basis of topologies $\mathbf{T}_\Pi$ we can construct an inductive system of topologies associated with the class of all pruning sets. For this purpose, we shall introduce the notion of partial sentence. Although our primary interest concerns partial sentences, for convenience of definition and for ease of exposition, we shall also define the notion of prepartial sentence.

If K is a phrase-structure tree language over (V, W), a tree $\bar{\tau}$ over the union of V and W is said to be a *prepartial sentence* of K if it satisfies the following conditions:

P1. There exists a sentence τ of K such that $\bar{\tau} < \tau$.

P2. $\bar{\tau}$ contains at least one branch whose terminal nodes are all labeled with terminal symbols (i.e., elements of W). Such a branch is called a *phrase of* $\bar{\tau}$.

From condition P1 it follows that a phrase of a prepartial sentence of K is a phrase of K.

A phrase of a prepartial sentence $\bar{\tau}$ may be contained in another phrase of $\bar{\tau}$ as a proper branch; if not, it is called a *maximal phrase of* $\bar{\tau}$. Evidently, each phrase of $\bar{\tau}$ is contained in one and only one maximal phrase of $\bar{\tau}$. A prepartial sentence is a sentence if and only if it is a maximal phrase of itself.

A prepartial sentence $\bar{\tau}$ is called a *partial sentence* if it satisfies the following condition:

P3. $\bar{\tau}$ is either a sentence or else there exists a maximal phrase π_0 with the following property: Let $\bar{\tau}_0$ be the tree obtained from $\bar{\tau}$ by pruning all the maximal phrases of $\bar{\tau}$ except for π_0; there exist infinitely many phrases π_i, $i \geqslant 1$, of K such that the tree $\bar{\tau}_i$ obtained from $\bar{\tau}_0$ by replacing π_0 by π_i is a prepartial sentence of K.

Such a π_0 is called a *fertile phrase* of $\bar{\tau}$, and its root, a *fertile node* of $\bar{\tau}$. Note that if $\bar{\tau}$ is a partial sentence, $\bar{\tau}_i$, $i \geqslant 0$, satisfying condition P3 are also all partial sentences. Note also that sentences of K are partial sentences.

The set of those sentences of K of which a prepartial sentence $\bar{\tau}$ is a rooted subtree is called the *closure* of $\bar{\tau}$ and is denoted by $C(\bar{\tau})$. If $\bar{\tau}$ is a sentence, $C(\bar{\tau}) = \{\bar{\tau}\}$.

A prepartial sentence may be considered as a specification of a sort for a sentence to contain a specified phrase, or phrases, of K in particular grammatical relations. The closure of a prepartial sentence, then, is the set of sentences of K that satisfy such a specification.

We now define the topology of (pre)partial sentences of K, $\mathbf{T}^*(K)$ (or $\mathbf{T}^{(*)}(K)$) as the topology on K whose class of closed sets is generated by the class of the closures of (pre)partial sentences of K. (If K is understood, we simply write $\mathbf{T}^*$ and $\mathbf{T}^{(*)}$, instead of $\mathbf{T}^*(K)$ and $\mathrm{T}^{(*)}(K)$.) The open sets of $\mathbf{T}^*$ (or $\mathbf{T}^{(*)}$) are then those whose complements are intersections of finite unions of closures of (pre)partial sentences.

Given a set of prepartial sentences $\bar{\tau}_\lambda$, consider the intersection C of $C(\bar{\tau}_\lambda)$. Assume that C is not void, and let σ be a sentence in C. Then, $\bar{\tau}_\lambda$ are all rooted subtrees of σ, and, hence, there are only finitely many different $\bar{\tau}_\lambda$'s. Their union is also a rooted subtree of σ; call it $\bar{\tau}$. It is certainly a prepartial sentence. Thus, we have

THEOREM 2. *The intersection of the closures of a nonvoid collection of prepartial sentences is either void or the closure of a prepartial sentence.*

Assume, now, that K is context-free and that one of the $\bar{\tau}_\lambda$'s is a partial sentence. Consider a fertile phrase of this partial sentence. It is a phrase in $\bar{\tau}$. If it is a maximal phrase in $\bar{\tau}$, it is a fertile phrase of $\bar{\tau}$; if it is not maximal in $\bar{\tau}$, the maximal phrase of $\bar{\tau}$ which contains it is a fertile phrase of $\bar{\tau}$. In any case, then, $\bar{\tau}$ is a partial sentence. Hence,

COROLLARY. *Let K be a context-free tree language. The intersection of the closures of a nonvoid collection of partial sentences of K is either void or the closure of a partial sentence.*

4. *Prepartial Sentences and Topologies Associated with Pruning Sets*

Let K be a phrase-structure tree language. Consider a pruning set Π of K and an extension Π' of Π. Let σ and τ be two sentences of K and assume that $(\tau)_\Pi$ is a rooted subtree of $(\sigma)_\Pi$, but $(\tau)_{\Pi'}$ is not a rooted subtree of $(\sigma)_{\Pi'}$. Thus there is a node p of $(\tau)_{\Pi'}$ which is not a node of $(\sigma)_{\Pi'}$. Since $(\tau)_{\Pi'} < (\tau)_\Pi$ we have $(\tau)_{\Pi'} < (\sigma)_\Pi$, and also, p is a node of $(\sigma)_\Pi$. On the other hand, $(\sigma)_{\Pi'} < (\sigma)_\Pi$. Consequently, there is a terminal node q of $(\sigma)_{\Pi'}$ which properly dominates p in $(\sigma)_\Pi$. The phrase of σ at q is a phrase in Π'. Let q_i denote such terminal nodes of $(\sigma)_{\Pi'}$, that is, those which properly dominate a node of $(\tau)_{\Pi'}$ which is not a node of $(\sigma)_{\Pi'}$. Clearly, the q_i's do not properly dominate each other.

Consider now a terminal node of $(\sigma)_{\Pi'}$ which is not any of these q_i's. It is a node of $(\sigma)_\Pi$, and since $(\tau)_{\Pi'} < (\sigma)_\Pi$, either it is a node of $(\tau)_{\Pi'}$ (in which case it is a terminal node of $(\tau)_{\Pi'}$, for otherwise it would have to be one of q_i's), or else there is one and only one terminal node of $(\tau)_{\Pi'}$ which properly dominates it. Denote by s_j those terminal nodes of $(\sigma)_{\Pi'}$ of the former case and denote by r_k those of the latter; denote by t_k the terminal node of $(\tau)_{\Pi'}$ which properly dominates r_k. Note that none of the nodes s_j and t_k properly dominate each other, as they are all terminal nodes of $(\tau)_{\Pi'}$. Furthermore, no q_i properly dominates any s_j or t_k and vice versa. For if q_i, which is a terminal node of $(\sigma)_{\Pi'}$, properly dominated s_j or t_k, it would properly dominate a terminal node of $(\sigma)_{\Pi'}$, s_j or r_k, a contradiction; s_j cannot properly dominate any q_i for the same reason; finally, if t_k, which is a terminal node of $(\tau)_{\Pi'}$, properly dominated q_i, it would properly dominate a node of $(\tau)_{\Pi'}$, again a contradiction. So, we have a set of nodes q_i, s_j, t_k of $(\tau)_{\Pi'}$ which do not properly dominate each other.

We now see that this set of nodes of $(\tau)_{\Pi'}$ defines a rooted subtree of $(\tau)_{\Pi'}$, that is, there is a rooted subtree of $(\tau)_{\Pi'}$ whose terminal nodes are the nodes in the set. Take a node p of $(\tau)_{\Pi'}$. Since $(\tau)_{\Pi'} < \sigma$, p is a node of σ. Hence, p either dominates, or is properly dominated by, some terminal node of $(\sigma)_{\Pi'}$, i.e., one of q_i, s_j, or r_k. But since p is a node of $(\tau)_{\Pi'}$, if it dominates r_k, it must also dominate t_k, the terminal node of $(\tau)_{\Pi'}$ dominating r_k. On the other hand, it is impossible that any r_k properly dominates p (for, if so, it would have to be some q_i). Hence p either dominates or is properly dominated by, some q_i, s_j, or t_k. It follows that the nodes of $(\tau)_{\Pi'}$ that dominates q_i, s_j, or t_k form a rooted subtree of $(\tau)_{\Pi'}$ (Section 1, Lemma 2).

Call $\bar{\tau}$ this rooted subtree of $(\tau)_{\Pi'}$. Let $\bar{\sigma}$ be the tree obtained from $\bar{\tau}$ by grafting at each q_i the phrase of σ at q_i. In other words, $\bar{\sigma}$ may be characterized as follows: $\bar{\tau}$ is a rooted subtree of $\bar{\sigma}$; for each q_i, the subtree of $\bar{\sigma}$ associated with the full branch at q_i is the phrase of σ at q_i; all s_j's and t_k's are terminal. Now, from our assumption on τ, there exists at least one q_i, hence $\bar{\sigma}$ is a prepartial sentence. Note that σ is in the closure of $\bar{\sigma}$. We have the following

LEMMA. *Let K, Π, Π', be as above and let τ be a sentence of K. There exists a finite (possibly zero) number of prepartial sentences the union of whose closures contains the set of sentences σ such that $(\tau)_\Pi$ is a rooted subtree of $(\sigma)_\Pi$ but $(\tau)_{\Pi'}$ is not a rooted subtree of $(\sigma)_{\Pi'}$.*

We have seen that each such σ is in the closure of a $\bar{\sigma}$ constructed as above, and there can only be a finite number of trees that can be such a $\bar{\sigma}$.

For each pruning set Π of K, we now define $\mathbf{T}_\Pi^{(*)}(K)$, (or simply $\mathbf{T}_\Pi^{(*)}$, when K is understood) as the topology generated by $\mathbf{T}_\Pi$ and $\mathbf{T}^{(*)}$. We have the following

THEOREM 3. *If Π' is an extension of a pruning set Π of a phrase-structure tree language K, the identity map of K is continuous from $\mathbf{T}_\Pi^{(*)}$ to $\mathbf{T}_{\Pi'}^{(*)}$; i.e., $\mathbf{T}_{\Pi'}^{(*)}$ is weaker than $\mathbf{T}_\Pi^{(*)}$.*

Let τ be a sentence of K. An open neighborhood of τ in $\mathbf{T}_{\Pi'}^{(*)}$ contains a set of the form $V' = V_{\Pi'}(\tau) \cap E$, where E is the complement of the union C of the closures of a finite number of prepartial sentences $\bar{\tau}_j$, $1 \leqslant j \leqslant m$. Let $\bar{\sigma}_i$, $1 \leqslant i \leqslant n$, be the prepartial sentences determined by the preceding lemma for τ. Put $V_0 = V_\Pi(\tau) \cap E_0$ where E_0 is the complement of the union of the closures of $\bar{\sigma}_i$, $1 \leqslant i \leqslant n$. From the lemma, we have $V_0 \subset V_{\Pi'}(\tau)$. Put $V = V_0 \cap E$. V is an open neighborhood of τ in $\mathbf{T}_\Pi^{(*)}$ and is contained in V'. It follows that the identity map of K is continuous from $\mathbf{T}_\Pi^{(*)}$ to $\mathbf{T}_{\Pi'}^{(*)}$ at τ, an arbitrary point of $\mathbf{T}_\Pi^{(*)}$, hence the theorem.

Let $\mathbf{P}$ be an arbitrary pruning system of K. The class of topologies $\mathbf{T}_\Pi^{(*)}$, where Π ranges over $\mathbf{P}$, together with the identity maps of K from $\mathbf{T}_\Pi^{(*)}$ to $\mathbf{T}_\Pi^{(*)}$ is called the *phrase-structure* (*)-*topological system of K associated with* $\mathbf{P}$. If this system is an inductive system of topologies, $\mathbf{P}$ is called (*)-*inductive.*

COROLLARY. *For $\mathbf{P}$ to be (*)-inductive, it is sufficient that $\mathbf{P}$ is a directed set.*

In particular, the class of topologies $\mathbf{T}_{\Pi}^{(*)}$ *for all pruning sets is an inductive system of topologies.*

5. *Partial Sentences and Topologies Associated with Pruning Sets*

Let K be as above. For each pruning set Π of K, let us define $\mathbf{T}_{\Pi}{}^{*}(K)$ (or simply $\mathbf{T}_{\Pi}{}^{*}$) as the topology on K generated by $\mathbf{T}^{*}$ and $\mathbf{T}_{\Pi}$. Unlike the case with $\mathbf{T}_{\Pi}^{(*)}$, the continuity of the identity of K from $\mathbf{T}_{\Pi}{}^{*}$ to $\mathbf{T}_{\Pi'}$, where Π' is an extension of Π, is not automatically guaranteed, because the prepartial sentences referred to in the lemma in the preceding section may not be partial sentences. However, by assigning somewhat different meaning to $\mathbf{T}_{\Pi}{}^{*}$ from the one given above, we can obtain an inductive system of topologies $\mathbf{T}_{\Pi}{}^{*}$ over the class of all pruning sets.

Let us look back to the discussion that led us to the lemma in the preceding section. Let τ and σ be as defined at the beginning of Section 4, and we also use the notations q_i, $\bar{\sigma}$, etc. with the same meaning as there. Recall that q_i is a node of $\bar{\sigma}$ and τ, and that as a node of τ it properly dominates a node of $(\tau)_{\Pi'}$. The subtree associated with the full branch of $\bar{\sigma}$ at q_i is a phrase of $\bar{\sigma}$ and it is a phrase in Π'. Now assume that $\bar{\sigma}$ is not a partial sentence. Then, the maximal phrase that contains q_i cannot be fertile. The root q_i' of this maximal phrase may or may not be q_i, but in any case it dominates q_i, and hence, is a node of τ which properly dominates a node of $(\tau)_{\Pi'}$.

Let us at this point introduce the following notions. A nonterminal symbol A is called *sterile*, if there exists a prepartial, nonpartial sentence, the root of one of whose maximal phrases is labeled with A. A phrase of K is called *sterile* if its root is labeled with a sterile symbol. Let Ω^{*} be the set of all sterile phrases.

Assume, for the sake of argument, that Ω^{*} is finite and that Π (and hence also Π') contains Ω^{*}. Let us return to the discussion of $\bar{\sigma}$. Node q_i' is labeled with a sterile symbol, and hence the phrase of τ at q_i' is sterile and in Π'. But then this phrase must be pruned in $(\tau)_{\Pi'}$. On the other hand, q_i' is a node of τ which properly dominates a node of $(\tau)_{\Pi'}$, which is a contradiction. Thus, under the assumption that Π' contains Ω^{*}, $\bar{\sigma}$ cannot be nonpartial; i.e., it must be a partial sentence.

With the above discussion in mind, in order to generalize the lemma of Section 4, we shall at this point introduce the following way of "relativizing" the theory with respect to a (finite or infinite) set Ω of phrases of K. When the theory is relativized with respect to Ω, given a pruning set Π, we prune trees using not only phrases in Π, but also those in Ω. That is, for each sentence τ of K, $(\tau)_{\Pi}$ in the relativized sense is equal to $(\tau)_{\Pi\cup\Omega}$ in the sense defined in Section 2. On the basis of this revised notion, we define the preorder $<_{\Pi}$, the topologies $\mathbf{T}_\Pi$, $\mathbf{T}_\Pi^{(*)}$, $\mathbf{T}_\Pi{}^*$ in the relativized theory as in the "absolute" theory.

Let us return to our problem of comparing the two topologies $\mathbf{T}_\Pi{}^*$ and $\mathbf{T}_{\Pi'}^*$. τ, σ, $\bar{\sigma}$ being as above, if the theory is relativized with respect to the set of all sterile phrases, Ω^*, each q_i cannot be contained in a sterile phrase, and $\bar{\sigma}$ must be a partial sentence. Furthermore, the phrase of $\bar{\sigma}$ at q_i must be in Π'; consequently, as in the proof of the lemma in Section 4, there are only a finite number of trees that can be $\bar{\sigma}$. Thus, in the theory relativized with respect to Ω^*, the lemma and the theorem corresponding to the lemma and the theorem of Section 4 hold with respect to partial sentences instead of prepartial sentences.

We now make a convention once and for all that *when topologies* $\mathbf{T}_\Pi{}^*$ *are mentioned it is assumed that we are dealing with the theory relative to* Ω^* without explicitly stating so. We state

THEOREM 4. *If* Π' *is an extension of a pruning set of a phrase-structure tree language* K, *then the identity map of* K *is continuous from* $\mathbf{T}_\Pi{}^*$ *to* $\mathbf{T}_{\Pi'}^*$; *i.e.*, $\mathbf{T}_{\Pi'}^*$ *is weaker than* $\mathbf{T}_\Pi{}^*$.

The class of topologies $\mathbf{T}_\Pi{}^*$, where Π ranges over a pruning system $\mathbf{P}$, is said to be the *structural *-topological system of* K *associated with* $\mathbf{P}$. If this system is an inductive system of topologies, $\mathbf{P}$ is called **-inductive.*

COROLLARY. *For a pruning system* $\mathbf{P}$ *to be *-inductive, it is sufficient that it be a directed set. In particular, the structural *-topological system of* K *associated with the pruning system of all pruning sets of* K *is an inductive topological system.*

In general, it is possible that Ω^* is so uncontrollably complex or so inclusive that it deprives the theory relative to Ω^* of any substantial significance. Note, in particular, that a fertile phrase of a partial sentence can be a sterile phrase, because the label of its root can be the label of the root of some maximal phrase of a prepartial, nonpartial sentence. For "generative capacity" of a node may depend not simply on its label, but also on its contexts in a sentence it is contained in. If there exists a fertile phrase of a prepartial sentence which is a sterile phrase, Ω^* is infinite.

For the important class of context-free languages, the fertility of a phrase of a prepartial sentence is determined solely by its node. More specifically, Ω^* consists of phrases whose roots are certain nonterminal symbols which can "generate" only a finite number of phrases. Hence, Ω^* itself is finite. But a phrase whose root is a nonterminal symbol which can generate only a finite number of phrases may not be "sterile" in the sense we have defined. This is because it is possible that such a phrase is always contained in a larger "sterile" phrase when it appears in a sentence. But to see this it is convenient to introduce the notion of almost terminal symbol.

Let us divide the nonterminal symbols of a context-free grammar G into two subclasses, one consisting of nonterminals A such that there are infinitely many phrases of G whose root is labeled with A, and the other consisting of nonterminals B such that there are only finite numbers of phrases of G whose root is labeled with B. Let us call the former *fertile symbols* and the latter *almost terminal symbols*. Informally, fertile symbols may be considered as those with infinite "generative capacity," and almost terminal ones as those with finite "generative capacity." Then a maximal phrase of a prepartial sentence is fertile (and, hence, the prepartial sentence is a partial sentence) if and only if the root of the maximal phrase is labeled with a fertile symbol.

The relation between sterility and almost-terminality, however, is not quite straightforward due to certain marginal situations. Certainly the root of a sterile phrase must be labeled with an almost terminal symbol. In fact, all nonterminal nodes of a sterile phrase must be labeled with almost terminal symbols. But a phrase whose root is labeled with an almost terminal symbol may not be sterile. Assume A is an almost terminal symbol of a context-free grammar G. Assume furthermore that $B \rightarrow A$ is the only rule in which A appears on the right-hand side. Then a phrase whose root is labeled with A cannot be a maximal phrase of any prepartial sentence, and, hence, cannot be a sterile phrase. Moreover if B is fertile, a phrase whose root is labeled with A cannot be even a subphrase of a sterile phrase.

We shall later have occasions to refer to the notion of almost-terminality in the proof of Theorem 6 for technical reasons. Since I formulate Theorem 6 in a slightly more general form than applicable solely to context-free languages, I shall now extend the notion of almost terminal symbol for phrase-structure tree languages in general as follows. Let us call a terminal node p of a tree T belonging to a tree language K *almost terminal* if there exists only a finite number of phrases π such that there is a sentence σ, $T < \sigma$, and the phrase of σ at p is π. A nonterminal symbol A of K is called *almost terminal* if there is an almost terminal node of a tree labeled with A. Finally, a phrase of K is called *almost terminal* if its root is labeled with an almost terminal symbol. Then, as in the context-free case, sterile phrases are almost terminal.

Before closing this section let us make some general remarks on relativization of the theory. When the theory is not relativized, the difference between the heights of τ and $(\tau)_\Pi$ is bounded for any given pruning set Π. Also, given a tree T, the number of sentences τ such that $(\tau)_{\Pi} = T$ is bounded. These finitary properties follow from the finiteness of Π. If Ω, with respect to which we relativize the theory, is infinite, finitary properties like these may cease to hold. Hence, theorems proved in the absolute case may not be carried over to the relative case.

Consider Theorem 1, Section 2, which states that for any extension Π' of a pruning set Π, there exists an extension Π'' of Π such that $\mathbf{T}_{\Pi''}$ is weaker than $\mathbf{T}_{\Pi}$. Π'' is the set of phrases π'' such that there exists a phrase π' in Π' with the property $(\pi'')_{\Pi} < (\pi')_{\Pi}$. Now, if Ω is finite, we can apply the same theorem to $\Omega \cup \Pi$ and $\Omega \cup \Pi'$ and obtain Π''. The same procedure does not work for an infinite Ω.

However, if Ω consists of all the phrases whose root is labeled with a nonterminal belonging to some specified subset of the nonterminal vocabulary, then Theorem 1 hold for the theory relative to Ω, even if Ω is infinite. To see this, let Π, Π', Π'', and so on, be as in Theorem 1. The phrase π'' of σ at q is now either in Ω or in Π''. If it is in Π'', then the proof proceeds in the same way. Assume, then, that it is in Ω. But then the phrase π of τ at q is also in Ω; hence, it cannot properly dominate a node p in $(\tau)_{\Pi''}$. Hence, $(\tau)_{\Pi''}$ must be a rooted subtree of $(\sigma)_{\Pi''}$.

In particular, then, Theorem 1 holds for the theory relative to Ω^*, whether it is finite or not.

6. *Structural Continuity and Structural Equivalence*

Let K and K' be two phrase-structure tree languages and let f be a function from K to K'. Furthermore, let $\mathbf{P}$ and $\mathbf{P}'$ be pruning systems of K and K', respectively.

Assume, for the moment, that $\mathbf{P}$ and $\mathbf{P}'$ are inductive and let Π and Π' be pruning sets of $\mathbf{P}$ and $\mathbf{P}'$, respectively. Assume f is continuous from $\mathbf{T}_{\Pi}$ to $\mathbf{T}_{\Pi'}$. If $\bar{\Pi}'$ is an extension of Π' in $\mathbf{P}'$, f is continuous from $\mathbf{T}_{\Pi}$ to $\mathbf{T}_{\bar{\Pi}'}$. On the other hand, if $\bar{\Pi}$ is an extension of Π in $\mathbf{P}$, f may not be continuous from $\mathbf{T}_{\bar{\Pi}}$ to $\mathrm{T}_{\Pi'}$, but perhaps f might be continuous from $\mathbf{T}_{\bar{\Pi}}$ to $\mathbf{T}_{\bar{\Pi}'}$, if $\bar{\Pi}'$ is an appropriate extension of Π' so that $\mathbf{T}_{\bar{\Pi}'}$ is sufficiently weak. Hence, it is meaningful to ask whether for an arbitrary extension $\bar{\Pi}$ of Π, there exists an extension $\bar{\Pi}'$ of Π' such that f is continuous from $\mathbf{T}_{\bar{\Pi}}$ to $\mathbf{T}_{\bar{\Pi}'}$.

More generally, given pruning systems $\mathbf{P}$ and $\mathbf{P}'$ of K and K', respectively (whether they are inductive or not), we say that f is *phrase-structurally*, or, simply, *structurally continuous* from K to K' (or, *structurally* (*)- or *-*continuous*) with respect to $\mathbf{P}$ and $\mathbf{P}'$, if for any pruning set Π of $\mathbf{P}$, there exists a pruning set Π' of $\mathbf{P}'$ such that f is continuous from $\mathbf{T}_{\Pi}$ to $\mathbf{T}_{\Pi'}$ (or $\mathbf{T}_{\Pi}^{(*)}$ to $\mathbf{T}_{\Pi'}^{(*)}$, $\mathbf{T}_{\Pi}{}^{*}$ to $\mathbf{T}_{\Pi'}^{*}$).

A one-to-one map f from K onto K' is called a *phrase-structural* or, simply, *structural homeomorphism* (or *structural* (*)- or *-*homeomorphism*) with respect to $\mathbf{P}$ and $\mathbf{P}'$, if f and f^{-1} are structurally continuous (or (*)- or *-continuous) with respect to $\mathbf{P}$ and $\mathbf{P}'$ and with respect to $\mathbf{P}'$ and $\mathbf{P}$, respectively.

In what follows, we shall deal mostly with the case in which $\mathbf{P}$ and $\mathbf{P}'$ are the pruning systems of all pruning sets of K and K', and we will not mention $\mathbf{P}$ and $\mathbf{P}'$ explicitly when they are those systems.

K and K' are said to be *phrase-structurally homeomorphic* ((*)- or *-*homeomorphic*) if there exists a structural homeomorphism ((*)- or *-homeomorphism) from K onto K'.

Although in the present work we shall not be concerned with the study of phrase-structure string languages, some definitions concerning them might be of interest in relating the present theory to some familiar notions in algebraic linguistics.

Let K be a tree language and L its surface string language. For each pruning set Π of K, we can define topologies $\mathbf{T}_{\Pi}(L)$, $\mathbf{T}_{\Pi}^{(*)}(L)$, and $\mathbf{T}_{\Pi}{}^*(L)$ on L as the strongest topologies that make the yield function continuous from $\mathbf{T}_{\Pi}(K)$, $\mathbf{T}_{\Pi}^{(*)}(K)$, and $\mathbf{T}_{\Pi}{}^*(K)$, respectively. Let K' be another tree language and let L' be its surface string language. Let f be a function from K to K'. Denote by η and η' the yield functions of K and K', respectively. If a function φ from L to L' satisfies the relation $\varphi \cdot \eta = \eta' \cdot f$, φ is said to be *induced* by f and may with no likelihood of confusion be identified with f. With this convention, if f is continuous from $\mathbf{T}_{\Pi}(K)$ to $\mathbf{T}_{\Pi'}(K')$ (or $\mathbf{T}_{\Pi}^{(*)}(K)$ to $\mathbf{T}_{\Pi'}^{(*)}(K')$, $\mathbf{T}_{\Pi}{}^*(K)$ to $\mathbf{T}_{\Pi'}^{*}(K')$), it is continuous from $\mathbf{T}_{\Pi}(L)$ to $\mathbf{T}_{\Pi'}(L')$ (or $\mathbf{T}_{\Pi}^{(*)}(L)$ to $\mathbf{T}_{\Pi}^{(*)}(L')$, $\mathbf{T}_{\Pi}{}^*(L)$ to $\mathbf{T}_{\Pi'}^{*}(L')$). In particular, if $K = K'$ and f is the identity map of K, f induces a map in L which is the identity of L. Hence, the analogs of Theorems 1, 3, and 4 hold for the phrase-structure string language (L, K).

A function φ from L to L' is said to be *structurally continuous* ((*)- or *-*continuous*) *from* (L, K) *to* (L', K') if for any pruning set Π of K there is a pruning set Π' of K' such that φ is continuous from $\mathbf{T}_{\Pi}(L)$ to $\mathbf{T}_{\Pi'}(L)$ ($\mathbf{T}_{\Pi}^{(*)}(L)$ to $\mathbf{T}_{\Pi'}^{(*)}(L')$, $\mathbf{T}_{\Pi}{}^*(L)$ to $\mathbf{T}_{\Pi'}^{*}(L')$). If a function f from K to K' is structurally continuous ((*)- or *-continuous), and if it induces a map from L to L', it is structurally continuous ((*)- or *-continuous) from (L, K) to (L', K').

Finally, the following concepts may be introduced as generalizations of Chomsky's notion of strong equivalence of grammars. Let G and G' be two grammars (finitary devices like context-free grammars) that generate tree languages K and K', respectively, such that K and K' yield the same string language L. Thus, G and G' are weakly equivalent in Chomsky's sense. G and G' are said to be *structurally equivalent* ((*)- or *-*equivalent*) if the identity map of L is a structural homeomorphism ((*)- or *-homeomorphism) from (L, K) onto (L, K').

Obviously, this is not the only possible generalization of Chomsky's strong equivalence of grammars. Phrase-structure tree languages (or phrase-structure string languages) may be classified into structurally homeomorphic ((*)- or *-homeomorphic) types and each such classification defines an equivalence of grammars. But the equivalence of grammars defined in the preceding paragraph is in a sense more interesting, as it is not an automatic consequence of homeomorphic language types. Also, it could be said to be a true generalization of Chomsky's strong equivalence in that it gives a subclassification of weakly equivalent grammars.

All the notions defined above may be understood either absolutely or relatively. That is, for each tree language K a specific set of phrases $\Omega(K)$ may be given, with respect to which the theory is relativized. Following the convention introduced in the preceding section, if the topologies of partial sentences are involved in a discussion, we assume that we are dealing with the theory relativized with respect to the set of all sterile phrases for each language under discussion.

More generally, we can (and in fact, it is more convenient to assume that we) relativize the theory with respect to certain equivalence classes of sets of phrases. Let Ω_1 and Ω_2 be two sets of phrases of K. Ω_1 and Ω_2 are said to be *equivalent* with respect to a pruning system $\mathbf{P}$ of K if the identity map of K is structurally homeomorphic with respect to $\mathbf{P}$ from K relativized with Ω_1 onto K relativized with Ω_2. Let us restrict ourselves to the case where $\mathbf{P}$ is the set of all pruning sets. Now, if a map f is structurally ((*)-, or *-) continuous from K relativized with Ω_1 to K' relativized with Ω_1' and if Ω_2 and Ω_2' are quivalent to Ω_1 and Ω_1', respectively, then f is structurally ((*)- or *-) continuous from K relativized with Ω_2 to K' relativized with Ω_2'. In particular, if K relativized with Ω_1 is structurally ((*)- or *-) homeomorphic to K' relativized with Ω_1', K relativized with Ω_2 is structurally ((*)-, *-) homeomorphic to K' relativized with Ω_2'. This means that as far as invariants under structural ((*)-, *-) homeomorphism are concerned, equivalent sets of phrases may be substituted for each other freely.

Assume that Ω_0 is a set of phrases of K such that Theorem 1 holds in the theory relative to Ω_0. Assume further that $\Omega \supset \Omega_0$ and the difference between Ω and Ω_0 is finite. Then Ω is equivalent to Ω_0. To see this, let Π be an arbitrary pruning set, and put $\Pi_0 = \Pi \cup (\Omega - \Omega_0)$; Π_0 is also a pruning set. The preorder $<_\Pi$ defined relative to Ω is nothing but the preorder $<_{\Pi_0}$ defined relative to Ω_0. Hence, the identity of K is homeomorphic from $\mathbf{T}_\Pi$ relative to Ω onto $\mathbf{T}_{\Pi_0}$ relative to Ω_0. Conversely, let Π_1 be an extension of

Π_0 such that the identity of K is continuous from $\mathbf{T}_\Pi$ to $\mathbf{T}_{\Pi_1}$, relative to Ω_0. Since $\Pi_1 \supset \Omega - \Omega_0$, the identity map is homeomorphic from $\mathbf{T}_{\Pi_1}$ relative to Ω onto $\mathbf{T}_{\Pi_1}$ relative to Ω_0; hence, it is continuous from $\mathbf{T}_\Pi$ relative to Ω_0 onto $\mathbf{T}_{\Pi_1}$ relative to Ω. This proves our contention. Furthermore, Theorem 1 holds relative to Ω. For, let Π and Π' be two pruning sets such that $\Pi \subset \Pi'$, and put, as above, $\Pi_0 = \Pi \cup (\Omega - \Omega_0)$. Put $\Pi_0' = \Pi' \cup \Pi_0$ and let Π_0'' be a extension of Π_0' such that the identity of K is continuous from $\mathbf{T}_{\Pi_0}$ to $\mathbf{T}_{\Pi_0''}$ relative to Ω_0. Since $\Pi''_0 \supset \Omega - \Omega_0$, the identity of K is homeomorphic from $\mathbf{T}_{\Pi_0''}$ relative to Ω_0 onto $\mathbf{T}_{\Pi_0''}$ relative to Ω. Consequently, it is continuous from $\mathbf{T}_\Pi$ onto $\mathbf{T}_{\Pi_0''}$ relative to Ω, and Π_0'' is an extension of Π'.

From the above observation, there follows, in particular, the following.

THEOREM 5. *If K is context-free, the set of almost terminal phrases and the set of sterile phrases are both equivalent to the void set of phrases. In the theory relative to each of these theories Theorem* 1 *holds.*

Parenthetically, let us consider the difference between structural continuity, as defined above, and the continuity with respect to inductive limits of inductive systems of topologies. Let $\mathbf{P}$ be an inductive ((*)-inductive, *-inductive) pruning system. Then, one can define the inductive limit $\mathbf{T}(\mathbf{T}^{(*)}, \mathbf{T}^*)$ of the structural ((*)-, *-) topological system associated with $\mathbf{P}$. Assume $\mathbf{P}$ is maximal and "converges" to the set of all phrases of K. Then, for each sentence τ in K, $\mathbf{T}(\mathbf{T}^{(*)}, \mathbf{T}^*)$ is locally homeomorphic at τ to $\mathbf{T}_\Pi(\mathbf{T}_\Pi^{(*)}, \mathbf{T}_\Pi{}^*)$, for some Π in $\mathbf{P}$. For there exists a pruning set Π in $\mathbf{P}$ which contains τ, and for such Π, the smallest neighborhood of τ in $\mathbf{T}_\Pi$ is the set of all sentences whose root is labeled with the same symbol as τ. In particular, if K is uni-rooted, this set is K itself; $\mathbf{T}$ is the weakest topology of K ($\mathbf{T}^{(*)}$ and $\mathbf{T}^*$ are the topologies of prepartial and partial sentences of K, justifying our notation). Let $\mathbf{P}'$ be a maximal inductive ((*)-, *-inductive) pruning system of K' that "converges" to the set of all phrases of K', and let $\mathbf{T}'(\mathbf{T}'^{(*)}, \mathbf{T}'^*)$ be the inductive limit of the structural ((*)-, *-) topological system associated with $\mathbf{P}'$. Let f be a map from K to K'. f naturally induces a map from $\mathbf{T}$ to $\mathbf{T}'$ ($\mathbf{T}^{(*)}$ to $\mathbf{T}'^{(*)}$, $\mathbf{T}^*$ to $\mathbf{T}'^*$). Since everywhere $\mathbf{T}(\mathbf{T}^{(*)}, \mathbf{T}^*)$ is locally homeomorphic to some $\mathbf{T}_\Pi(\mathrm{T}_\Pi^{(*)}, \mathbf{T}_\Pi{}^*)$ and everywhere $\mathbf{T}'(\mathbf{T}'^{(*)}, \mathbf{T}'^*)$ is locally homeomorphic to some $\mathbf{T}_{\Pi'}(\mathrm{T}_{\Pi'}^{(*)}, \mathbf{T}_{\Pi'})$, f is continuous from $\mathbf{T}$ to $\mathbf{T}'$ ($\mathbf{T}^{(*)}$ to $\mathbf{T}'^{(*)}$, $\mathbf{T}^*$ to $\mathbf{T}'^*$) at τ, if and only if for any Π in $\mathbf{P}$ there is a Π' in $\mathbf{P}'$ such that f is continuous from $\mathbf{T}_\Pi$ to $\mathbf{T}_{\Pi'}$ ($\mathbf{T}_\Pi^{(*)}$ to $\mathbf{T}_{\Pi'}^{(*)}$, $\mathbf{T}_\Pi{}^*$ to $\mathbf{T}_{\Pi'}^*$) at τ. The apparent similarity of this condition with the definition of the structural ((*)-, *-) continuity of f, however, does not imply their equivalence. The crucial difference between them lies in the fact that in the former case for each Π the existence of Π' such that f is continuous from $\mathbf{T}_\Pi$ to $\mathbf{T}_{\Pi'}$

($\mathbf{T}_{\Pi}^{(*)}$ to $\mathbf{T}_{\Pi'}^{(*)}$, $\mathbf{T}_{\Pi}{}^{*}$ to $\mathbf{T}_{\Pi'}^{*}$) is required only "pointwise," while in the latter "globally" or "uniformly." Indeed, from what is said above, it follows directly that when K and K' are uni-rooted, and $\mathbf{P}$ and $\mathbf{P}'$ maximal, the continuity of f from the inductive limit with respect to $\mathbf{P}$ to that with respect to $\mathbf{P}'$ means nothing more than the continuity of f from the weakest topology of K to that of K' (the topology of prepartial sentences of K to that of K', the topology of partial sentences of K to that of K'), an uninteresting condition.

EXAMPLE. (Chomsky normal form.) A context-free grammar is said to be in the Chomsky normal form if all rules are of the form $A \to BC$ or $A \to a$, where A, B, and C are nonterminals and a is a terminal. It is well known that any ϵ-free context-free grammar G is weakly equivalent (in Chomsky's sense) to a Chomsky normal grammar (cf. (Chomsky, 1959; Hopcroft and Ullman, 1969)). We shall sketch a demonstration that a context-free grammar can be normalized within structural equivalence.

First assume that G does not have a rule of the form $A \to B$, where A and B are nonterminals. Iteration of the following two processes converts G into a Chomsky normal grammar.

(1) Let $A \to \alpha_1\alpha_2 \cdots \alpha_m$, $m \geqslant 2$ be a rule of a given grammar, where α_i, $1 \leqslant i \leqslant m$, is either a terminal or nonterminal. Replace this rule by rules $A \to B_1B_2 \cdots B_m$ and $B_i \to \alpha_i$ for each i such that α_i is a terminal, where $B_j = \alpha_j$, $1 \leqslant j \leqslant m$, if α_j is a nonterminal and B_j is a new nonterminal if α_j is a terminal.

(2) Let $A \to B_1B_2 \cdots B_m$, $m > 2$, be a rule of a given grammar, where B_i, $1 \leqslant i \leqslant m$, is a nonterminal. Replace this rule by $A \to CB_3 \cdots B_m$ and $C \to B_1B_2$, where C is a new nonterminal.

Each of these processes converts a given grammar G_1 into another, G_2, which is structurally equivalent to G_1.

Assume first that G_2 is obtained from G_1 by one application of process (1). Define a function f from the phrases of G_2 onto the phrases of G_1 as follows: If a phrase π of G_2 contains a branch of the form $B_i(\alpha_i)$, where $B_i \to \alpha_i$ is a new rule in G_2, replace it by α_i and define the tree thus obtained as $f(\pi)$; otherwise put $f(\pi) = \pi$. The restriction of f to $K_2 = K(G_2)$ is one-to-one and onto $K_1 = K(G_1)$. For an arbitrary pruning set Π_2 of K_2, put

$$\Pi_1 = f(\Pi_2) = \{\pi; \pi = f(\pi'), \pi' \in \Pi_2\};$$

then, f is continuous (in fact, homeomorphic) from $\mathbf{T}_{\Pi_2}$ onto $\mathbf{T}_{\Pi_1}$. Conversely if an arbitrary pruning set Π_1 of K_1 is given, put

$$\Pi_2 = f^{-1}(\Pi_1) = \{\pi; f(\pi) \in \Pi_1\};$$

then, f is continuous (in fact, homeomorphic) from $\mathbf{T}_{\Pi_2}$ onto $\mathbf{T}_{\Pi_1}$. Hence, f is structurally homeomorphic and G_2 is structurally equivalent to G_1.

Next, assume that G_2 is obtained from G_1 by an application of process (2). Define a function f from the phrases of G_1 to those of G_2 inductively as follows: If $\pi = A(\pi_1, \pi_2, ..., \pi_m)$, where each π_i, $1 \leqslant i \leqslant m$, is a phrase whose root is B_i, put $f(\pi) = A(\pi_0, \pi_3, ..., \pi_m)$ where $\pi_0 = C(\pi_1, \pi_2)$; otherwise for $\pi = B(\pi_1, \pi_2, ..., \pi_n)$, put $f(\pi) = B(f(\pi_1), f(\pi_2), ..., f(\pi_n))$. The restriction of f to $K_1 = K(G_1)$ is one-to-one and onto $K_2 = K(G_2)$. Let Π_1 be an arbitrary pruning set of G_1; put $\Pi_2 = f(\Pi_1)$. Then f is homeomorphic from $\mathbf{T}_{\Pi_1}$ onto $\mathbf{T}_{\Pi_2}$. Conversely, assume Π_2 is an arbitrary pruning set of G_2; put $\Pi_1 = f^{-1}(\Pi_2)$ if Π_2 does not contain a phrase of the form $\pi = C(\pi_1, \pi_2)$, where the roots of π_1 and π_2 are B_1 and B_2, respectively; if Π_2 does contain such a phrase, let Π_1 be the union of $f^{-1}(\Pi_2)$ and the set of phrases $f^{-1}(\pi_1)$ and $f^{-1}(\pi_2)$ such that $\pi = C(\pi_1, \pi_2)$ is in Π_2. Then f^{-1} is homeomorphic from $\mathbf{T}_{\Pi_2}$ onto $\mathbf{T}_{\Pi_1}$. It follows from the above results that f is a structural homeomorphism; G_1 and G_2 are structurally equivalent.

From the above we can conclude that if an ϵ-free context-free grammar G does not contain a rule of the form $A \to B$, where A and B are nonterminals, there exists a Chomsky normal grammar G' which is structurally equivalent to G. In fact, there exists a structural homeomorphism from $K(G)$ onto $K(G')$, which induces the identity map on $L(G) = L(G')$.

But if G has a rule of the form $A \to B$, normalization of G gives rise to a somewhat different situation. For each sequence of rules of G, $A \to B_1$, $B_1 \to B_2, ..., B_n \to C$, and a rule $C \to \omega$, where ω is either a terminal or a string of length more than one, introduce a new rule $A \to \omega$. Remove all rules of G of the form $A \to B$. The grammar G' obtained from G in this way is weakly equivalent to G and free of rules of the form $A \to B$. Let f be a function on the phrases of G defined inductively as follows: if π is a terminal, put $f(\pi) = \pi$; if $\pi = A(\pi_0)$, where π_0 is not a terminal, and if $f(\pi_0) = B(\pi_1, \pi_2, ..., \pi_n)$, put $f(\pi) = A(\pi_1, \pi_2, ..., \pi_n)$; otherwise, for $\pi = A(\pi_1, \pi_2, ..., \pi_n)$ put $f(\pi) = A(f(\pi_1), f(\pi_2), ..., f(\pi_n))$. The $f(\pi)$ defined this way may not be a phrase of G or G', but the restriction of f to $K = K(G)$ is a function onto $K' = K(G')$. However, in general, f may not be one-to-one from K onto K'. (That is, G is more ambiguous than G'.) Yet the identity map of $L = L(G) = L(G')$, which is induced on L by f, is a structural homeomorphism.

To see this, let us first introduce the following notions. Let us call a phrase π of K *normal* if π is a terminal symbol or if its root directly dominates more than one node. (In other words, π is normal if it is of height 0 or else node 1 has at least one sister.) In general, given a phrase $\bar{\pi}$, we can write

$\bar{\pi} = B_1(B_2(\cdots (B_n(\pi)) \cdots)$, $n \geqslant 0$, where $B_i \to B_{i+1}$, $1 \leqslant i < n$, is a rule of G and π is normal. (By convention, if $n = 0$, we understand that $\bar{\pi} = \pi$.) We call π the *normal kernel* of $\bar{\pi}$, and we call $\bar{\pi}$, a *nonbranching extension* of π. $f(\bar{\pi})$ is a phrase of K' if there is a rule in G of the form $A \to \omega_1 B_1 \omega_2$, $\omega_1\omega_2 \neq \epsilon$. Note that given π there may be infinitely many $\bar{\pi}$ such that π is the normal kernel of $\bar{\pi}$. However, there can exist only a finite number of different $f(\bar{\pi})$'s; for $f(\bar{\pi})$ is determined by its root B_1 and its kernel π and there is only a finite number of nonterminals.

Now, let Π be an arbitrary pruning set of K. Define a pruning set Π' of K' as the set of all $f(\bar{\pi})$'s which are phrases of K', where $\bar{\pi}$ is a nonbranching extension of the normal kernel of a phrase of Π. Then f is continuous from $\mathbf{T}_{\Pi}$ onto $\mathbf{T}_{\Pi'}$, and, hence, structurally continuous from K onto K'. The identity of L, which is induced by f on L, is structurally continuous from $L = L(G)$ onto $L = L(G')$.

In order to see, conversely, that the identity of L is structurally continuous from $L = L(G')$ onto $L = L(G)$, let Π' be an arbitrary pruning set of K' and take a pruning set Π of K such that $f(\Pi) = \Pi'$. Let τ' be an arbitrary sentence of K'. Assume that a phrase π' of Π' is pruned from τ' to obtain $(\tau')_{\Pi'}$. Then there are τ in K and π in Π such that $f(\tau) = \tau'$ and π is pruned from τ to obtain $(\tau)_{\Pi}$. That is, $f(\tau) = \tau'$, and if $(\tau')_{\Pi'} < (\sigma')_{\Pi'}$, then there is σ such that $(\tau)_{\Pi} < (\sigma)_{\Pi}$ and $f(\sigma) = \sigma'$. Now, take an open set E of $\mathbf{T}_{\Pi}(L(G))$, and let x be a sentence in E. Assume that $\eta'(\tau') = x$, where η' is the yield function of K'. Let τ, σ, σ' be determined as above. Then, $\sigma \in \eta^{-1}(E)$ where η is the yield function of K. Hence, $\sigma' = f(\sigma) \in f\eta^{-1}(E) = \eta'^{-1}(E)$. It follows that $V_{\Pi'}(\tau') \subset \eta'^{-1}(E)$. This relation can be considered to hold for an arbitrary element of $\eta'^{-1}(E)$. This means that $\eta'^{-1}(E)$ is open in K', and, consequently, E is open in $\mathbf{T}_{\Pi'}(L(G'))$. The identity of $L = L(G) = L(G')$ is thus continuous from $T_{\Pi'}(L)$ to $T_{\Pi}(L)$. Since Π' is an arbitrary pruning set of G', it follows that the identity of L is structurally continuous from $L = L(G')$ onto $L = L(G)$. This completes the proof that G and G' are structurally equivalent.

We now conclude: Any ϵ-free context-free grammar is structurally equivalent to a Chomsky-normal grammar.

For an example of structurally *-equivalent, but not structurally equivalent context-free grammars, see Part II, Section 6, Example 1.

7. *Context-Free Languages, Weak and Strong Structural Equivalence*

The objective of this section is to see that for context-free languages structural homeomorphism implies structural *-homeomorphism. Thus, it might be said, at least for context-free languages, that topologies $\mathbf{T}_{\Pi}{}^*$ are strong enough to make the system of topologies $\mathbf{T}_{\Pi}{}^*$ for the entire class of pruning sets inductive, but yet weak enough to be compatible with the structure equivalence defined in terms of topologies $\mathbf{T}_{\Pi}$.

Looking ahead to this result, it might be useful to introduce the following terminology for the class of context-free languages. We call a function from a context-free language onto another a *strong structural homeomorphism* (*weak structural homeomorphism*) if it is structural homeomorphism (structural $*$-homeomorphism). We use the similar terminology for context-free string languages and for context-free grammars. Thus, context-free tree (or string) languages are classified into weak and strong homeomorphic types, the latter being subclassification of the former. Context-free grammars may be classified according to these homeomorphic types of tree (or string) languages. Finally, weakly equivalent grammars (in Chomsky's sense) are classified according to weak and strong structural equivalence.

To state our theorem in a somewhat more general form than indicated above, and also for purposes of proof, we shall introduce some notions. Let K be a phrase-structure tree language. K is said to satisfy the *finite branching condition* if there is an upper bound for the number of sisters a node of a phrase belonging to K can have. K is said to be *lexically finite* if the terminal and the nonterminal vocabulary of K are both finite. If K is lexically finite and satisfies the finite branching condition, then for any tree T and any integer h there exist at most a finite number of trees belonging to K of which T is a subtree and whose height relative to T is bounded by h.

Note that a context-free language is lexically finite and satisfies the finite branching condition. Note also that there are only a finite number of almost terminal phrases in a context-free language.

THEOREM 6. *Let K be a lexically finite phrase-structure tree language satisfying the finite branching condition and assume that its set of almost terminal phrases is finite. Let K' be context-free. Then, if f is a structural homeomorphism from K onto K', f is structurally $*$-continuous from K to K'.*

Note that for either K or K', the set of all sterile phrases, being a subset of the set of almost terminal phrases, is also finite. Note also that a node in a sterile phrase of a prepartial sentence is almost terminal, because the set of sterile phrases is finite. From the theorem it follows, in particular, the

COROLLARY. *If K and K' are context-free and structurally homeomorphic, they are structurally $*$-homeomorphic.*

We shall first prove the following

LEMMA. *Let K' be context-free and let g be a structurally continuous function from K' to K. If Π' is a sufficiently large pruning set of K' and $\overline{\Pi}$ is a pruning set of K such that g is continuous from $\mathbf{T}_{\Pi'}$ to $\mathbf{T}_{\overline{\Pi}}$, then for any tree T' belonging to K', there exist a finite number of sentences of K, σ_j, $1 \leqslant j \leqslant l$, such that $\sigma_j = g(\sigma_j')$, $\sigma_j' \in \overline{V}_{\Pi'}(T')$ and $g(\overline{V}_{\Pi'}(T')) \subset \bigcup_{1\leqslant j\leqslant l} V_{\overline{\Pi}}(\sigma_j)$.*

In fact, if Π' is so large that for any nonterminal A' of K', there exists at least one phrase in Π' whose root is labeled with A', then the lemma holds with respect to Π'. (Note that such Π' indeed exists, since the nonterminal vocabulary of K' is finite.)

Let h' be the maximum of the heights of the phrases in Π'. Take an arbitrary element σ' in $\overline{V}_{\Pi'}(T')$ and consider $(\sigma')_{\Pi'}$. If the height of $(\sigma')_{\Pi'}$ relative to T' is not greater than h', put $T'_{\sigma'} = (\sigma')_{\Pi'}$. If it is greater than h', for each node p of $(\sigma')_{\Pi'}$ whose distance from T' is h', replace the phrase of σ' at p by a phrase in Π' whose label has the same label as p in σ'. Name $\hat{\sigma}'$ the tree obtained from σ' in this way and put $T'_{\sigma'} = (\hat{\sigma}')_{\Pi'}$. Note that since K' is context-free, $\hat{\sigma}'$ is a sentence of K'. Note also that T' is a rooted subtree of $T'_{\sigma'}$, the height of $T'_{\sigma'}$ relative to T' is not greater than h', and $\sigma' \in \overline{V}_{\Pi'}(T'_{\sigma'})$.

If we take all trees T_j', $1 \leqslant j \leqslant l$, of which T' is a subtree and whose heights relative to T' is not greater than h' and such that $T_j' = (\sigma_j')_{\Pi'}$ for some σ_j' in K', then for any σ' in $\overline{V}_{\Pi'}(T')$, $T'_{\sigma'}$ is identical to one of these T_j', $1 \leqslant j \leqslant l$. Hence, we have

$$\overline{V}_{\Pi'}(T') = \bigcup_{1\leqslant j\leqslant l} V_{\Pi'}(\sigma_j') \qquad \text{and} \qquad g(\overline{V}_{\Pi'}(T')) = \bigcup_{1\leqslant j\leqslant l} g(V_{\Pi'}(\sigma_j')).$$

Now, if g is continuous from $\mathbf{T}_{\Pi'}$ to $\mathbf{T}_{\overline{\Pi}}$, we have $g(V_{\Pi'}(\sigma_j')) \subset V_{\overline{\Pi}}(\sigma_j)$, where we put $\sigma_j = g(\sigma_j')$. Hence,

$$g(\overline{V}_{\Pi'}(T')) \subset \bigcup_{1\leqslant j\leqslant l} V_{\overline{\Pi}}(\sigma_j).$$

We now proceed to prove the theorem. Let f be a structural homeomorphism from K onto K'. We shall derive a contradiction from the assumption that f is not structurally *-continuous from K to K'.

If f were not structurally *-continuous from K to K', there would exist a pruning set Π of K such that for any pruning set Π' of K', f is not continuous from $\mathbf{T}_{\Pi}{}^*$ to $\mathbf{T}^*_{\Pi'}$. Since f is structurally continuous from K to K', we may take Π' such that f is continuous from $\mathbf{T}_{\Pi}$ to $\mathbf{T}_{\Pi'}$. We may furthermore assume that Π' is large enough to make the preceding lemma valid with respect to $g = f^{-1}$ (cf. Theorem 1). Since f is continuous from $\mathbf{T}_{\Pi}$ to $\mathbf{T}_{\Pi'}$, but not from $\mathbf{T}_{\Pi}{}^*$ to $T^*_{\Pi'}$, it is not continuous from $\mathbf{T}_{\Pi}{}^*$ to $\mathbf{T}'^*$, $\mathbf{T}'^* = \mathbf{T}^*(K')$

being the topology of partial sentences of K'. Hence, there exists a partial sentence $\bar{\sigma}'$ of K' such that the inverse image D by f of its closure $D' = C(\bar{\sigma}')$ is not closed in $\mathbf{T}_{\Pi}{}^*$. Consequently, there is τ in the complement of D such that any of its neighborhoods in $\mathbf{T}_{\Pi}{}^*$ meets D. In particular, if C_i, $1 \leqslant i \leqslant k$, are closures of partial sentences that do not contain τ, concerning their complements E_i we have

$$V_{\Pi}(\tau) \cap E \cap D \neq \varnothing, \qquad \text{where} \qquad E = \bigcap_{1 \leqslant i \leqslant k} E_i\,. \tag{1}$$

Let σ be a sentence in $V_{\Pi}(\tau)$ such that $\sigma \neq \tau$. The closure of σ is the singleton set $\{\sigma\}$, which does not contain τ. From this, it follows that $V_{\Pi}(\tau)$ cannot be finite; if it were, let C_i above be the closures of the sentences in $V_{\Pi}(\tau)$ different from τ, and we would get $V_{\Pi}(\tau) \cap E \cap D = \{\tau\} \cap D \neq \varnothing$; but τ is assumed not to belong to D. Hence, we can now assume that $V_{\Pi}(\tau)$ is not finite.

Consider partial sentences $\bar{\sigma}$ of K that satisfy the following conditions:

(2.1) $C(\bar{\sigma})$ does not contain τ;

(2.2) $V_{\Pi}(\tau)$ and $C(\bar{\sigma})$ meet.

From (2.2) it follows that $(\tau)_{\Pi}$ and $\bar{\sigma}$ have intersection and union. We define $h_{\bar{\sigma}}$ as the maximum distance from a terminal node of a phrase of $\bar{\sigma}$ to the intersection of $(\tau)_{\Pi}$ and $\bar{\sigma}$.

We shall contend that for each natural number h, there exist a finite number of partial sentences $\bar{\sigma}_i{}^h$, $1 \leqslant i \leqslant k(h)$, such that each $\bar{\sigma}_i{}^h$ satisfies (2) and such that if a partial sentence $\bar{\sigma}$ satisfies (2) and $h_{\bar{\sigma}} \leqslant h$, then for some i, $C(\bar{\sigma}) \subset C(\bar{\sigma}_i{}^h)$. Consider such a $\bar{\sigma}$. Let $T_{\bar{\sigma}}$ be the intersection of $\bar{\sigma}$ and $(\tau)_{\Pi}$. We need to consider two cases: (a) Assume that the distance between $T_{\bar{\sigma}}$ and a terminal node of $\bar{\sigma}$ which is not labeled with a terminal symbol (i.e., a terminal node of $\bar{\sigma}$ which is not a terminal node of a phrase of $\bar{\sigma}$) is at most the maximum of the heights of phrases in Π and Ω^*. (We call this maximum the height of Π relative to Ω^*.) From the finite branching condition there can exist only a finite number of such $\bar{\sigma}$'s. (b) Otherwise, there is a node p of $\bar{\sigma}$ which is not the root of a phrase of $\bar{\sigma}$ and whose distance from $T_{\bar{\sigma}}$ is equal to the height of Π relative to Ω^*. Let $\hat{\sigma}$ be the tree obtained from $\bar{\sigma}$ by pruning the full branch of $\bar{\sigma}$ at p. $\hat{\sigma}$ is a partial sentence which satisfies (2.1). For if $C(\hat{\sigma})$ contained τ, p would be a nonterminal node of τ whose distance from $T_{\bar{\sigma}}$ is equal to the height of Π relative to Ω^*, which is impossible because this distance would be the same as the distance of p from $(\tau)_{\Pi}$. $\hat{\sigma}$ also satisfies (2.2). For obviously $C(\bar{\sigma}) \subset C(\hat{\sigma})$, and $C(\bar{\sigma})$ meets $V_{\Pi}(\tau)$. Furthermore, $h_{\hat{\sigma}} = h_{\bar{\sigma}} \leqslant h$. Repeating this procedure if necessary, we shall obtain $\hat{\sigma}$ satisfying (2), $h_{\hat{\sigma}} \leqslant h$, and $C(\bar{\sigma}) \subset C(\hat{\sigma})$ such that the distance between $(\tau)_\Pi$ and a terminal node of $\hat{\sigma}$ which is not labeled with a terminal symbol is at most the height of Π relative to Ω^*; i.e., $\hat{\sigma}$ is one of the $\bar{\sigma}$'s in case (a). Thus, our contention follows.

Let E_i^h be the complement of $C(\bar{\sigma}_i^h)$, $1 \leqslant i \leqslant k(h)$, and put

$$E_h = \bigcap_{1 \leqslant i \leqslant k(h)} E_i^h.$$

From (1) we have $F_h = V_\Pi(\tau) \cap E_h \cap D \neq \varnothing$. Let us note that if $\sigma \in V_\Pi(\tau)$, $\sigma \neq \tau$ and the height of σ relative to $(\tau)_\Pi$ is not greater that h, then σ is not in E_h; for σ is itself a partial sentence such that $h_\sigma \leqslant h$ and the closure of σ is $\{\sigma\}$.

It follows that the sequence $(F_i)_{i \geqslant 0}$ is infinitely decreasing and that $\bigcap_{i \geqslant 0} F_i$ is void. Hence, if we choose a sentence σ_i from each F_i, the sequence $(\sigma_i)_{i \geqslant 0}$ is essentially infinite, that is, there are infinitely many different sentences among the σ_i's.

Next, we shall construct a sequence of trees $(T_r)_{r \geqslant 0}$ that satisfies the following conditions:

(3.1) For each r, T_r is a rooted subtree of T_{r+1};

(3.2) $\bar{V}_\Pi(T_r)$ contains infinitely many σ_i's;

(3.3) if a terminal node of T_r is terminal in T_{r+1}, it is labeled with an almost terminal symbol;

(3.4) $| T_r |$ is not bounded.

For T_0, we put $T_0 = (\tau)_\Pi$; all the σ_i's belong to $\bar{V}_\Pi(T_0)$. Assume that for some k we have defined T_r, $1 \leqslant r \leqslant k$, which satisfy (3.1)–(3.3), and also assume $| T_k | - | T_0 | = k$. Consider all the trees T such that

(4.1) T_k is a rooted subtree of T;

(4.2) $| T | - | T_k | = 1$;

(4.3) $\bar{V}_\Pi(T) \neq \varnothing$;

(4.4) a terminal node of T_k which is dominated by a sterile node is terminal in T;

(4.5) if a terminal node of T_k is terminal in T, it is labeled with an almost terminal symbol.

From the finiteness of the vocabulary, the finite branching condition, and (4.2), there are only a finite number of such T's. On the other hand, there must exist at least one such T. First of all, since $\bar{V}_\Pi(T_k)$ is infinite and Ω^* is finite (recall that we deal with the theory relative to Ω^*) there obviously exists a T that satisfies (4.1)–(4.4). Assume that a terminal node p of T_k is not labeled with an almost terminal symbol and yet is terminal in T. Then there are infinitely many sentences σ's such that $T < \sigma$ and the height of the phrase of σ at p is not bounded. Hence, there exists among them a σ such that $T < (\sigma)_\Pi$ and p is not terminal in $(\sigma)_\Pi$. Obtain a tree from T by grafting a branch of $(\sigma)_\Pi$ at p of height 1, and rename it T; this new T still satisfies (4.1)–(4.4). Repeating this process, if necessary, we finally attain a tree T satisfying (4.1)–(4.5). Let T_j', $1 \leqslant j \leqslant k'$ be the trees satisfying (4.1)–(4.5). Assume now that a sentence σ in $\bar{V}_\Pi(T_k)$ does not belong to any of $\bar{V}_\Pi(T_j')$, $1 \leqslant j \leqslant k'$. Then there exists a terminal node p of $(\sigma)_\Pi$ which is a terminal node of T_k not dominated by a sterile node. Consider the tree $\hat{\sigma}$ obtained from T_k by grafting at p the phrase of σ at p. This is a partial sentence. Note that $C(\hat{\sigma})$ meets $V_\Pi(\tau)$, since σ belongs to both. Hence, $h_{\hat{\sigma}}$ is defined and it is less than k plus the height of Π relative to Ω^*. Let h be this number. Then σ is contained in one of $C(\bar{\sigma}_i^h)$, $1 \leqslant i \leqslant k(h)$, $\bar{\sigma}_i^h$ being as defined earlier. In sum, although there may be infinitely many σ in $\bar{V}_\Pi(T_k)$ which do not belong to any of $\bar{V}_\Pi(T_j')$, they are contained in the union of $C(\bar{\sigma}_i^h)$, and consequently, do not belong to F_h. But, then, since infinitely many σ_i's belong to F_h, at least one of $\bar{V}_\Pi(T_j')$ contains infinitely many σ_i's. Take such a j and put $T_{k+1} = T_j'$. From the construction of T_{k+1}, it satisfies (3.1)–(3.3). Note also $| T_{k+1} | - | T_0 | = k + 1$. We have constructed, then, a sequence of trees $(T_r)_{r \geqslant 0}$ satisfying (3.1)–(3.4).

Let $\bar{\Pi}$ be a pruning set of K such that f^{-1} is continuous from $\mathbf{T}_{\Pi'}$ to $\mathbf{T}_{\bar{\Pi}}$. By Theorem 1 and from the assumption that the set of almost terminal phrases of K is finite, we may further assume that $\bar{\Pi}$ is an extension of Π containing all the almost terminal phrases of K. We consider $\bar{H}_r = \bar{V}_{\bar{\Pi}}(T_r)$. We also put $H_r = \bar{V}_\Pi(T_r)$.

We first observe that $\bar{H}_r$ also contains infinitely many σ_i's. For if σ is in H_r but not in $\bar{H}_r$, let p be a terminal node of $(\sigma)_{\bar{\Pi}}$ which is not a terminal node of $(\sigma)_\Pi$, and define $\bar{\sigma}$ as the union of T_r and the tree obtained from $(\sigma)_{\bar{\Pi}}$ by grafting at p the phrase of σ at p. Note that p is not a sterile node; otherwise, it would be terminal in $(\sigma)_\Pi$, as we are dealing with the theory relative to Ω^*. Hence, $\bar{\sigma}$ is a partial sentence. We have $T_r < \bar{\sigma} < \sigma$. The height of $\bar{\sigma}$ relative to T_r is equal to or less than the height of $\bar{\Pi}$ relative to Ω^*. Now, $\sigma \in \bar{V}_\Pi(T_r) \cap C(\bar{\sigma}) \subset \bar{V}_\Pi(T_0) \cap C(\bar{\sigma}) = V_\Pi(\tau) \cap C(\bar{\sigma})$; i.e., $V_\Pi(\tau)$ and $C(\bar{\sigma})$ meet condition (2.2). Furthermore, if r is sufficiently large, T_r, hence $\bar{\sigma}$, cannot be a subtree of τ (condition (2.1)). Then, if $h_{\bar{\sigma}}$ is h, σ is not in E_h, and not in F_h. Hence, all the σ_i's in H_r which are in F_h are in $\bar{H}_r$, and, hence, there are infinitely many σ_j's in $\bar{H}_r$, for a sufficiently large r, and hence, for any r.

For each r we choose one σ_i that belongs to $\bar{H}_r$ and call it ρ_r. Put $H_r' = f(\bar{H}_r)$ and $\rho_r' = f(\rho_r)$. Since f^{-1} is continuous from $\mathbf{T}_{\Pi'}$ to $\mathbf{T}_{\bar{\Pi}}$ and $\bar{H}_r = \bar{V}_{\bar{\Pi}}(T_r)$ is open in $\mathbf{T}_{\bar{\Pi}}$, H_r' is an open set in $\mathbf{T}_{\Pi'}$ containing ρ_r'.

Each σ_i, hence each ρ_r, is in D, and ρ_r' is in D', the closure of the partial sentence $\bar{\sigma}'$; that is, $\bar{\sigma}'$ is a rooted subtree of ρ_r'. Obviously, $\bar{\sigma}'$ cannot be a sentence, for there are infinitely many ρ_r''s as f is one-to-one. Thus, there is a fertile phrase π' of $\bar{\sigma}'$; call its root p. Since K' is assumed to be context-free, there are infinitely many partial sentences identical to $\bar{\sigma}'$ except for the phrases at p. Let $(\hat{\bar{\sigma}}_k')_{k\geqslant 1}$ be an infinite sequence of such partial sentences. Furthermore, we may assume that for each k the phrase $\hat{\pi}_k'$ of $\hat{\bar{\sigma}}_k'$ at p is sufficiently large so that it is not a subphrase of any of the phrases in Π'. It is of course not sterile. It follows that for any σ' in the closure of $\hat{\bar{\sigma}}_k'$, p is a node in $(\sigma')_{\Pi'}$.

Since ρ_r' is in D', p is a node of ρ_r'. For each pair (r, k), we construct $\hat{\rho}'_{r,k}$ as follows: Replace the phrase of ρ_r' at p by $\hat{\pi}_k'$. Since K' is context-free $\hat{\rho}'_{r,k}$ is a sentence of K'. Note that $\hat{\rho}'_{r,k}$ is in the closure $C(\hat{\bar{\sigma}}_k')$ of $\hat{\bar{\sigma}}_k'$. Furthermore, $\hat{\rho}'_{r,k}$ is in $V_{\Pi'}(\rho_r')$; hence, it is also in H_r', since H_r' is open in $\mathbf{T}_{\Pi'}$ and contains ρ_r', as has been remarked earlier.

Let $\bar{T}'$ be the tree obtained from $\bar{\sigma}'$ by deleting its phrase at p. Then, both $\bar{T}'$ and $(\hat{\rho}'_{r,k})_{\Pi'}$ are rooted subtrees of $\hat{\rho}'_{r,k}$. Hence, they have the intersection tree $\bar{T}'_{r,k}$, which is a rooted subtree of $\bar{\sigma}'$. Let T' be the intersection of all $\bar{T}'_{r,k}$. Note that p is a terminal node of each $\bar{T}'_{r,k}$, and hence, a terminal node of T'. Finally, let $\hat{T}_k'$ be a tree obtained from T' by grafting the full branch of $(\hat{\rho}'_{r,k})_{\Pi'}$ at p; that is, $\hat{T}_k'$ is identical to T' except for its full branch at p, which is the same as that of $(\hat{\rho}'_{r,k})_{\Pi'}$ at p. (From the condition imposed on $\hat{\bar{\sigma}}_k'$, this branch is the same for all r and identical to the full branch of $(\hat{\bar{\sigma}}_k')_{\Pi'}$ at p.) Then, $\hat{T}_k'$ is a rooted subtree of $(\hat{\rho}'_{r,k})_{\Pi'}$. Putting $\hat{H}_k' = \bar{V}_{\Pi'}(\hat{T}_k')$, we have $\hat{\rho}'_{r,k} \in \hat{H}_k'$. Combining this with the result of the preceding paragraph, we have $\hat{\rho}'_{r,k} \in H_r' \cap \hat{H}_k'$.

We are now going to map the entities defined above in K' back into K. We put $\hat{\rho}_{r,k} = f^{-1}(\hat{\rho}'_{r,k})$ and $\hat{H}_k = f^{-1}(\hat{H}_k')$. Recall that $\bar{H}_r = \bar{V}_{\bar{\Pi}}(T_r)$ and $H_r' = f(\bar{H}_r)$. We have $\hat{\rho}_{r,k} \in \bar{H}_r \cap \hat{H}_k$.

Recalling that $\hat{H}_k' = \bar{V}_{\Pi'}(\hat{T}_k')$ and the condition imposed on Π' at the beginning of the proof, we conclude from the lemma stated above that for some sentence $\hat{\sigma}_j'^{(k)}$, $1 \leqslant j \leqslant l(k)$, of K, we have $\hat{\sigma}_j'^{(k)} = f(\hat{\sigma}_j^{(k)}) \in \hat{H}_k'$ and $\hat{H}_k \subset \bigcup_{1\leqslant j\leqslant l(k)} V_{\bar{\Pi}}(\hat{\sigma}_j^{(k)})$. Since all $\hat{\rho}_{r,k}$ belong to $\hat{H}_k$, at least one of the factors of this union contains infinitely many $\hat{\rho}_{r,k}$. Let us take such $V_{\bar{\Pi}}(\hat{\sigma}_j^{(k)})$ and put $\hat{\sigma}_j^{(k)} = \hat{\sigma}_k$, $\hat{\sigma}_k' = f(\hat{\sigma}_k)$. Then, for infinitely many r, $V_{\bar{\Pi}}(\hat{\sigma}_k) \cap \bar{H}_r$ is not empty; that is, for infinitely many r, $(\hat{\sigma}_k)_{\bar{\Pi}}$ and T_r have their union.

Consider such infinitely many T_r's. Assume that for every r, $(\hat{\sigma}_k)_{\bar{\Pi}}$ is not a rooted subtree of T_r. Then there is a terminal node q of $(\hat{\sigma}_k)_{\overline{\Pi}}$ which is not a node of T_r in the union of $(\hat{\sigma}_k)_{\overline{\Pi}}$ and T_r, for any r. Then there must be a node $\bar{q}$ of $\hat{\sigma}_k$ properly dominating q such that $\bar{q}$ is terminal in T_r for any sufficiently large r. From the construction of T_r, such $\bar{q}$ must be labeled with an almost terminal symbol. But this is a contradiction, since $\overline{\Pi}$ contains all the almost terminal phrases and a node of $(\hat{\sigma}_k)_{\overline{\Pi}}$ labeled with an almost terminal symbol can only be terminal. It follows that there is r_k such that $(\hat{\sigma}_k)_{\overline{\Pi}}$ is a rooted subtree of T_{r_k}. We conclude $\rho_{r_k} \in V_{\overline{\Pi}}(\hat{\sigma}_k)$. (Recall that for each r, $\rho_r \in \overline{\overline{H}}_r$.)

Take now a pruning set $\overline{\Pi}'$ of K' such that f is continuous from $\mathbf{T}_{\overline{\Pi}}$ to $\mathbf{T}_{\overline{\Pi}'}$. From $\rho_{r_k} \in V_{\overline{\Pi}}(\hat{\sigma}_k)$ we have $\rho'_{r_k} \in V_{\overline{\Pi}'}(\hat{\sigma}_k')$, i.e., $(\hat{\sigma}_k')_{\overline{\Pi}'} < (\rho'_{r_k})_{\overline{\Pi}'}$. Recall on the other hand $\hat{\sigma}_k' \in \hat{H}_k' = \overline{V}_{\overline{\Pi}'}(\hat{T}_k')$. Hence, $\hat{T}_k' < (\hat{\sigma}_k')_{\Pi'} < \hat{\sigma}_k'$.

Since ρ'_{r_k} is in D', the node p of $\bar{\sigma}'$ which we referred to in the construction of $\hat{T}_k'$ is a node of ρ'_{r_k}. On the other hand, p is also a node of $\hat{T}_r'$, hence, from the preceding paragraph, also a node of $\hat{\sigma}_k'$. But the height of the full branch of $\hat{T}_k'$ at p, and, hence, also that of the phrase of $\hat{\sigma}_k'$ at p, is not bounded. Consequently, if we choose k so large that the height of the phrase of $\hat{\sigma}_k'$ at p is greater than the height of the phrase of $\bar{\sigma}'$ at p plus the height of $\overline{\Pi}'$, then $(\hat{\sigma}_k')_{\overline{\Pi}'}$ cannot be a rooted subtree of $(\rho'_{r_k})_{\overline{\Pi}'}$, which contradicts the relation $(\hat{\sigma}_k')_{\overline{\Pi}'} < (\rho'_{r_k})_{\overline{\Pi}'}$ established also in the preceding paragraph.

This completes the proof of the theorem.

Remark. The assertions of Theorem 6 and its corollary do not hold if "structurally *-continuous" and "structural *-homeomorphism" are replaced by "structurally (*)-continuous" and "structural (*)-homeomorphism". Consider, for example, two context-free grammars G and G' defined by the following rules.

$$\begin{aligned} G:\quad & S \to aS \\ & S \to a \\ G':\quad & S \to AS \\ & S \to A \\ & A \to a. \end{aligned}$$

Put $K = K(G)$, $K' = K(G')$. Denote by τ_n and τ_n' the sentences of K and K' such that $\eta(\tau_n) = \eta(\tau_n') = a^n$, where η is the yield function. Define f by $f(\tau_n) = \tau_n'$. Then f is obviously one-to-one and onto. To see that f is a structural homeomorphism, let Π be an arbitrary pruning set of K. Note that phrases of K are sentences except for a itself. $\overline{\Pi}' = \{\pi'; \pi' = f(\pi), \pi \in \Pi\}$ is a pruning set of K'. Then, f is a homeomorphism from $\mathbf{T}_{\Pi}$ onto $\mathbf{T}_{\overline{\Pi}'}$. Next let Π' be an arbitrary pruning set of K'. Note that phrases of K' are sentences except for $A(a)$ and a. Let Π_1' be the subset of Π' consisting of all the sentences in Π'. Then the identity map of K' is homeomorphic from $\mathbf{T}_{\Pi'}$ onto $\mathbf{T}_{\Pi_1'}$. Put $\Pi_1 = \{\pi; f(\pi) \in \Pi_1'\}$. From what was said above f is a homeomorphism from $\mathbf{T}_{\Pi_1}$ onto $\mathbf{T}_{\Pi_1'}$. It follows that f^{-1} is a homeomorphism from $\mathbf{T}_{\Pi'}$ onto $\mathbf{T}_{\Pi_1}$. Hence, K and K' are structurally homeomorphic.

Note that a prepartial sentence of K is a sentence. It follows that the topologies of partial sentences and of prepartial sentence of K are both identical to the topology of cofinite sets (i.e., the topology whose open sets are those whose complement is finite or K). Let now Π_0 be the singleton set $\{S(a)\}$. Then for each sentence τ_n of K, $V_{\Pi_0}(\tau_n) = \{\tau_m; m \geqslant n\}$, it follows that no nonempty open set in $\mathbf{T}^{(*)}_{\Pi_0}$ is finite and a fortiori $\{\tau_n\}$ is not open in $\mathbf{T}^{(*)}_{\Pi_0}$.

On the other hand, consider now an arbitrary sentence τ_n' of K', and let $\bar{\tau}_n'$ be the tree obtained from τ_n' by pruning the phrase $S(A(a))$. Unless $n = 1$, $\bar{\tau}_n'$ is a prepartial sentence, since it contains the phrase $A(a)$ as a branch. We have $C(\bar{\tau}_n') = \{\tau_m'; m \geqslant n\}$. It follows that $\{\tau_k'; 1 \leqslant k \leqslant n\}$ is open for $n > 1$ in $\mathbf{T}^{(*)}(K')$. Let Π' be an arbitrary pruning set of K' that contains at least one phrase other than $A(a)$, and take n, $n > 1$, so large that for no m, $m \geqslant n$, is τ_m' in Π'. Then we have $V_{\Pi'}(\tau_n') = \{\tau_m'; m \geqslant n\}$. Hence, $\{\tau_n'\}$, which is the intersection of $\{\tau_m'; m \geqslant n\}$ and $\{\tau_k'; 1 \leqslant k \leqslant n\}$ is open in $\mathbf{T}^{(*)}_{\Pi'}$. On the other hand, if Π' is empty or $\Pi' = \{A(a)\}$, $\mathbf{T}_{\Pi'}$ is discrete and $\{\tau_n'\}$ is trivially open in $\mathbf{T}^{(*)}_{\Pi'}$.

From the preceding two paragraphs it follows that for an arbitrary pruning set Π' of K', there exists n such that f is not continuous from $\mathbf{T}^{(*)}_{\Pi_0}$ to $\mathbf{T}^{(*)}_{\Pi'}$ at τ_n; f is not structurally (*)-continuous.

II. Regular Languages

1. *Introduction*

We are now going to deal with a restricted class of languages, languages that can be related to regular sets in a certain way. As is well known, regular sets are string languages that are characterized as sets acceptable by finite automata, representable by regular expressions, or as sets generatable by one-sided linear grammars. But we are interested in tree languages. The last characterization relates regular sets to a class of tree languages, in fact, to a subclass of the class of context-free tree languages. This subclass seems to represent naturally the counterpart, in the present framework, of regular sets. One-sided linear grammars, however, seem to impose a certain unnecessary limitation on us for a reason that will become clear directly. We shall introduce instead what we shall call regular grammars. This represents a generalization, in a certain sense, of one-sided linear grammars.

Let us first indicate what consequence the direct application of our general theory to one-sided linear grammars would have. Let us consider a right linear grammar G, and for the sake of simplicity let us assume each rule of G is of the form $A \rightarrow aB$ or $A \rightarrow a$, where a is a single terminal letter. Then, as is well known, G can be directly converted into a finite automaton $A(G)$ by interpreting the rules of G, $A \rightarrow aB$, as the transition function of $A(G)$. A sentence of G may then be identified with a labeled path from the initial to the final state of $A(G)$ where each unit path from a state to another is labeled with a transition symbol. It is easy to see that, given a pruning set Π of G, σ belongs to $V_{\Pi}(\tau)$ if and only if the path corresponding to τ is identical, including labels attached to unit paths (i.e., transition symbols), to an initial section of that corresponding to σ, except for certain last states of σ and τ specified by Π. But it might also be interesting to compare the paths corresponding to σ and τ disregarding what symbols are emitted at each transition. The regular grammar defined below will accommodate both of these situations.

A *right regular grammar* is defined as a context-free grammar G that satisfies the following conditions.

R1. The nonterminal vocabulary consists of two disjoint subvocabularies, that of those symbols called *preterminals* and that of those symbols called *nonpreterminals*.

R2. The initial (i.e., sentence) symbol of G is a nonpreterminal.

R3. Rules of G are one of the following forms:

(a) $P \rightarrow AQ$
(b) $P \rightarrow Q$
(c) $P \rightarrow A$
(d) $A \rightarrow x$

where P and Q are nonpreterminals, A a preterminal, and x a nonempty string over the terminal vocabulary.

If R3(a) is replaced by

(a′) $P \rightarrow QA$

G is by definition *left regular*. In what follows we shall only deal with right regular grammars, and we shall use "*regular grammar*" in place of "right regular grammar."

A regular grammar is called *standard* if it meets the following conditions.

S1. G has no rule of the form R3(b).

S2. For any nonpreterminal P there exists at most one preterminal A such that $P \rightarrow A$.

S3. For any pair of nonpreterminals (P, Q) there exists at most one preterminal A such that $P \rightarrow AQ$.

S4. No preterminal A appears in the right-hand side of more than one rule.

A preterminal appearing in a rule $P \rightarrow A$ is called an *end preterminal.*

A (*standard*) *regular language* is by definition a language generated by a (standard) regular grammar.

A tree belonging to a regular grammar G (or, in brief, a tree of G) is called *preterminal* (*nonpreterminal*) if its root is labeled with a preterminal (nonpreterminal). A tree of G is a *phrase of* G if its terminal nodes are labeled with terminal symbols. A phrase of G is *preterminal* (*nonpreterminal*) if it is preterminal (nonpreterminal) as a tree. A tree is called *terminated* if it contains a nonpreterminal phrase, or equivalently, if it contains a phrase of the form $P(A(x))$. (A terminated tree may not itself be a phrase, since some of its preterminal nodes may be terminal and not dominate terminal symbols).

Obviously the string language covered by a regular language is a regular set which does not contain the empty word. Conversely, any such regular set can be obtained as the string language covered by some regular language, in fact, by some standard regular language.

Standard regular languages are those which, from our topological point of view, disregard the difference between transition symbols. On the other hand, if we impose on regular languages the restriction that for each terminal string x, there exists at most one pair of nonpreterminal P and preterminal A such that $P \rightarrow AQ$ for some Q or $P \rightarrow A$ and $A \rightarrow x$, then we have a subclass of regular grammars that are, so to speak, sensitive to different transition symbols and, in this sense, equivalent to right linear grammars; that is, they define the same topological structures that right linear grammars would if we were to apply our general theory directly to the latter. By imposing still other conditions on the relations between the preterminals of a regular grammar and the terminal strings they generate, one can assign still different roles to transition symbols.

In what follows, however, we restrict ourselves to standard regular grammars. But, in a sense, this does not introduce any essential restriction. For if a regular grammar contains, say, two rules $P \rightarrow AQ$ and $P \rightarrow A'Q$, with the same pair (P, Q) of nonpreterminals but two different preterminals A and A', then one could replace them by two rules $P \rightarrow AQ$ and $P \rightarrow A'Q$, with different pairs (P, Q) and (P, Q') of nonpreterminals. Of course, this is not an exact account and I am not going into details, but it is perhaps enough of a hint to justify the present restriction of our discussion to standard regular grammars.

Remark. The last of the conditions in the definition of a standard regular grammar, namely that no preterminal may appear in the right-hand side of more than one rule ($P \rightarrow AQ$ or $P \rightarrow A$), is, in a sense, immaterial. If a regular grammar G satisfies the other conditions for a standard regular grammar, we can easily convert it into a standard regular grammar by introducing enough new preterminals to differentiate different occurrences of preterminals in rules of G. The thus standardized grammar is structurally equivalent to the original one, trivially. The reason to impose the condition in question on standard regular grammars is to facilitate technically the process of construction of standard regular grammars on the basis of others in our following discussion.

2. *Structural Topologies and Finite State Diagrams*

We shall associate a finite state diagram with each standard regular grammar; we shall describe topologies defined on standard regular languages in terms of the associated finite state diagrams. A *finite state diagram* D, as we understand it here, is defined on a pair of two finite sets, the set of states Γ and the transition alphabet Δ, and satisfies the following conditions.

F1. With each ordered pair (P, Q) of states is associated a (possibly void) finite subset of Δ^*, namely the set of *transition words* for (P, Q), which we denote by $\delta(P, Q)$.

F2. One state is designated as the *initial* state and one state is designated as the *final* state.

If $\delta(P, Q)$ is not void, we call the ordered pair (P, Q) a *unit path* of D. A *path* θ of D is a finite sequence $(P_0, P_1, \ldots, P_n)$, $n \geqslant 0$, of states such that each (P_i, P_{i+1}), $0 \leqslant i \leqslant n$, is a unit path; n is called the *length* of θ. P_0 and P_n are called the *origin* and the *endpoint* of θ, respectively; θ may also be said to be a path *from* P_0 to P_n. Paths of lenght 0 may be identified with states and paths of lenght 1 with unit paths. Let $\theta_1 = (P_0, P_1, \ldots, P_n)$ and $\theta_2 = (Q_0, Q_1, \ldots, Q_m)$ be two paths. If $P_n = Q_0$, their *product* $\theta_1 \cdot \theta_2$ is defined as the path $(P_0, P_1, \ldots, P_n, Q_1, \ldots, Q_m)$. On the other hand, if $n \leqslant m$ and for each i, $1 \leqslant i \leqslant n$, $P_i = Q_i$, then the *left quotient* $\theta_1 \backslash \theta_2$ is defined as the path $(Q_{m-n}, Q_{m-n+1}, \ldots, Q_m)$. Finally, if $m \leqslant n$ and for each i, $1 \leqslant i \leqslant m$,

$Q_i = P_{n-m+i}$, the *right quotient* θ_1/θ_2 is defined as the path $(P_0, P_1, \ldots, P_{n-m})$. We will use set theoretical notions and symbols with respect to paths considered as ordered sets of states; for example, $\theta' \subset \theta$ if and only if $\theta = \theta_1 \cdot \theta' \cdot \theta_2$ for some paths θ_1 and θ_2.

Let G be a standard regular grammar and let us construct a finite state diagram $D(G)$ to be associated with G in the following way. The set of states of $D(G)$ is the set of nonpreterminal symbols of G plus the final state F. The sentence symbol of G is the initial state of $D(G)$. For each ordered pair (P, Q) of nonpreterminals of G, let $\delta(P, Q)$ be the set of terminal strings x such that for some preterminal A, $P \to AQ$ and $A \to x$ are rules of G; schematically $\delta(P, Q) = \{x; P \to AQ, A \to x\}$. (Hence, $\delta(P, Q)$ is void if there is no rule $P \to AQ$.) Similarly, for each nonpreterminal P, let $\delta(P, F) = \{x; P \to A, A \to x\}$. (Hence, $\delta(P, F)$ is void if there is no rule $P \to A$.)

We now assign a path $\theta(T)$ in $D(G)$ to each nonpreterminal tree T belonging to G inductively as follows: If T is of height 0, it is a nonpreterminal symbol, say, P, which is a path of length 0; we put $\theta(T) = P$. In general, if T is of height $n > 0$, either $T = P(T_0)$ or $T = P(T_0, T_1)$, where P is a nonpreterminal, T_0 is a preterminal tree, and T_1 is a nonpreterminal tree whose height is less than n; if $T = P(T_0)$, we put $\theta(T) = (P, F)$; if $T = P(T_0, T_1)$, we put $\theta(T) = (P, Q) \cdot \theta(T_1)$, where Q is the label of the root of T_1.

Thus, we have a function $\theta(T)$ from the set of nonpreterminal trees belonging to G into the set of paths of $D(G)$. The image of θ is the set of those paths of $D(G)$ whose origin is not F.

Let T_1 and T_2 be two nonpreterminal trees of G. $\theta(T_1) = \theta(T_2)$ if and only if T_1 and T_2 are identical except for leaves (i.e., the nodes of either of them that do not belong to the other are labeled with terminal symbols). We prove this by induction on the length of $\theta(T_1)$. If the length of $\theta(T_1)$ is zero, put $\theta(T_1) = P$; then $\theta(T_1) = \theta(T_2)$ if and only if $T_2 = P$, i.e., $T_1 = T_2$. If the length of $\theta(T_1)$ is $n > 0$, put $\theta = \theta(T_1) = (P_0, P_1, \ldots, P_n)$, $\theta_1 = (P_0, P_1)$, and $\theta_2 = (P_1, \ldots, P_n)$. If $P_1 = F$, then $\theta = \theta_1$, and $T_1 = P_0(T_0^1)$, where T_0^1 is a preterminal tree whose root A is such that $P_0 \to A$ is a rule of G. For this case, $\theta(T_2) = \theta$ if and only if $T_2 = P_0(T_0^2)$, where T_0^2 is a tree whose root is A; hence, $\theta(T_2) = \theta$ if and only if T_1 and T_2 are identical except for leaves. Finally, if $P_1 \neq F$, then $T_1 = P_0(T_0^1, T_1^1)$, where T_1^1 is a tree whose root is P_1; from the definition of θ, $\theta(T_1^1) = \theta_2$. $\theta(T_2) = \theta$ if and only if $T_2 = P_0(T_0^2, T_1^2)$ and $\theta(T_1^2) = \theta_2$. By an induction hypothesis, we conclude that $\theta(T_1^2) = \theta_2$ if and only if T_1^1 and T_1^2 are identical except for leaves. Consequently, (since T_0^1 and T_0^2 are identical except for leaves) $\theta(T_2) = \theta$ if and only if T_1 and T_2 are identical except for leaves.

It follows that if in particular T_1 and T_2 are leafless nonpreterminal trees (i.e., no node is labeled with a terminal symbol), then $\theta(T_1) = \theta(T_2)$ if and only if $T_1 = T_2$. Hence, so long as no confusion is likely, we may agree to identify paths of $D(G)$ whose origin is not the final state with leafless nonpreterminal trees of G.

We shall apply our general theory to standard regular grammars. We shall consider structural topologies relative to the set of all preterminal phrases Ω_0. From Part I, Theorem 5, Ω_0 is equivalent to the void set of phrases; that is, so far as structural homeomorphic invariants are concerned, the theory relative to Ω_0 is equivalent to the absolute theory. It is also equivalent to the theory relative to Ω^*, the set of all sterile phrases. Note also that each preterminal phrase is sterile; hence, Ω_0 is contained in Ω^*.

Relativization with respect to Ω_0 allows us to conveniently discuss structural topologies of a standard regular grammar G in terms of the finite state diagram $D(G)$ associated with it. Indeed, for any pruning set Π and for any sentence τ of K, $(\tau)_\Pi$ is leafless, since all the branches of τ of the form $A(x)$ are pruned by Ω_0. Accordingly, as argued above, we may regard $(\tau)_\Pi$ as a path in $D(G)$ by identifying it with $\theta((\tau)_\Pi)$. (We may also write $\theta(\tau)_\Pi$ instead of $\theta((\tau)_\Pi)$ for the sake of brevity.) Furthermore, if σ is another sentence of G, we have $(\tau)_\Pi < (\sigma)_\Pi$ if and only if $\theta(\tau)_\Pi \subset \theta(\sigma)_\Pi$. More generally, consider a tree T of G whose root is labeled with the initial symbol S of G. If T is not leafless, $\overline{V}_\Pi(T)$ is empty. (Recall that $\overline{V}_\Pi(T)$ is defined relative to Ω_0. Hence, for each σ, $(\sigma)_\Pi$ is leafless. If T is not leafless, $T < (\sigma)_\Pi$ is impossible.) On the other hand, assume T is leafless; then $T < (\sigma)_\Pi$ if and only if $\theta(T) \subset \theta(\sigma)_\Pi$. The relation $<_\Pi$ among leafless trees whose roots are labeled with S may be identified with the relation $\subset$ among paths whose origin is S. In particular, if for any path θ whose origin is S one defines $\overline{V}_\Pi(\theta) = \{\sigma; \theta \subset \theta(\sigma)_\Pi\}$, the class of sets $\overline{V}_\Pi(\theta)$, where θ ranges over such paths coincides with the class of sets $\overline{V}_\Pi(T)$, T ranging over the leafless trees with root S. In other words topology $\mathbf{T}_\Pi$ can be described in terms of paths in $D(G)$ instead of trees of G.

As far as our discussion here is concerned, then, a standard regular grammar G and a finite state diagram $D(G)$ may be identified in many respects. Accordingly, we may talk about notions which are originally defined with reference to G as if they were notions belonging to $D(G)$, and vice versa; e.g., we may speak of a path or state of G and a sentence or phrase of $D(G)$, etc.

In particular, if π_1 and π_2 are nonpreterminal phrase, and if π_2 is a subphrase of π_1, then one may say that π_2 is a subpath of π_1; in fact, the path $\theta(\pi_2)$ is a subpath of the path $\theta(\pi_1)$. But it is not the case when $\theta(\pi_2)$ is a subpath of $\theta(\pi_1)$ that π_2 is necessarily a subphrase of π_1; we can only say that π_2 is identical to some subphrase of π_1 except for leaves.

Given an arbitrary τ in K, there either does or not exist in $\Pi \cup \Omega_0$ a nonpreterminal phrase which is a subphrase of τ. In the former case, there exists among such phrases the largest one (i.e., such that the others are subphrases of it); denote this phrase by $\pi(\tau, \Pi)$. Then we have $\theta(\tau)_\Pi = \theta(\tau)/\theta(\pi(\tau, \Pi))$. In conformity with the convention of identifying the tree $(\tau)_\Pi$ with the path $\theta(\tau)_\Pi$, we may also write $(\tau)_\Pi = \tau/\pi(\tau, \Pi)$. (On the right-hand side of the equation, the operation / is not, properly speaking, defined between trees τ and $\pi(\tau, \Pi)$; but $\tau/\pi(\tau, \Pi)$ as a whole has a definite meaning as defined above.)

In the latter case (i.e., where no nonpreterminal phrase is pruned from τ by $\Pi \cup \Omega_0$) we have $\theta(\tau)_\Pi = \theta(\tau)$. To facilitate our exposition, we still agree to write $\theta(\tau)_\Pi = \theta(\tau)/\theta(\pi(\tau, \Pi))$ or $(\tau)_\Pi = \tau/\pi(\tau, \Pi)$, by interpreting $\theta(\pi(\tau, \Pi))$ as the path (F) of length 0, although $\pi(\tau, \Pi)$ itself is not given any definite meaning as a phrase. We call $\pi(\tau, \Pi)$ with this meaning the singular phrase of G. In any case, singular $\pi(\tau, \Pi)$ may be considered to represent a path with its endpoint F (as nonsingular $\pi(\tau, \Pi)$'s may), as long as it appears in the expression $\tau/\pi(\tau, \Pi)$. More generally, we shall agree to say that the singular phrase π_0 is such that for any phrase π (singular or not), $\pi/\pi_0 = \theta(\pi)/\theta(\pi_0) = \theta(\pi)$, and to say that π_2 is a subpath of π_1 if and only if $\pi_1/\pi_2 \subset \theta(\pi_1)$, whether π_1 and π_2 are singular or not. Hence, the singular phrase is a subpath of any phrase, and no phrase except for the singular phrase is a subpath of the singular phrase. (One might interpret the singular phrase as an entity which, so to speak, appears in each nonpreterminal phrase π of G in the shape of the end preterminal phrase of π. Since all such preterminal phrases are pruned by Ω_0, this ambiguous interpretation of the singular phrase does not prevent it from being considered as a "subpath" of any nonpreterminal phrase π. We shall not say, however, that the singular phrase is a "subphrase" of π.)

Let us now consider a prepartial sentence $\bar{\sigma}$ of G. We have two cases. First, assume that there exists a maximal phrase π of $\bar{\sigma}$ which is a nonpreterminal phrase. Then a sentence σ of G which is in the closure $C(\bar{\sigma})$ of $\bar{\sigma}$ is identical to $\bar{\sigma}$ except for leaves; this is because no terminal node of $\bar{\sigma}$ is nonpreterminal. If $\bar{\sigma}$ is not a sentence, it is a partial sentence just in case π is not an almost terminal phrase. Second, assume that no maximal phrase of $\bar{\sigma}$ is a nonpreterminal phrase. Then $\bar{\sigma}$ is not a partial sentence, since no preterminal symbol can "generate" infinitely many phrases.

It follows that for any partial sentence $\bar{\sigma}$, $C(\bar{\sigma})$ is a finite set of sentences of G. Conversely, every finite set of sentences of G is a closed set in $\mathbf{T}^*$, the topology of partial sentences on $K = K(G)$, because each sentence of G is a partial sentence and constitutes by itself its closure. Hence, we have

THEOREM 1. *For a standard regular language K, the topology of partial sentences $\mathbf{T}^*(K)$ is the topology of cofinite sets (i.e., the topology whose open sets are the empty set and those sets whose complements are finite).*

3. *Elimination of a Not Directly Self-Connected Inner State*

We shall now try to normalize standard regular languages in various ways by means of structural homeomorphisms.

As the first step toward this end we shall prove in this section a theorem of a technical nature. The meaning of the theorem is that given a standard regular grammar C, one can construct a standard regular grammar structurally homeomorphic to G by eliminating a specified state of G, possibly at the expense of introducing some new states. Thus, in this general form the theorem does not contribute to simplifying grammars. But it will later serve to eliminate states with certain specified properties from a standard regular grammar.

Before stating the theorem we shall introduce some notions. A state of a standard regular grammar G is called *inner* if it is neither the initial nor the final state of G. A state P is said to be *directly connected forward* (or, simply, directly connected) to another state Q, if (P, Q) is a unit path; P is said to be *connected* (*forward*) to Q if P is related to Q by the transitive closure of the relation "directly connected to." A state P is called *directly self-connected* if P is directly connected to itself, and *self-connected* if it is connected to itself.

In the theorem to be stated directly we shall be concerned with a specified inner state O not directly self-connected, with those state which are directly connected to O, and with those states to which O is directly connected. Now, O may or may not be directly connected to the final state. To deal with both possibilities similarly as far as possible, we shall introduce the following convention. We denote by Q_j, $1 \leqslant j \leqslant l$, those states of G which are not the final state and to which O is directly connected. If O is directly connected to the final state, we shall agree to let Q_0 refer to the final state. We shall use [ ] to enclose statements, or portions thereof, which are relevant only in the case where O is directly connected to the final state. For example, we say that $[Q_0]$, $Q_1, \ldots, Q_l$, or Q_j, $0 \leqslant j \leqslant l$ $[0 < j \leqslant l]$, are all the states which O is directly connected to, meaning that all the states which O is directly connected to are Q_j, $0 \leqslant j \leqslant l$, if O is directly connected to the final state, and Q_j, $0 < j \leqslant l$, otherwise.

On the other hand, we denote the states which are directly connected to O by $P_1, P_2, \ldots, P_k$. Note that some of the Q_j's, $1 \leqslant j \leqslant l$, may be identical to some P_i, $1 \leqslant i \leqslant k$, as there may exist a state which is directly connected to O and to which O is directly connected.

We shall introduce another convention before stating the following theorem and its proof. We can eliminate the state O, as has been mentioned above, at the expense of duplicating the states Q_j, $1 \leqslant j \leqslant l$, i.e., by introducing "primed copies" Q_j' of Q_j. However, in the case in which no P_i is directly connected to any Q_j, we can obtain a sharper result, namely, that we can eliminate O without introducing any new states. Intuitively, this situation may be described as follows: If no P_i is directly connected to any Q_j, then we might introduce new unit paths (P_i, Q_j) to replace in the new grammar the paths (P_i, O, Q_j) of length 2 in the original grammar without interfering with the rest of the structure of the latter. On the other hand, if (P_i, Q_j) is already a unit path in the original grammar, this procedure does not work and we need new "primed states" Q_j' and new unit paths (P_i, Q_j') to replace (P_i, O, Q_j). Now, in order to prove the general and the special case simultaneously, we shall introduce the following convention. Statements, or portions thereof, may be enclosed in ⟨ ⟩ when they are relevant only in the special case in question. We may also enclose in ⟨ ⟩ some remarks, whenever it seems advisable to do so, as to how some statements or paragraphs are to be interpreted for the special case.

⟨Thus, in the special case in question, a primed letter Q_j' is understood as having the same referent as Q_j. The new grammar has the same states as the original one except for O, which is eliminated. A primed letter may simply be interpreted as a mnemonic sign for an occurrence of Q_j which P_i directly dominates in some tree, or, to put it differently, to which P_i is directly connected in some path. "A phrase or tree whose root is Q_j'" may be understood to mean "a phrase or tree whose root is Q_j considered as a subphrase or subtree of another phrase in which the former is directly dominated by some P_i."⟩

As the last preliminary before stating the theorem, let us classify the rules of G as follows. We shall use letters P and Q as generic terms for nonpreterminal symbols, and A and B for preterminal symbols. On the other hand, P, Q (or A, B) with a subscript i, j, etc., is a proper name for a certain nonpreterminal (or preterminal) symbol of G.

(a) $A \to x$.

(b) $P \to A$, where $P \neq O$.

(c) $P \to AQ$, where $P \neq O$ and $Q \neq O$.

(d1) $P_i \to A_i O$, $1 \leqslant i \leqslant k$.

(d2) $O \to B_j Q_j$, $1 \leqslant j \leqslant l$ (then, $Q_j \neq O$, since O is not directly selfconnected).

[(e) $O \rightarrow B_0$. (This means that O is directly connected to the final state.)]

THEOREM 2. *Let O be an inter state of a standard regular grammar G which is not directly self-connected. Let P_i, $1 \leqslant i \leqslant k$, be the states that are directly connected to O and let Q_j, $0 \leqslant j \leqslant l$ $[0 < j \leqslant l]$, be the states to which O is directly connected. ⟨When statements inside the angular parentheses are read, following the convention introduced above, assume that no P_i is directly connected to any Q_j.⟩ Construct a new standard regular grammar G' as follows. The states of G' are the states of G, except for O, plus the "primed copies" Q_j' for each j, $0 < j \leqslant l$, which are different from any states of G. ⟨The primed copies Q_j' are simply another name for Q_j.⟩ The initial and the final state of G are the initial and the final state of G', respectively. The preterminals of G' are the preterminals of G [except for B_0 such that $O \rightarrow B_0$] plus A_{ij}, $1 \leqslant i \leqslant k$, $0 \leqslant j \leqslant l$, where A_{ij} are new symbols for $1 \leqslant i \leqslant k$, $1 \leqslant j \leqslant l$ [and where A_{i0}, $1 \leqslant i \leqslant k$, is the preterminal of G such that $P_i \rightarrow A_{i0}$, if such a preterminal already exists in G; otherwise A_{i0} is also a new symbol]. The terminal symbols of G' are the terminal symbols of G [plus a new symbol u. ⟨If, however, no P_i is directly connected to Q_j, u is interpreted simply as a mnemonic whose referent may be considered to be the empty (i.e., identity) string in the free semi-group over the terminal alphabet; u is then to indicate that the terminal string to which it is attached is directly dominated by some P_i, $1 \leqslant i \leqslant k$.⟩] The rules of G' are given as follows:*

(a′–1) $A \rightarrow x$, *for each rule* (a) *of G [except for B_0 such that $O \rightarrow B_0$].*

(a′–2) $A_{ij} \rightarrow xy$, *for each pair of rules $A_i \rightarrow x$, $B, \rightarrow y$ of G, $1 \leqslant i \leqslant k$, $1 \leqslant j \leqslant l$.*

[(a′–3) $A_{i0} \rightarrow xyu$, *for each pair of rules $A_i \rightarrow x$, $1 \leqslant i \leqslant k$, and $B_0 \rightarrow y$ of G.*]

(b′) $P \rightarrow A$, *for each rule* (b) *of G.*

(c′) $P \rightarrow AQ$, *for each rule* (c) *of G.*

(d′) $P_i \rightarrow A_{ij}Q_j'$ *for each pair of rules* (d1) *and* (d2) *of G, $1 \leqslant i \leqslant k$, $1 \leqslant j \leqslant l$.*

[(e′) $P_i \rightarrow A_{i0}$, $1 \leqslant i \leqslant k$. *(This rule may already be a rule in G; then it is enumerated twice, once here and once in* (b′))].

(f′–b) $Q_j' \rightarrow B$, *for each rule $Q_j \rightarrow B$ of G, $1 \leqslant j \leqslant l$.*

(f′–c) $Q_j' \rightarrow AQ$, *for each rule $Q_j \rightarrow AQ$ of G, where $Q \neq O$ and $1 \leqslant j \leqslant l$.*

(f′–d) $Q'_{j_0} \rightarrow A_{i_0 j} Q_j'$, $1 \leqslant j \leqslant l$, *if* $Q_{j_0} = P_{i_0}$ *for some* i_0.

[(f′–e) $Q'_{j_0} \rightarrow A_{i_0 0}$, *if* $Q_{j_0} = P_{i_0}$ *for some* i_0 *and* $P_{i_0} \rightarrow A_{i_0}$ *is a rule of* G.]
Then G' *constructed as above is structurally homeomorphic to* G.

⟨*Remark* 1. In (d′), the primed sign is a mnemonic for the fact that the Q_j generated by the rule is directly dominated by P_i; the same for (f′–d), since the left-hand side Q'_{j_0} is a name for P_{i_0}. The other rules containing Q_j' contain it on the left-hand side; they may be considered simply as redundant repetitions of the rules with Q_j. According to interpretation given to u, $(a' - 3)$ is the same as $A_{i0} \rightarrow xy$.⟩

For the proof of the theorem we are going to construct a function f from $K = K(G)$ to $K' = K(G')$ and show that it is a structural homeomorphism. Some preliminary steps are required. We shall first define the degrees $\delta(\pi)$ and $\delta'(\pi')$ of nonpreterminal phrases π and π' of G and G', respectively. For a phrase π of G, if π is not a phrase whose root is labeled with O (in brief, an O-phrase), $\delta(\pi)$ is the number of occurrences of nonpreterminals in π (i.e., the number of occurrences of nonfinal states in the path $\theta(\pi)$); if π is an O-phrase, $\delta(\pi)$ is not defined. For a phrase π' of G', $\delta'(\pi')$ is the number of occurrences of nonpreterminals plus the number of occurrences of primed Q_j', $1 \leqslant j \leqslant l$ (i.e., each Q_j' contained in π' is counted twice ⟨i.e., each Q_j directly dominated by some P_i (mnemonic Q_j') is counted twice⟩) [plus 1, if π' contains an occurrence of u].

Remark 2. (1) (a′) rules, and no others, generate preterminal phrases of G'.

(2) (b′) rules, and no others, generate phrases of degree 1 of G' [except that if $P_i \rightarrow A_i$, for some i, already exists in G, such a rule is at the same time of type (e′), and may generate a phrase of degree 2].

(3) (c′) and (d′) rules, and no others, generate phrases of G' of degree more than one whose roots are not primed Q_j''s.

[(4) (e′) rules, followed by single (a′–3) rules, generate phrases of degree 2.]

(5) (f′) rules, and no others, generate phrases of G' whose roots are primed Q_j''s.

Let $P_n(G)$ and $P_n(G')$ be the sets of phrases of G and G' of degree n, respectively; let $P_n'(G')$ be the subset of $P_n(G')$ consisting of those phrases whose root is not a primed Q_j', $1 \leqslant j \leqslant l$. Put

$$P(G) = \bigcup_{n \geqslant 1} P_n(G), \qquad P(G') = \bigcup_{n \geqslant 1} P_n(G')$$

and

$$P'(G') = \bigcup_{n \geqslant 1} P_n'(G').$$

Next, for each phrase π' of G' whose root is Q_j, for some j, $1 \leqslant j \leqslant l$, we define $\varphi(\pi')$ as the tree obtained from π' by replacing its root Q_j by Q_j'. $\langle\varphi$ is the identity map; $\varphi(\pi')$ refers to the same phrase as π'. It may be considered as a mnemonic sign for π' to be used when π' is regarded as a subtree directly dominated by some P_i in another phrase.$\rangle$ Then, because of rule (f′), $\varphi(\pi')$ is a phrase of G', and φ gives a one-to-one correspondence between the Q_j-phrases of G' and Q_j'-phrases of G'. Note that $\delta'(\varphi(\pi')) = \delta'(\pi') + 1$. $\langle$Hence, the degree of π' is considered larger by 1 when it is regarded as a subphrase directly dominated by some P_i in another phrase.$\rangle$

We can inductively characterize the phrases of G' in the following way.

Lemma 1. *$\pi' = A'(z')$ is a preterminal phrase of G' if and only if one of the following conditions* (a1)–(a3) *holds:*

(a1) *π' is a preterminal phrase of G [except for the case in which $O \rightarrow A'$ is a rule of G];*

(a2) *$A' = A_{ij}$ for some i and j, $1 \leqslant i \leqslant k$, $1 \leqslant j \leqslant l$, and $z' = xy$ such that $A_i \rightarrow x$ and $B_j \rightarrow y$ are rules of G;*

[(a3) *$A' = A_{i0}$, $z' = xyu$ such that $A_i \rightarrow x$ and $B_0 \rightarrow y$ are rules of G*].

π' is a nonpreterminal phrase of G' if and only if one of the following conditions (b)–(f) *holds:*

(b) *$\pi' = P(\pi_0')$, where π' is a phrase of G (hence, $P \neq O$, $P \neq Q_j'$ and π_0' is a preterminal phrase of G);*

(c) *$\pi' = P(\pi_0', \pi_1')$, where $P \neq Q_j'$ and there exists a phrase π_1 of G of the same degree and root as π_1' such that $\pi = P(\pi_0', \pi_1)$ is a phrase of G;*

(d) *$\pi' = P_i(\pi_0', \varphi(\pi_1'))$, where π_1' is a phrase of G' whose root is some Q_j; π_0' is a preterminal phrase of type* (a2);

[(e) *$\pi' = P_i(\pi_0')$, where π_0' is a preterminal phrase of G' of the form $\pi_0' = A_{i0}(z')$ of the type* (a3);]

(f) *$\pi' = \varphi(\pi_1')$, where π_1' is a phrase of G' whose root is Q_j, $1 \leqslant j \leqslant l$.*

Corollary. (i) *$\delta'(\pi') = 1$ if and only if π' is of type* (b).

(ii) *The root of π' is a primed Q_j' if and only if π' is of type* (f).

[(iii) *$\delta'(\pi') = 2$ if π' is of type* (e).]

To confirm that (a)–(f) inductively characterizes the phrases of G', one compares this characterization with the classification of rules of G' given earlier, and with Remark 2 above.

We shall now define by induction on the degree of phrases of G a function f from $P(G)$ to $P'(G')$. We shall at the same time prove that f is one-to-one from $P_n(G)$ onto $P_n'(G')$, and hence, from $P(G)$ onto $P'(G')$. The function f restricted to $K(G)$ will establish a structural homeomorphism between $K(G)$ and $K(G')$.

Let π be a phrase of G of degree 1. We put $f(\pi) = \pi$. Then f is one-to-one from $P_1(G)$ onto $P_1'(G')$ (Lemma 1, Corollary (i)).

We now assume that a one-to-one function f from $\bigcup_{i=1}^{n-1} P_i(G)$ onto $\bigcup_{i=1}^{n-1} P_i'(G')$ is defined, and that π and $f(\pi)$ have the same root. We shall define f for a phrase π of G of degree n, $n > 1$. We have $\pi = P(\pi_0, \pi_1)$, where $P \neq O$, π_0 is a preterminal phrase, and π_1 is a phrase of degree $n - 1$. Let us denote the root of π_1 by Q.

(i) If $Q \neq O$, $\delta(\pi_1)$ is defined and must be $n - 1$. By the induction assumption, $\pi_1' = f(\pi_1)$ is defined and $\delta'(\pi_1') = n - 1$. Consider $\pi' = P(\pi_0, \pi_1')$. Since π_1' has the same root as π_1, π' is a phrase of G' (cf. Lemma 1(c)). We have $\delta(\pi') = \delta(\pi_1') + 1 = n$.

(ii) If $Q = O$, then $P = P_i$ for some i, $1 \leqslant i \leqslant k$. We have $\pi_0 = A_i(x)$ (cf. rule (d1) of G). We distinguish two cases concerning π_1.

(iia) $\pi_1 = O(\pi_{10}, \pi_{11})$, where the root of π_{11} is Q_j for some j, $1 \leqslant j \leqslant l$. Then $\delta(\pi_{11}) = n - 2$; $\pi_{11}' = f(\pi_{11})$ is defined, $\delta'(\pi_{11}') = n - 2$, and the root of π_{11}' is Q_j. Put $\pi_1' = \varphi(\pi_{11}')$. We can write $\pi_{10} = B_j(y)$ (cf. rule (d2) of G). By Lemma 1(d), $\pi' = P_i(\pi_0', \pi_1')$ is a phrase of G' where $\pi_0' = A_{ij}(xy)$. Put $\pi' = f(\pi)$. Note that $\delta'(\pi') = n$ and that π' has the same root as π.

[(iib) $\pi_1 = O(\pi_{11})$ where $\pi_{11} = B_0(y)$. Then by Lemma 1(e), $\pi' = P_i(\pi_1')$, where $\pi_1' = A_{i0}(xyu)$ is a phrase of G'. We put $\pi' = f(\pi)$. Note that $\delta(\pi) = 2 = \delta(\pi')$ and that π' has the same root as π.]

We have defined $f(\pi)$ for all π of degree n. The map f is one-to-one. This follows from two observations. For one thing, those $f(\pi)$'s defined in case (i) are different from those defined in case (ii). ⟨Note that since $P \neq O$, Q in case (i) cannot be any Q_j⟩ [and those defined in case (iia) are different from those defined in case (iib)]. For another, since f is assumed to be one-to-one for phrases of degree less than n, f is one-to-one within both case (i) and case (ii).

The map f is onto the set of phrases π' of G' such that π' is of degree n, $n > 1$, and the root of π' is not a primed symbol Q_j'. To see this, let π' be such a phrase. We shall use the notation of Lemma 1. According to the lemma and its corollary, π' is either type (c), (d) [or (e)].

If π' is type (c), $\pi' = P(\pi_0', \pi_1')$ and the root Q of π_1' is not Q_j'. Hence, by an induction hypothesis, there is a phrase π_1 of G of degree $n - 1$ such that $\pi_1' = f(\pi_1)$, where the root of π_1 is Q. Thus $\pi = P(\pi_0', \pi_1)$ is a phrase of G of degree n. From the construction of f in case (i) above, we have $\pi' = f(\pi)$.

Next, if π' is of type (d), $\pi' = P_i(\pi_0', \varphi(\pi_1'))$ and π_1' is of degree $n - 2$. By an induction hypothesis there is a phrase π_1 of G of degree $n - 2$ such that $\pi_1' = f(\pi_1)$, where the root of π_1 is Q_j. Since π_0' is of type (a2), there must exist x and y such that $xy = z'$ and $\pi_0 = A_i(x)$ and $\pi_{10} = B_j(y)$ are phrases of G. Put $\bar{\pi}_1 = O(\pi_{10}, \pi_1)$ and $\pi = P_i(\pi_0, \bar{\pi}_1)$; π is a phrase of degree n of G. From the construction of f in case (iia) above, we have $\pi' = f(\pi)$.

[If π' is type (e), then $\pi' = P_i(\pi_1')$, $\pi_1' = A_{i0}(xyu)$ and $\pi_0 = A_i(x)$ and $\pi_{10} = B_0(y)$ are phrases of G. Put $\pi_1 = O(\pi_{10})$; this is a phrase of G. Put $\pi = P_i(\pi_0, \pi_1)$; this is a phrase of G of degree 2. From the construction of f in case (iib), we have $\pi' = f(\pi)$.]

We have completed the construction of a one-to-one function f from $P(G)$ onto $P'(G')$. Note that by restricting f to the sentences of G we obtain a one-to-one function from $K = K(G)$ onto $K' = K(G')$.

To facilitate the exposition of further necessary lemmas, we shall introduce one more function and a convention concerning the singular phrase. The inverse φ^{-1} of function φ is defined on a set of phrases of G' whose root is a Q_j'. We shall extend φ^{-1} to the entire $P(G')$ and define function ψ as follows $\psi(\pi') = \varphi^{-1}(\pi')$ if $\varphi^{-1}(\pi')$ is defined, otherwise $\psi(\pi') = \pi'$. ψ is a function from $P(G')$ onto $P'(G')$. Hence, $f^{-1}\psi$ is a function from $P(G')$ onto $P(G)$.

Finally, we shall agree to extend functions f and ψ to take values at the singular phrase. If π (or π') refers to the singular phrase of G (or G'), $f(\pi)$ (or $\psi(\pi')$) is, by definition, the singular phrase of G'. We also agree to say that the singular phrase of G (the singular phrase of G') is in $P(G)$ ($P'(G')$ and $P(G')$). We define $\delta(\pi) = \delta'(\pi) = 0$ for the singular phrases π and π' of G and G', respectively.

Lemma 2. *Let π_1 and π_2 be phrases of G in $P(G)$, and put $\pi_1' = f(\pi_1)$ and $\pi_2' = f(\pi_2)$. Then if π_2 is a subpath of π_1, either π_2' or $\varphi(\pi_2')$ is a subpath of π_1'. Conversely, let π_1' and π_2' be phrases of G' in $P(G')$ and put $\pi_1 = f^{-1}\psi(\pi_1')$ and $\pi_2 = f^{-1}\psi(\pi_2')$. Then, if π_2' is a subpath of π_1', π_2 is a subpath of π_1.*

Furthermore, in either case, π_2 is a proper subpath of π_1 if and only if π_2' or $\varphi(\pi_2')$ is a proper subpath of π_1'.

The lemma is trivial if either of π_1, $\pi_2(\pi_1', \pi_2')$ is singular. In the case in which neither of π_1, $\pi_2(\pi_2', \pi_2')$ is singular, the lemma will be proved by induction on $d = \delta(\pi_1) - \delta(\pi_2)$ (or on $d' = \delta'(\psi(\pi_1')) - \delta'(\psi(\pi_2')))$. When $d = 0$ (or $d' = 0$), the lemma is again trivial.

Assume, then, π_2 is a subtree of π_1 and $d = 1$. Since π_2 is in $P(G)$ and $d = 1$, the root of π_2 is not O. From the construction of f, $\pi_2' = f(\pi_2)$ is a phrase of $\pi_1' = f(\pi_1)$.

Conversely, assume that π_2' is a subpath of π_1' and $d' = 1$. Assume that the root of π_1' is not a primed Q_j'; then $\pi_1 = f^{-1}\psi(\pi_1') = f^{-1}(\pi_1')$. If we also have that the root of π_2' is not any Q_j', then $\pi_2 = f^{-1}\psi(\pi_2') = f^{-1}(\pi_2')$, and from the construction of f, π_2 must be a subpath of π_1. If, on the other hand, the root of π_2' is a Q_j', then $\pi_2 = f^{-1}\psi(\pi_2')$ is a subpath of π_1 (of degree smaller by two than that of π_1).

Assume that the root of π_1' is some Q_j'. Then the root of $\psi(\pi_1')$ is Q_j and π_2' is a phrase of $\psi(\pi_1')$. Hence, this case reduces to the previous case; i.e., applying the preceding discussion to $\psi(\pi_1')$ and π_2', we see that $\pi_2 = f^{-1}\psi(\pi_2')$ is a subpath of $\pi_1 = f^{-1}\psi(\pi_1') = f^{-1}\psi\psi(\pi_1')$.

Assume next that $d = 2$ and π_2 is a subpath of π_1. If the phase π_3 between π_1 and π_2 is not an O-phrase, our problem obviously reduces to the case with $d = 1$. If π_3 is an O-phrase, from the construction of f it is clear that $\varphi(\pi_2') = \varphi f(\pi_2)$ is a subpath of π_1'.

Finally, as the induction step, assume that $d = n > 2$, that π_2 is a subpath of π_1, and that the first assertion of the lemma is true for $d < n$. Then, there exists a phrase π_3 of π_1 whose root is not O, and such that $\delta(\pi_1) - \delta(\pi_3) < n - 1$ and $\delta(\pi_3) - \delta(\pi_2) \leqslant n - 1$, or $\delta(\pi_1) - \delta(\pi_3) \leqslant n - 1$ and $\delta(\pi_3) - \delta(\pi_2) < n - 1$. Hence, by the induction hypothesis either π_2' or $\varphi(\pi_2')$ is a subpath of π_3', and hence, also of $\varphi(\pi_3')$, and either π_3' or $\varphi(\pi_3')$ is a subpath of π_1'. In any case either π_2' or $\varphi(\pi_2')$ is a subpath of π_1'.

Conversely, assume that $d' = n > 1$, that π_2' is a subpath of π_1', and that the converse part of the lemma is true for $d' < n$. Then there exists a subphrase π_3' of π_1' such that $\delta'(\pi_1') - \delta'(\pi_3') < n$ and $\delta'(\pi_3') - \delta'(\pi_2') < n$. From the induction hypothesis, $\pi_3 = f^{-1}\psi(\pi_3')$ is a subpath of $\pi_1 = f^{-1}\psi(\pi_1')$, and $\pi_2 = f^{-1}\psi(\pi_2')$ is a subpath of π_3. Hence, π_2 is a subpath of π_1.

LEMMA 3. *Let π_1, π_2, $\bar{\pi}_1$, $\bar{\pi}_2$ be phrases of G in $P(G)$ and assume that π_2 and $\bar{\pi}_2$ are subpaths of π_1 and $\bar{\pi}_1$, respectively. Put $\pi_1' = f(\pi_1)$ and $\bar{\pi}_1' = f(\bar{\pi}_1)$. Let π_2' be whichever of $f(\pi_2)$ and $\varphi f(\pi_2)$ that is a subpath of π_1' and let $\bar{\pi}_2'$ whichever of $f(\bar{\pi}_2)$ and $\varphi f(\bar{\pi}_2)$ that is a subpath of $\bar{\pi}_1'$. Then if $\pi_1/\pi_2 = \bar{\pi}_1/\bar{\pi}_2$, we have $\pi_1'/\pi_2' = \bar{\pi}_1'/\bar{\pi}_2'$.*

Conversely, let π_1', π_2', $\bar{\pi}_1'$, $\bar{\pi}_2'$ be phrases of G' and put $\pi_1 = f^{-1}\psi(\pi_1')$, $\pi_2 = f^{-1}\psi(\pi_2')$, $\bar{\pi}_1 = f^{-1}\psi(\bar{\pi}_1')$, and $\bar{\pi}_2 = f^{-1}\psi(\bar{\pi}_2')$. Then, if $\pi_1'/\pi_2' = \bar{\pi}_1'/\bar{\pi}_2'$, we have $\pi_1/\pi_2 = \bar{\pi}_1/\bar{\pi}_2$.

The lemma can be proved by induction on the difference of the degrees of π_1 and π_2, and of π_1' and π_2' in much the same way as in the proof of the preceding lemma. Hence, the proof is here omitted.

Let Π be an arbitrary pruning set of G. We shall define a pruning set Π' of G' such that f is continuous from $\mathbf{T}_{\Pi}(K)$ onto $\mathbf{T}_{\Pi'}(K')$. Let π be a phrase in Π. If the root of π is not O, let $f(\pi)$ be in Π'; furthermore, if the root of π is some Q_j, also let $\varphi f(\pi)$ be in Π'. If the root of π is O, then we have $\pi = O(\pi_0, \pi_1)$, where the root of π_1 is some Q_j, and where $\pi_0 = B_j(y)$ [or $\pi = O(\pi_0)$, $\pi_0 = B_0(y)$]. Put $\bar{\pi}_1 = \varphi f(\pi_1)$. For each i, $1 \leqslant i \leqslant k$, and for each x such that $A_i \rightarrow x$ is a rule of G, put $\bar{\pi}_0^{(i,x)} = A_{ij}(xy)$, which is a phrase of G'. Finally, put $\pi'^{(i,x)} = P_i(\bar{\pi}_0^{(i,x)}, \bar{\pi}_1)$ [or put $\bar{\pi}_0^{(i,x)} = A_{i0}(xyu)$, which is a phrase of G'; finally put $\pi'^{(i,x)} = P_i(\bar{\pi}_0^{(i,x)})$]; let each $\pi'^{(i,x)}$ be in Π'. If $P_i = Q_j$ for some j, let $\varphi(\pi'^{(i,x)})$ be also in Π'. No other phrases of G' shall be in Π'.

Let Π and Π' be as above and consider an arbitrary sentence τ in K. We put $\pi = \pi(\tau, \Pi)$ and $\tau' = f(\tau)$. We shall try to characterize $\bar{\pi}' = \pi(\tau', \Pi')$. If the root of π is O, then there is a unique phrase $\hat{\pi}$ of τ such that we have $\hat{\pi} = P_i(\hat{\pi}_0, \pi)$ for some i, $1 \leqslant i \leqslant k$. Otherwise, put $\hat{\pi} = \pi$. From Lemma 2, either $f(\hat{\pi})$ or $\varphi f(\hat{\pi})$ is a subpath of τ. Call $\hat{\pi}'$ whichever of them is a subpath of τ'. Either $\hat{\pi}'$ is singular (i.e., $\hat{\pi}$ is singular), or else it is in Π', by the construction of Π'. Hence, $\hat{\pi}'$ is a subpath of $\bar{\pi}' = \pi(\tau', \Pi')$. From the converse part of the same lemma, $f^{-1}\psi(\hat{\pi}')$ is a subpath of $f^{-1}\psi(\bar{\pi}')$, and is a proper subpath if and only if $\hat{\pi}'$ is a proper subpath of $\bar{\pi}'$. Note, on the one hand, that $f^{-1}\psi(\hat{\pi}') = \hat{\pi}$. On the other hand, from the construction of Π', if $\bar{\pi}'$ is in Π' (i.e., if $\bar{\pi}'$ is not singular), then either $\bar{\pi} = f^{-1}\psi(\bar{\pi}')$ is in Π or else $\bar{\pi}$ is of the form $\bar{\pi} = P_i(\bar{\pi}_0, \bar{\pi}_1)$, where $\bar{\pi}_1$ is an O-phrase in Π. From the meaning of $\hat{\pi}$ it follows that $\hat{\pi} = \bar{\pi}$. If $\bar{\pi}'$ is singular, $\bar{\pi}$ and consequently $\hat{\pi}$, is singular, and $\hat{\pi} = \bar{\pi}$. Hence, in any case we have $\hat{\pi}' = \bar{\pi}'$. We state the result as

LEMMA 4. *Given an arbitrary τ in K and $\tau' = f(\tau)$, we have $\pi(\tau', \Pi') = \hat{\pi}'$, where $\hat{\pi}' = f(\hat{\pi})$ or $\varphi f(\hat{\pi})$ (whichever of them is a phrase of τ'), and where $\hat{\pi}$ is determined from $\pi(\tau, \Pi)$ as above.*

We are now in a position to prove the main lemma.

LEMMA 5. *f is continuous from $\mathbf{T}_{\Pi}(K)$ to $\mathbf{T}_{\Pi'}(K')$. Moreover, if Π does not contain any O-phrases, f is a homeomorphism.*

Let τ_1 and τ_2 be two sentences of K such that $(\tau_1)_{\Pi} < (\tau_2)_{\Pi}$. We put $\pi_1 = \pi(\tau_1, \Pi)$ and $\pi_2 = \pi(\tau_2, \Pi)$; we have $\tau_1/\pi_1 \subset \tau_2/\pi_2$. We put $\tau_1' = f(\tau_1)$

and $\tau_2' = f(\tau_2)$. Let $\hat{\pi}_1$, $\hat{\pi}_1'$, $(\hat{\pi}_2, \hat{\pi}_2')$ be to $\pi_1(\pi_2)$ what $\hat{\pi}$ and $\hat{\pi}'$ are to π in Lemma 4. We have $\tau_1/\hat{\pi}_1 \subset \tau_1/\pi_1 \subset \tau_2/\pi_2$. There exists a subpath $\bar{\pi}_2$ of τ_2 (of which π_2 is a subpath) such that $\tau_1/\hat{\pi}_1 = \tau_2/\bar{\pi}_2$. Since $\hat{\pi}_1$, and hence, also $\bar{\pi}_2$, are not O-phrases, according to Lemma 3 we have $\tau_1'/\hat{\pi}_1' = \tau_2'/\bar{\pi}_2'$, where $\bar{\pi}_2'$ is whichever of $f(\bar{\pi}_2)$ and $\varphi(f(\bar{\pi}_2))$ that is a subpath of τ_2'. From Lemma 4 we have $\hat{\pi}_1' = \pi(\tau_1', \Pi')$. Hence $(\tau_1')_{\Pi'} < (\tau_2')_{\Pi'}$ if and only if $\hat{\pi}_2' = \pi(\tau_2', \Pi')$ is a subpath of $\bar{\pi}_2'$. Now, $\hat{\pi}_2$ is a subpath of $\bar{\pi}_2$; both $\hat{\pi}_2$ and $\bar{\pi}_2$ are not O-phrases. Hence, from Lemma 2, $f(\hat{\pi}_2)$ or $\varphi f(\hat{\pi}_2)$ is a subpath of $f(\bar{\pi}_2)$, and hence, one of them is a subpath of $\varphi f(\bar{\pi}_2)$. Consequently, either $f(\hat{\pi}_2)$ or $\varphi f(\hat{\pi}_2)$ is a subpath of $\bar{\pi}_2'$. Whichever of them is a subpath of $\bar{\pi}_2'$ is a subpath of τ_2', since $\bar{\pi}_2'$ is a subpath of τ_2'. Hence this subpath is $\hat{\pi}_2'$ and from Lemma 4, $\hat{\pi}_2' = \pi(\tau_2', \Pi')$; i.e., $\pi(\tau_2', \Pi')$ is a subpath of $\bar{\pi}_2'$.

Conversely, assume that $(\tau_1')_{\Pi'} < (\tau_2')_{\Pi'}$. Then there exists a subpath $\bar{\pi}_2'$ (of which $\hat{\pi}_2'$ is a subpath) such that $\tau_1'/\hat{\pi}_1' = \tau_2'/\bar{\pi}_2'$. (Recall that $\hat{\pi}_1' = \pi(\tau_1', \Pi')$.) Hence, by the converse part of Lemma 3, $\tau_1/\hat{\pi}_1 = \tau_2/\bar{\pi}_2$, where $\bar{\pi}_2 = f^{-1}\psi(\bar{\pi}_2')$. Since $\hat{\pi}_2 = f^{-1}\psi(\hat{\pi}_2')$, from the fact that $\hat{\pi}_2'$ is a subpath of $\bar{\pi}_2'$, by Lemma 2 it follows that $\hat{\pi}_2$ is a subpath of $\bar{\pi}_2$. Now, if Π does not contain an O-phrase, then $\pi_1 = \hat{\pi}_1$, and $\pi_2 = \hat{\pi}_2$. Hence, we have $\tau_1/\pi_1 = \tau_2/\bar{\pi}_2$, and π_2 is a subpath of $\bar{\pi}_2$. It follows that $(\tau_1)_{\Pi} < (\tau_2)_{\Pi}$; i.e., f^{-1} is continuous. The lemma is proved.

Lemma 5 shows that f is structurally continuous. What is left to be proved is that f^{-1} is structurally continuous. We now extend φ over all of $P(G')$; we define $\bar{\varphi}(\pi) = \pi$ if the root of π is neither Q_j nor Q_j', $\bar{\varphi}(\pi) = \varphi(\pi)$ if the root of π is any Q_j, and $\bar{\varphi}(\pi) = \varphi^{-1}(\pi)$ if the root of π is any Q_j'. Let $\bar{\Pi}'$ be an arbitrary pruning set of K'. Define $\bar{\Pi}_1' = \bar{\Pi}' \cup \bar{\varphi}(\bar{\Pi}') \cup \{\bar{\pi}_1'; (\bar{\pi}_1')_{\bar{\Pi}'} < \bar{\pi}'$ for some $\bar{\pi}' \in \bar{\Pi}' \cup \bar{\varphi}(\bar{\Pi}')\}$. $\bar{\Pi}_1'$ is obviously an extension of $\bar{\Pi}'$. Let Π' be the extension of $\bar{\Pi}_1'$ constructed from $\bar{\Pi}'$ as in Theorem 1, Part I. The identity map of K' is then continuous from $\mathbf{T}_{\bar{\Pi}'}$ to $\mathbf{T}_{\Pi'}$. We note that $\Pi' = \bar{\varphi}(\Pi')$. Indeed, take an arbitrary π' in Π'. By the construction of Π', there exists $\bar{\pi}_1'$ in $\bar{\Pi}_1'$ such that $(\pi')_{\bar{\Pi}'} < (\bar{\pi}_1')_{\bar{\Pi}'}$. (i) Assume $\bar{\pi}_1' \in \bar{\Pi}'$. Then $(\bar{\varphi}(\pi'))_{\bar{\Pi}'} < \bar{\varphi}(\bar{\pi}_1')$; hence, $\bar{\varphi}(\pi') \in \bar{\Pi}_1'$ and a fortiori $\bar{\varphi}(\pi') \in \Pi'$. (ii) Assume $\bar{\pi}_1' \in \bar{\varphi}(\bar{\Pi}')$. Then $(\bar{\varphi}(\pi'))_{\bar{\Pi}'} < \bar{\varphi}(\bar{\pi}_1')$ and $\bar{\varphi}(\bar{\pi}_1') \in \bar{\Pi}'$; hence, $\bar{\varphi}(\pi') \in \Pi'$. (iii) Assume $(\bar{\pi}_1)_{\bar{\Pi}'} < \bar{\pi}'$ for some $\bar{\pi}' \in \bar{\Pi}' \cup \bar{\varphi}(\bar{\Pi}')$. Then $(\bar{\varphi}(\bar{\pi}_1'))_{\bar{\Pi}'} < \bar{\varphi}(\bar{\pi}')$; hence, $\bar{\varphi}(\bar{\pi}_1') \in \bar{\Pi}_1'$ and $\bar{\varphi}(\pi') \in \Pi'$.

Now define a pruning set Π of G as follows: $\Pi = \{\pi; f(\pi) \in \Pi'\}$. Π does not contain any O-phrase. Moreover, the Π' corresponding to Π defined for Lemma 5 is nothing but the Π' given here. Hence, from that lemma it follows that f is a homeomorphism from $\mathbf{T}_{\Pi}(K)$ onto $\mathbf{T}_{\Pi'}(K')$, and f^{-1} is continuous from $\mathbf{T}_{\bar{\Pi}'}(K')$ onto $\mathbf{T}_{\Pi}(K)$. This means that f^{-1} is structurally continuous from K' onto K. This completes the proof of Theorem 2.

⟨COROLLARY. *Let us use notations in the same way as in Theorem 2. Assume that G has no rule of the form* $P_i \to AQ_j$. *Let* $\bar{G}$ *be a standard regular grammar obtained from G by removing the nonterminal symbol O (and all the rules in which O appears), by adding preterminal symbols* $\bar{A}_{ij}$, *and by adding the following rules: For each pair of rules of G,* $P_i \to A_i O$ *and* $O \to B_j Q_j$, *a new rule* $P_i \to \bar{A}_{ij} Q_j$; *for each pair of rules* $A_i \to x$ *and* $B_j \to y$ *of G, a new rule* $\bar{A}_{ij} \to xy$. *Then* $K(\bar{G})$ *is structurally homeomorphic to* $K(G)$.

By the convention introduced before, Q_j' may be considered as another name for Q_j under the assumption of the corollary. Then the grammar $\bar{G}$ of the corollary is just the grammar G' of Theorem 2.⟩

4. *Applications of the Elimination Theorem*

We shall first introduce the notion of simple cycle. Let G be a standard regular grammar. A path of G is said to be *closed* if its origin and endpoint are the same state; a closed path is said to be *simple* if no two of its states are identical, except for the origin and endpoint. Two closed paths are called *equivalent* if they are of the same length and contain the same set of unit paths. An equivalence class of equivalent closed paths is called a *cycle*. (Thus the notion of cycle is obtained from that of closed path by leaving the identity of the origin-endpoint unspecified.) We may talk about the length, states, and unit paths of a cycle in the obvious sense. Generally, we may without fear of confusion denote a cycle by a closed path which is a representative of it; for example, we may say a cycle $(P_0, P_1, ..., P_n)$, where $P_0 = P_n$. A *simple cycle* is by definition a cycle of which a representative is a simple closed path (and consequently, of which all the representatives are simple closed paths). Finally, a simple cycle is said to be *separated* if for each of its states it is the only simple cycle that passes through that state.

THEOREM 3. *For each standard regular grammar G, there exists a standard regular grammar G′ such that G′ has no separated simple cycle of length more than* 1, *and such that K(G) is structurally homeomorphic to K(G′).*

Assume that G has a separated simple cycle γ of length more than 1. Let O be a state on γ. Since γ is of length more than 1 and it is the only simple cycle passing through O, O cannot be directly self connected. Note also that we may take O as a state which is not the initial state. Therefore, we may apply Theorem 2 with respect to O. Let G' be the grammar thus obtained. We will use the same notations as in the statement of Theorem 2. Let γ be represented as $\gamma = (R_0, R_1, ..., R_n)$, $n > 1$, $R_0 = R_n$, $R_g \neq R_h$, $0 < g < h < n$, and assume $R_{n-1} = O$. Then, $R_0 = R_n = Q_{j_0}$ for some j_0, $1 \leqslant j_0 \leqslant l$. In

G', γ is replaced by a cycle of length $n - 1$, i.e., by $\gamma' = (R_0', R_1',..., R'_{n-1})$, where $R_0' = R'_{n-1} = Q'_{j_0}$, $R_i' = R_i$, $1 \leqslant i \leqslant n - 2$. Furthermore, γ' is the only simple cycle of G' which is not a simple cycle in G. For if a simple cycle $\bar{\gamma}'$ of G' does not contain any primed state Q_j', then $\bar{\gamma}'$ is certainly a cycle in G. Assume then that $\bar{\gamma}'$ contains some primed state Q_j'. Let $\bar{\gamma}'$ be represented as $\bar{\gamma}' = (\bar{R}_0', \bar{R}_1',..., \bar{R}_m')$, $\bar{R}_0' = \bar{R}_m'$, $\bar{R}_g' \neq \bar{R}_h'$, $0 \leqslant g < h < m$, and assume that $\bar{R}_0' = \bar{R}_m'$, $\bar{R}'_{i_1}$, $\bar{R}'_{i_2}$,..., $\bar{R}'_{i_h}$, $0 < i_1 < i_2 < \cdots < i_h < m$, are those $\bar{R}_i'$ that are primed Q_j''s. Then, in the expression $\bar{\gamma}' = (\bar{R}_0', \bar{R}_1',..., \bar{R}_m')$ replace $\bar{R}_0' = \bar{R}_m'$, $\bar{R}'_{i_1}$,..., $\bar{R}'_{i_h}$ by the corresponding nonprimed Q_j's, and insert O in front of each of $\bar{R}'_{i_1}$, $\bar{R}'_{i_2}$,..., $\bar{R}'_{i_h}$, and $\bar{R}_m'$. The expression thus obtained represents a closed cycle $\bar{\gamma}$ on G passing through O $h + 1$ times. From the assumption that γ is a separated simple cycle, we must have $\bar{\gamma} = \gamma^{h+1}$. But then $\bar{R}'_{i_1} = \bar{R}'_{i_2} = \cdots = \bar{R}'_{i_h} = \bar{R}_m' = Q'_{j_0}$, and, since $\bar{\gamma}'$ is simple, $h = 0$, and consequently $\bar{\gamma} = \gamma$ and $\bar{\gamma}' = \gamma'$.

From this argument it follows that G' has as many separated simple cycles as G. One of them has length less by 1 than that of the corresponding cycle on G, and the others have the same length as their corresponding cycles. Hence, by continuing this process we can eliminate all separated simple cycles of G of length more than 1. The theorem is proved.

Let P be a nonfinal, nonself-connected state of a standard regular grammar G, and denote by $\Sigma(P)$ the set of those nonfinal states to which P is connected. Assume that $\Sigma(P)$ contains another nonself-connected state P'. Then $\Sigma(P')$ is a proper subset of $\Sigma(P)$. Let us say that a nonfinal nonself-connected state P is *minimal* if $\Sigma(P)$ contains no nonself-connected state. Clearly, if there exists a nonfinal nonself-connected state in G, there exists a minimal one. As another application of the elimination theorem we are now going to eliminate noninitial nonfinal nonself-connected states. For this purpose it suffices to show that a noninitial minimal nonself-connected state can be eliminated without creating new nonself-connected states.

Let O be a minimal nonselfconnected state of G. We cannot apply Theorem 2 directly to G and O. Suppose that we try to do so. Using the same notations as in Theorem 2, consider one Q_j', $1 \leqslant j \leqslant l$. Q_j' is not self-connected. For if it were, it would be connected to P_i, $1 \leqslant i \leqslant k$, because these are the only states connected to Q_j'. It follows then that Q_j is connected to P_i in G, and hence, O is self-connected, contrary to assumption. Hence, the effect of eliminating a nonself-connected state O is to create some other nonself-connected states Q_j'.

Assume, however, that no P_i is directly connected to any Q_j in G. Then we may apply the corollary of Theorem 2 and construct a grammar $\bar{G}$ such that O is eliminated without introducing any new states, and such that $K(G)$ and

$K(\bar{G})$ are structurally homeomorphic. $\bar{G}$ has one less nonself-connected state than G by one. Note, incidentally, that no new cycles are introduced in this process of reduction; if G has no separated simple cycle of length more than one, neither does $\bar{G}$.

To return to the general case, let us prove the following.

LEMMA. *Let O be a nonself-connected state of a standard regular grammar G, and let $\Sigma(O)$ be the set of nonfinal states to which O is connected. Let a standard regular grammar G' be defined as follows: The states of G' are the states of G and primed copies R''s for each R in $\Sigma(O)$. The rules of G' are* (i) *the rules of G except for those of the form $O \rightarrow AQ$;* (ii) *$O \rightarrow AQ'$ for each rule $O \rightarrow AQ$ of G;* (iii) *$R_k' \rightarrow AR_l'$ (or $R_k' \rightarrow A$) for each rule $R_k \rightarrow AR_l$ (or $R_k \rightarrow A$) of G, respectively, where R_k and R_l are states in $\Sigma(O)$. The initial state of G' is the initial state of G. Then $K(G')$ is structurally homeomorphic to $K(G)$.*

A structural homeomorphism from $K' = K(G')$ onto $K = K(G)$ can be obtained from the map f which is defined as follows on trees belonging to G': For each tree T' belonging to G', let $f(T')$ be the tree obtained from T' by replacing each primed symbol in T' by the corresponding nonprimed symbol. The map f is not one-to-one, but its restriction to K' is one-to-one and onto K. For each pruning set Π' of K', f is continuous from $\mathbf{T}_{\Pi'}$ to $\mathbf{T}_{f(\Pi')}$; for each pruning set Π of K, f^{-1} is continuous from $\mathbf{T}_{\Pi}$ to $\mathbf{T}_{f^{-1}(\Pi)}$.

Clearly G' does not contain any cycle of length greater than the maximum length for cycles in G. In particular, if G does not contain any separated simple cycle of length more than 1, then neither does G'.

It is obvious that each primed state R' is self-connected in G' if and only if R is self-connected in G, and that state P of G is self-connected in G if and only if it is self-connected in G'. In particular, O is a minimal nonself-connected state in G' if and only if it is a minimal nonself-connected state in G and the number of minimal nonself-connected states in G is the same as in G'. Note that in G' the states to which O is directly connected are primed states and no states other than O is directly connected to them. Hence, we may apply Theorem 2 to eliminate O without introducing a new nonself-connected state. We conclude

THEOREM 4. *For each standard regular grammar G, there exists a standard regular grammar G' such that noninitial nonfinal states of G' are all self-connected, and such that $K(G)$ and $K(G')$ are structurally homeomorphic.*

COROLLARY. *We may specify in addition that G' does not contain any separated simple cycle of length more than one.*

5. *Principal Standard Regular Grammars*

Let G be a standard regular grammar and let Π be a pruning set of G. A nonfinal state P of G is called *principal with respect to* Π if for any path θ from the initial state of G to P, there exists a sentence τ of $K(G)$ such that $\theta = \theta(\tau)_\Pi$. Π is called *principal* if all nonfinal states of G are principal with respect to Π. Finally, G is called *principal* if the pruning set $\Pi_0(G)$ is principal, where $\Pi_0(G)$ consists of all the phrases π of G such that the root of π is the only node of π labeled with a nonpreterminal (i.e., which are of the form $P(A(x))$). Clearly G is principal if and only if each nonfinal state of G is directly connected to the final state.

Again, let G be a standard regular grammar. For each nonfinal state P of G, take a phrase $\pi(P)$ whose root is P. To the rules of G, add rules $P \rightarrow A_P$ and $A_P \rightarrow x$, where x is the terminal string of $\pi(P)$. The grammar G' thus obtained is principal. It might appear that $K(G)$ and $K(G')$ are canonically structurally homeomorphic; each sentence τ of G which contains a $\pi(P)$ can be mapped to the sentence τ' of G' obtained from τ by replacing $\pi(P)$ by the phrase $P(A_P(x))$ of G'. However, this procedure does not establish a homeomorphism between $K(G)$ and $K(G')$. For if some $\pi(P)$ is a subphrase of another $\pi(P')$, a sentence τ of G containing $\pi(P')$ as a phrase is, so to speak, generated twice by G', once using the rule $P \rightarrow A_P$ and once more using the rule $P' \rightarrow A_{P'}$. We need somewhat more elaborate means to construct a principal grammar G' such that $K(G')$ is structurally homeomorphic to $K(G)$.

LEMMA 1. *A pruning set* Π *is principal if* (1) *for each nonfinal state* P *there exists in* Π *at least one phrase whose root is* P, *and* (2) *no phrase of* Π *is a proper subphrase of any other phrase of* Π.

Given a path θ whose origin is the initial state and whose endpoint is P, there is a phrase π whose root is P and a sentence τ, of which π is a phrase, such that $\theta(\tau) = \theta \cdot \theta(\pi)$. Then $\theta(\tau)_\Pi = \theta$. Hence, Π is principal. Q.E.D.

Let G be a standard regular grammar such that the noninitial nonfinal states of G are all self-connected. Let Π be a pruning set of G which satisfies condition (1) of the above lemma and also the condition that no two phrases in it have the same root. If Π is not principal, Π does not satisfy condition (2) of the lemma and there must exist a series of phrases of Π, $(\pi_i)_{1 \leqslant i \leqslant k}$, $k \geqslant 2$, such that π_i is a proper subphrase of π_{i+1}. Let us call such a series *singular*. We shall try to eliminate singular series from Π without abandoning condition (1).

Assume that a singular series $(\pi_i)_{1 \leqslant i \leqslant k}$ is maximal (i.e., there is no singular series of which it is a proper subset). Let P_i be the root of π_i. If P_1 is not the initial state, then it must be self-connected, by our assumption on G above. If P_1 is the initial state, π_1 is then a sentence and is a proper subphrase of π_2, which in turn is a subphrase of some sentence; thus the initial state is self-connected. In any case, then, P_1 is self-connected. Hence, there exists a closed path from P_1 to P_1.

First we assume that there are two different closed paths from P_1 to P_1, γ_0 and γ_1, neither of which is contained in the other. Let us encode the numbers 1 through k in the binary notations; i.e., take a sufficiently large h and let $\epsilon(i) = \epsilon_{i1}\epsilon_{i2} \cdots \epsilon_{ih}$, $\epsilon_{ij} = 0$ or 1, $\epsilon(i) \neq \epsilon(j)$ for $i \neq j$, $1 \leqslant i, j \leqslant k$. Define $\Gamma_i = \gamma_{\epsilon_{i1}}\gamma_{\epsilon_{i2}} \cdots \gamma_{\epsilon_{ih}}$ for i, $1 \leqslant i \leqslant k$. Γ_i is a closed path from P_1 to P_1, and $\Gamma_i \neq \Gamma_j$ for $i \neq j$. Put $\theta_i = \theta(\pi_i)/\theta(\pi_1)$ and $\eta_i = \theta_i\Gamma_i\theta(\pi_1)$. η_i is a path from P_i to the final state. Take π_i' such that $\theta(\pi_i') = \eta_i$. Then, for any $i \neq j$, π_i' is not a subphrase of π_j'. Hence, if we replace π_i by π_i', we seem to eliminate one maximal singular series of Π, keeping the condition (1) intact.

It may happen, however, that an inclusion relationship holds between some π_i' and some π in Π which is not contained in the series $(\pi_i)_{1 \leqslant i \leqslant k}$; that is, either π_i' is a subphrase of π or π is a subphrase of π_i'. In that case, we eliminate one maximal singular series of Π at the expense of introducing some new ones. We want to avoid this situation.

Assume that π contains π_i' as a subphrase; then $\theta(\pi)$ must be of the form $\theta\Gamma_i\theta(\pi_1)$ where $\theta \supset \theta_i$ and the endpoint of θ is P_1. Assume, on the other hand, that π is contained in π_i' as a subphrase; then $\theta(\pi)$ must be either of the form $\theta\Gamma_i\theta(\pi_1)$, where $\theta \subset \theta_i$ and the endpoint of θ is P_1, or else of the form $\Gamma_i'\theta(\pi_1)$, where $\Gamma_i' \subset \Gamma_i$ and the endpoint of Γ_i' is P_1. (Note that we may assume π not to be contained in π_1; otherwise the singular series $(\pi_i)_{1 \leqslant i \leqslant k}$ could be extended, contrary to the assumption that it is maximal.)

Consider, then, the set of all phrases $\bar{\pi}$ in Π that satisfy the following condition: $\theta(\bar{\pi}) = \bar{\theta}\bar{\Gamma}\theta(\pi_1)$, where $\bar{\theta}$ is a path such that $\bar{\theta} \supset \theta_i$ or $\bar{\theta} \subset \theta_i$ for some i, or $\bar{\theta} \subset \gamma_0$, or $\bar{\theta} \subset \gamma_1$, and where, for some l, $l \geqslant 0$, $\bar{\Gamma} = \gamma_{i_1}\gamma_{i_2} \cdots \gamma_{i_l}$, $\gamma_{i_j} = \gamma_0$ or γ_1 for $1 \leqslant j \leqslant l$. (For $l = 0$, we assume that $\bar{\Gamma}$ is the null path.) Let $\bar{\pi}_1, \bar{\pi}_2, \ldots, \bar{\pi}_{\bar{k}}$ be the phrases of this set. Our original π_i, $1 \leqslant i \leqslant k$, are among these phrases. We apply to this set the same technique as above. Thus, let $\bar{\epsilon}(i)$ be a binary coding of numbers 1 through $\bar{k}$ and let $\bar{\Gamma}_i = \gamma_{\bar{\epsilon}_{i1}}\gamma_{\bar{\epsilon}_{i2}} \cdots \gamma_{\bar{\epsilon}_{ih}}$, where $\bar{\epsilon}(i) = \bar{\epsilon}_{i1}\bar{\epsilon}_{i2} \cdots \bar{\epsilon}_{i\bar{h}}$. We consider the paths $\bar{\eta}_i = \bar{\theta}_i\bar{\Gamma}_i\theta(\pi_1)$, where $\bar{\theta}_i$ is given from the decomposition $\theta(\bar{\pi}_i) = \bar{\theta}_i\bar{\Gamma}\theta(\pi_1)$, and we take a phrase $\bar{\pi}_i'$ such that $\theta(\bar{\pi}_i') = \bar{\eta}_i$. No $\bar{\pi}_i'$ is contained in any $\bar{\pi}_j'$, $j \neq i$. More generally, no $\bar{\pi}_i'$ contains, or is contained in, any other phrase π which is not any $\bar{\pi}_i$; for, if $\bar{\pi}_i'$ contains π (or is contained in π), π must be one of $\bar{\pi}_i$'s. Hence, by replacing $\bar{\pi}_i$ by $\bar{\pi}_i'$ for all i, $1 \leqslant i \leqslant \bar{k}$, we can eliminate one maximal singular series of Π without creating any new ones.

Next we assume that there do not exist two closed paths γ_0 and γ_1 from P_1 to P_1 such that neither of them is contained in the other. Since P_1 is self-connected, there must exist one, and only one, simple cycle through P_1; call it γ. Furthermore, it must also be true of each state on γ that γ is the only simple cycle through it; that is, γ is a separated simple cycle.

Assume that for some i, $1 \leqslant i \leqslant k$, P_i is not on γ. Then, for all j, $i \leqslant j \leqslant k$, P_j is not on γ. For, supposing P_j were on γ, let θ be the subpath of γ from P_1 to P_j and put $\gamma' = \theta \cdot \theta_j$, where $\theta_j = \theta(\pi_j)/\theta(\pi_1)$. γ' is (or, more exactly, represents) a cycle through P_1 which is not a multiple of γ (since P_i is on γ' but not on γ). This contradicts the assumption that γ is a separated simple cycle.

Hence $(\pi_i)_{1 \leqslant i \leqslant k}$ may be divided into two series $(\pi_i)_{1 \leqslant i \leqslant k_1}$ and $(\pi_j)_{k_1 < j \leqslant k}$ in such a way that each P_i, $1 \leqslant i \leqslant k_1$, is on γ and each P_j, $k_1 < j \leqslant k$, is not on γ. (Here, $k_1 \geqslant 1$ and the subseries $(\pi_i)_{1 \leqslant i \leqslant k_1}$ is not empty, since P_1 is on γ; on the other hand, it may be that $k_1 = k$, in which case the subseries $(\pi_j)_{k_1 < j \leqslant k}$ is empty.)

Let us use the symbol γ to also denote the closed path from P_1 to P_1 representing the cycle γ. Then, for each π_i, $1 \leqslant i \leqslant k_1$, we can have the following expression: $\theta(\pi_i) = \bar{\gamma}_i \gamma^{n_i} \theta(\pi_1)$, where $\bar{\gamma}_i \subset \gamma$, $\bar{\gamma}_i \neq \gamma$, and the endpoint of $\bar{\gamma}_i$ is P_1 and where n_i is the number of times $\theta(\pi_i)$ circles around γ. Similarly, for each π_j, $k_1 < j \leqslant k$, we put $\theta(\pi_j) = \bar{\theta}_j \bar{\gamma}_j \gamma^{\bar{n}} \theta(\pi_1)$, where $\bar{\gamma}_j \subset \gamma$, $\bar{\gamma}_j \neq \gamma$, and the endpoint of $\bar{\gamma}_j$ is P_1 and where $\bar{n}$ is the number of times $\theta(\pi_{k_1+1})$, and hence also $\theta(\pi_j)$, circles around γ.

Now let us consider those π's in Π, whether they are in $(\pi_i)_{1 \leqslant i \leqslant k}$ or not, which are of the form either (a) $\theta(\pi) = \bar{\gamma} \gamma^n \theta(\pi_1)$ or (b) $\theta(\pi) = \theta \bar{\gamma} \gamma^n \theta(\pi_1)$, where $\bar{\gamma} \subset \gamma$, $\bar{\gamma} \neq \gamma$, the endpoint of $\bar{\gamma}$ is P_1, n is the number of times $\theta(\pi)$ circles around γ, and where θ is a path whose endpoint is on γ but whose other states are not on γ. We note that π_i, $1 \leqslant i \leqslant k_1$ are among the phrases of type (a) and π_j, $k_1 < j \leqslant k$, are among those of type (b). Let the phrases of type (a) be enumerated as $\kappa_s = \bar{\gamma}_s \gamma^{n_s} \theta(\pi_1)$, $1 \leqslant s \leqslant h$, and those of type (b) as $\lambda_t = \theta_t \bar{\gamma}_t \gamma^{n_t} \theta(\pi_1)$, $h < t \leqslant g$. Let N be the maximum of n_t, $h < t \leqslant g$. Put $\bar{\kappa}_s = \bar{\gamma}_s \gamma^{N+1} \theta(\pi_1)$, $1 \leqslant s \leqslant h$. Note that $\bar{\kappa}_s$ has the same root as κ_s and that no $\bar{\kappa}_s$ contains any λ_t as a subphrase and, conversely, that no λ_t contains any $\bar{\kappa}_s$ as a subphrase.

Let us replace κ_s by $\bar{\kappa}_s$ in Π. This replacement has the following effect on the maximal singular series $(\pi_i)_{1 \leqslant i \leqslant k}$. Each π in the subseries $(\pi_i)_{1 \leqslant i \leqslant k_1}$ is replaced by some $\bar{\kappa}_s$, $1 \leqslant s \leqslant h$. The subseries $(\pi_j)_{k_1 < j \leqslant k}$ now constitutes a maximal singular series with fewer terms than the original $(\pi_i)_{1 \leqslant i \leqslant k}$. (Recall that $k_1 \geqslant 1$.) On the other hand, $\bar{\kappa}_s$, $1 \leqslant s \leqslant h$, constitutes a maximal singular series, and the subseries $(\pi_i)_{1 \leqslant i \leqslant k_1}$, now replaced by certain $\bar{\kappa}_s$'s, is absorbed in that maximal singular series.

From these observations we conclude

Lemma 2. *Let G be a standard regular grammar whose noninitial nonfinal states are all self-connected. Then there exists a pruning set Π of G, satisfying condition* (1) *of Lemma* 1, *such that for each singular series*, $(\pi_i)_{1 \leqslant i \leqslant k}$ *of Π, there is a separated simple cycle γ such that the roots of π_i are all on γ.*

Theorem 5. *Let G be a standard regular grammar whose noninitial and nonfinal states are all self-connected and which does not have a separated simple cycle of length more than one. Then G has a principal pruning set.*

Theorem 5 follows directly from Lemma 2, for no singular series can lie on a cycle of length 1. (Note that a singular series must have at least two terms.)

Corollary 1. *For any standard regular grammar G, there exists a standard regular grammar G' such that $K(G)$ and $K(G')$ are structurally homeomorphic and G' has a principal pruning set.*

This corollary follows from Theorem 5 together with Theorem 4 and its corollary.

Corollary 2. *In Corollary* 1 *we may further assume that the principal pruning set mentioned satisfies conditions* (1) *and* (2) *of Lemma* 1.

This follows from the proof of Theorem 5.

Theorem 6. *For each standard regular grammar G there exists a principal standard grammar G' such that $K(G)$ and $K(G')$ are structurally homeomorphic.*

Let $\bar{G}$ be a standard regular grammar such that $K(G)$ and $K(\bar{G})$ are struturally homeomorphic and such that it has a principal pruning set Π satisfying the conditions of Lemma 1. (Cf. Theorem 5, Corollary 2.) We may assume that $\pi(\tau, \Pi)$ is singular for only a finite number of sentences τ of $\bar{G}$. For, assume $\pi(\tau, \Pi)$ is singular. Then no phrase of Π is a phrase of τ. Now if the height of τ is larger than the maximum H of the heights of the phrases of Π, let κ be the phrase of τ of height H. No phrase of Π is a subphrase of κ, and conversely κ is not a subphrase of any phrase of Π. Consider all the phrases of height H obtained in this way. There obviously exist a finite number of them, and none of them is a proper subphrase of any other, as their heights are all equal. Furthermore, none of them is a subphrase of any phrase of Π, and no phrase of Π is a subphrase of any of them. Hence, if we add all those phrases of height H to Π, then Π, thus expanded, still satisfies the conditions of Lemma 1; it is principal. And with respect to this new Π, no sentence τ of height more than H has singular $\pi(\tau, \Pi)$. It follows that there are only a finite number of sentences, τ_i, $1 \leqslant i \leqslant m$, such that $\pi(\tau_i, \Pi)$ is singular.

We shall construct a new grammar G' from $\bar{G}$ as follows. The states of G' are the states of $\bar{G}$ plus the new initial state S'. (The initial state $\bar{S}$ of $\bar{G}$ is a state of G', but is not the initial state of G'.) The preterminals of $\bar{G}$ which are not end preterminals (i.e., those which appear in rules of the form $P \rightarrow AQ$) are preterminals of G'. Besides these we introduce one more nonend preterminal $\bar{A}$ (which will appear in the rule $S' \rightarrow \bar{A}\bar{S}$) and, for each nonpreterminal P of G', an end preterminal A_P. For each τ_i, $1 \leqslant i \leqslant m$, let a_i be a new terminal symbol; for each phrase π_j, $1 \leqslant j \leqslant n$, of Π, let b_j also be a new terminal symbol; furthermore, we add one more new terminal symbol c. (Hence, the terminal vocabulary of G' is the union of the terminal vocabulary of $\bar{G}$ and the set $\{a_i, b_j, c;\ 1 \leqslant i \leqslant m,\ 1 \leqslant j \leqslant n\}$.) The rules of G' are:

(1) $S' \rightarrow \bar{A}\bar{S}$;

(2) $P \rightarrow AQ$, if it is a rule of $\bar{G}$;

(3) $P \rightarrow A_P$ for each nonpreterminal P;

(4) $A_{S'} \rightarrow a_i$, $1 \leqslant i \leqslant m$;

(5) $A_{P_j} \rightarrow b_j$, $1 \leqslant j \leqslant n$, where P_j is the root of π_j;

(6) $A \rightarrow x$, if it is a rule of $\bar{G}$ and A is a nonend preterminal of $\bar{G}$;

(7) $\bar{A} \rightarrow c$.

We define a function f from the set of phrases of $\bar{G}$ into the set of phrases of G' as follows. If $\kappa = \tau_i$, $1 \leqslant i \leqslant m$, put $f(\kappa) = S'(A_{S'}(a_i))$. If some π_j is a subphrase of κ, replace π_j by $P_j(b_j)$ in κ and let the tree thus obtained be $f(\kappa)$. Otherwise put $f(\kappa) = \kappa$. The function f is well defined; in fact, if $\kappa = \tau_i$, no π_j is a subphrase of κ, and for any κ at most one π_j is a subphrase of κ.

Using f we next define a function $\bar{f}$ from $\bar{K} = K(\bar{G})$ onto $K' = K(G')$ as follows. We put $\bar{f}(\tau_i) = f(\tau_i)$ for i, $1 \leqslant i \leqslant m$, and $\bar{f}(\tau) = S'(\bar{A}(c), f(\tau))$ if $\tau \neq \tau_i$. The function $\bar{f}$ thus defined is one-to-one and onto K'. For if a sentence τ' of K' is of the form $\tau' = S'(A_{S'}(a_i))$, then $\bar{f}(\tau_i) = \tau'$ and clearly τ_i is the only sentence τ of $\bar{K}$ such that $\tau' = \bar{f}(\tau)$. Otherwise, $\tau' = S'(\bar{A}(c), \tau_1')$, where τ_1' is a phrase of G' whose root is $\bar{S}$. Then, the end preterminal phrase of τ_1' must be of the form $P_j(A_{P_j}(b_j))$. Let τ be a tree obtained from τ_1' by replacing this phrase by π_j; τ is a sentence of $\bar{G}$ and $f(\tau) = \tau'$. It follows that $\bar{f}(\tau) = \tau'$. Clearly τ is the only sentence of K satisfying this relation.

To establish that $\bar{f}$ is structurally homeomorphic, we first prove the following lemma.

LEMMA 3. *Let $\bar{\Pi}$ be a pruning set of $\bar{K}$ that contains $\Pi \cup \{\tau_i; 1 \leqslant i \leqslant m\}$. Put $\Pi' = \{\pi'; \pi' = f(\pi)$ or $\bar{f}(\pi), \pi \in \bar{\Pi}\}$. (That is, $\Pi' = f(\bar{\Pi}) \cup \bar{f}(\bar{\Pi} \cap \bar{K})$.) Then $\bar{f}$ is homeomorphic from $\mathbf{T}_{\bar{\Pi}}(\bar{K})$ onto $\mathbf{T}_{\Pi'}(K')$.*

Let σ and τ be two sentences of $\bar{K}$ and put $\sigma' = \bar{f}(\sigma)$ and $\tau' = \bar{f}(\tau)$. First, we assume $\tau = \tau_i$ for some i, $1 \leqslant i \leqslant m$. Then $(\tau)_{\bar{\Pi}} = \bar{S}$ and $(\tau')_{\Pi'} = S'$. Hence, whatever σ is, $(\tau)_{\bar{\Pi}} < (\sigma)_{\bar{\Pi}}$ and $(\tau')_{\Pi'} < (\sigma')_{\Pi'}$. Next, we assume $\sigma = \tau_i$ for some i. Then $(\sigma)_{\bar{\Pi}} = \bar{S}$. Hence $(\tau)_{\bar{\Pi}} < (\sigma)_{\bar{\Pi}}$ if and only if $(\tau)_{\bar{\Pi}} = \bar{S}$, i.e., $(\tau)_{\bar{\Pi}} < (\sigma)_{\bar{\Pi}}$ if and only if τ is in $\bar{\Pi}$; hence $(\tau)_{\bar{\Pi}} < (\sigma)_{\bar{\Pi}}$ if and only if τ' is in Π', i.e., if and only if $(\tau')_{\Pi'} = S'$. It follows then that $(\tau)_{\Pi} < (\sigma)_{\Pi}$ if and only if $(\tau')_{\Pi'} < (\sigma')_{\Pi'}$, because $(\sigma')_{\Pi'} = S'$. Finally, we assume that $\tau \neq \tau_i$, $\sigma \neq \tau_i$ for any i. Then for some j (and in fact only one j) $\pi_j \subset \pi(\tau, \bar{\Pi})$, and for some k (and for only one k) $\pi_k \subset \pi(\sigma, \bar{\Pi})$. Put $\pi = \pi(\tau, \bar{\Pi})$ and $\kappa = \pi(\sigma, \bar{\Pi})$. From the definition of Π', if $\pi = \tau$, then $\pi(\tau', \Pi') = \bar{f}(\pi)$, otherwise, $\pi(\tau', \Pi') = f(\pi)$. Similarly, if $\kappa = \sigma$, $\pi(\sigma', \Pi') = \bar{f}(\kappa)$; otherwise, $\pi(\sigma', \Pi') = f(\kappa)$. From this it follows that $(\tau)_{\bar{\Pi}} < (\sigma)_{\bar{\Pi}}$ if and only if $(\tau')_{\Pi'} < (\sigma')_{\Pi'}$.

Returning to the proof that $\bar{f}$ is structurally homeomorphic, let Π_1 be an arbitrary pruning set of $\bar{K}$. Take an extension $\bar{\Pi}$ of $\Pi_1 \cup \Pi \cup \{\tau_i; 1 \leqslant i \leqslant m\}$ that makes the identity map of $\bar{K}$ continuous from $\mathbf{T}_{\Pi_1}$ to $\mathbf{T}_{\bar{\Pi}}$. Let Π' be the pruning set of K' defined from $\bar{\Pi}$ as in the lemma. Then f is continuous from $\mathbf{T}_{\bar{\Pi}}$ to $\mathbf{T}_{\Pi'}$, and hence, from $\mathbf{T}_{\Pi}$ to $\mathbf{T}_{\Pi'}$.

Conversely, let Π_1' be an arbitrary pruning set of K'. Let π' be a phrase in Π_1' whose root is S'. Then $f\bar{f}^{-1}(\pi')$ is a phrase of K', which may or may not be in Π_1'. Similarly, if the root of a phrase π' is $\bar{S}$, then $\bar{f}f^{-1}(\pi')$ is a phrase (in fact, a sentence) of K', which may or may not be in Π_1'. We expand Π_1' by adding all such $f\bar{f}^{-1}(\pi')$ and $\bar{f}f^{-1}(\pi')$ and call the new pruning set Π_2'. Π_2' is, so to speak, closed under $f\bar{f}^{-1}$ and $\bar{f}f^{-1}$. Now let Π_3' be

$$\Pi_2' \cup f(\Pi) \cup \bar{f}(\Pi \cap \bar{K}) \cup \{f(\tau_i); 1 \leqslant i \leqslant m\} \cup \{\bar{f}(\tau_i); 1 \leqslant i \leqslant m\}.$$

Π_3' is an extension of Π_1' and is closed under $f\bar{f}^{-1}$ and $\bar{f}f^{-1}$. Finally, let Π'

be the extension of Π_3' that is constructed from Π_1' and Π_3' in the way described in the proof of Part I, Theorem 1 (i.e., π' is in Π' if and only if for some π'' in Π_3' we have $(\pi')_{\Pi_1'} < (\pi'')_{\Pi_3'}$.) The identity map of K' is continuous from $\mathrm{T}_{\Pi_1'}$ to $\mathbf{T}_{\Pi'}$ and Π' is also closed under $f\bar{f}^{-1}$ and $\bar{f}f^{-1}$. Define a pruning set $\overline{\Pi}$ of $\bar{K}$ by $\overline{\Pi} = f^{-1}(\Pi') \cup \bar{f}^{-1}(\Pi')$. Then $\overline{\Pi}$ contains $\Pi \cup \{\tau_i; 1 \leqslant i \leqslant m\}$ and we have $\Pi' = \{\pi'; \pi' = f(\pi) \text{ or } \bar{f}(\pi), \pi \in \overline{\Pi}\}$. Hence, by the preceding lemma, $\bar{f}$ is homeomorphic from $\mathbf{T}_{\overline{\Pi}}$ to $\mathbf{T}_{\Pi'}$, and consequently, $\bar{f}^{-1}$ is continuous from $\mathbf{T}_{\Pi_1'}$ to $\mathbf{T}_{\overline{\Pi}}$. This completes the proof that $\bar{f}$ is a structural homeomorphism from $\bar{K}$ onto K'.

COROLLARY. *For any standard regular grammar G, there exists a principal standard regular grammar G' such that* (i) *$K(G)$ and $K(G')$ are structurally homeomorphic,* (ii) *the noninitial nonfinal states of G' are all self-connected, and* (iii) *G' does not have a separated simple cycle of length more than one.*

In the proof of the theorem, if $\bar{G}$ satisfies condition (iii), so does G'. If, moreover, all the nonfinal states (including the initial state) of $\bar{G}$ are self-connected, then G' also satisfies condition (ii). Hence, from Theorems 4 and 5, our corollary follows.

If, on the other hand, the initial state of $\bar{G}$ is not self-connected, let us redefine G' as follows. We do not introduce the new initial state S'; the initial state $\bar{S}$ of $\bar{G}$ remains the initial state in G'. Rules (1), (4) and (7) are not introduced. Instead we introduce rule

$$A_{\bar{S}} \rightarrow a_i, \qquad 1 \leqslant i \leqslant m. \tag{8}$$

We redefine the function $\bar{f}$ as follows: For each i, $1 \leqslant i \leqslant m$, $\bar{f}(\tau_i) = \bar{S}(A_{\bar{S}}(a_i))$; for $\tau \neq \tau_i$, $1 \leqslant i \leqslant m$, $\bar{f}(\tau) = f(\tau)$, where $f(\tau)$ is as defined in the proof of the theorem. By our assumption that $\bar{S}$ is not self-connected, $\bar{f}$ thus redefined is a one-to-one map from $K(\bar{G})$ onto $K(G')$, and is a structural homeomorphism. Furthermore, if $\bar{G}$ satisfies conditions (ii), then only S' is not self-connected. Then G' satisfies condition (ii) and the corollary follows.

6. *A Finitary Condition for a Weak Structural Homeomorphism*

Let G and G' be two standard regular grammars and assume that a function f from $K = K(G)$ to $K' = K(G')$ is structurally continuous (or structurally $*$-continuous). Then, by definition, for any pruning set Π of G, there exists a pruning set Π' of G' such that f is continuous from $\mathbf{T}_{\Pi}$ to $\mathbf{T}_{\Pi'}$($\mathbf{T}_{\Pi}{}^*$ to $\mathbf{T}^*_{\Pi'}$). We will now consider the problem of whether we can choose specific pruning sets such that the continuity of f with respect to these pruning sets is sufficient to imply the structural continuity (structural *-continuity) of f. Our goal is a theorem stating for two standard regular principal grammars G and G' to be weakly structurally equivalent, it is sufficient that a one-to-one map and its inverse between $K = K(G)$ and $K' = K(G')$ be strongly continuous at the "bottom" of the hierarchies of pruning sets for G and G'. The meaning we intend to give to "bottom" will be specified below.

LEMMA. *Let G and G' be standard regular grammars and let G' be principal. Let f be a one-to-one map from $K = K(G)$ onto $K' = K(G')$. Let Π_0 and Π_0' be the pruning sets of G and G', respectively, consisting of the phrases of the form $P(A(x))$. Put $\mathbf{T}_0 = \mathbf{T}_{\Pi_0}(K)$. Assume that there exist pruning system Π and Π' of G and G' containing Π_0 and Π_0', respectively, such that f is continuous from $\mathbf{T}_0$ to $\mathbf{T}_{\Pi'}$ and f^{-1} is continuous from $\mathbf{T}_{\Pi'}$ to $\mathbf{T}_{\Pi}$. Then:*

(1) *For any natural number k, there exists a natural number k' that satisfies the following condition: For any pair of sentences σ' and τ' of G', if $(\tau')_{\Pi'} < (\sigma')_{\Pi'}$ and $|(\tau)_{\Pi}\backslash(\sigma)_{\Pi}| = k$, then $|(\tau')_{\Pi'}\backslash(\sigma')_{\Pi'}| < k'$, where $\sigma = f^{-1}(\sigma')$ and $\tau = f^{-1}(\tau')$.*

(2) *For any natural number k, there exists a natural number k' that satisfies the following condition: For any pair of sentences σ and τ of G, if $(\tau)_{\Pi_0} < (\sigma)_{\Pi_0}$ and $|(\tau)_{\Pi_0}\backslash(\sigma)_{\Pi_0}| = k$, then $|(\tau')_{\Pi'}\backslash(\sigma')_{\Pi'}| < k'$, where $\sigma' = f(\sigma)$ and $\tau' = f(\tau)$.*

In the proof we write (σ), $V(\sigma)$, etc., instead of $(\sigma)_{\Pi_0}$, $V_{\Pi_0}(\sigma)$, etc.

(1) Put $|(\sigma')_{\Pi'}| - |(\tau')_{\Pi'}| = h$. Let

$$V_{\Pi'}(\sigma_0') = V_{\Pi'}(\tau') \supset V_{\Pi'}(\sigma_1') \supset \cdots \supset V_{\Pi'}(\sigma'_{h_1}) = V_{\Pi'}(\sigma')$$

$$V_{\Pi'}(\sigma_i') \neq V_{\Pi'}(\sigma_{i+1}),\ 0 \leqslant i \leqslant h_1,$$

be the maximal chain of open sets of $\mathbf{T}_{\Pi'}$ between $V_{\Pi'}(\tau')$ and $V_{\Pi'}(\sigma')$; we put $\sigma_0' = \tau'$, $\sigma'_{h_1} = \sigma'$. Put $H' = \mathrm{Max}_{\pi\in\Pi'} |\pi'|$. Since G' is principal, for any i, $0 < i \leqslant h_1$, there is $\bar{\rho}'$ such that $(\bar{\rho}') < (\sigma_i')_{\Pi'}$ and $|(\sigma_i')_{\Pi'}| - |(\bar{\rho}')| = 1$. If $(\bar{\rho}')_{\Pi'} = (\bar{\rho}')$, then $(\sigma'_{i-1})_{\Pi'} = (\bar{\rho}')_{\Pi'}$ and $|(\sigma_i')_{\Pi'}| - |(\sigma'_{i-1})_{\Pi'}| = 1 \leqslant H' + 1$. If $(\bar{\rho}')_{\Pi'} < (\bar{\rho}')$ but $(\bar{\rho}')_{\Pi'} \neq (\bar{\rho}')$, then

$$\begin{aligned}|(\sigma_i')_{\Pi'}| - |(\sigma'_{i-1})_{\Pi'}| &\leqslant |(\sigma_i')_{\Pi'}| - |(\bar{\rho}')_{\Pi'}| \\ &= (|(\sigma_i')_{\Pi'}| - |(\bar{\rho}')|) + (|(\bar{\rho}')| - |(\bar{\rho}')_{\Pi'}|) \leqslant H' + 1.\end{aligned}$$

Hence, in any case $|(\sigma_i')_{\Pi'}| - |(\sigma'_{i-1})_{\Pi'}| \leqslant H_1' = H' + 1$. Thus we have $h \leqslant h_1 H_1'$.

Since f^{-1} is continuous from $\mathbf{T}_{\Pi'}$ to $\mathbf{T}_{\Pi}$, we have

$$V_{\Pi}(\tau) = V_{\Pi}(\sigma_0) \supset V_{\Pi}(\sigma_1) \supset \cdots \supset V_{\Pi}(\sigma_{h_1}) = V_{\Pi}(\sigma)$$

where $\sigma_i = f^{-1}(\sigma_i')$, $1 \leqslant i \leqslant h_1$. Among these open sets there are at most k distinct ones, since $|(\tau)_{\Pi} \backslash (\sigma)_{\Pi}| = k$. Consider now $V(\sigma_0)$, $V(\sigma_1), \ldots, V(\sigma_{h_1})$. There may be more than k distinct sets among these. However, there is a natural number g (depending only on Π) such that at most g different $V(\sigma_i)$'s yield the same $V_{\Pi}(\sigma_i)$'s. (Take as g, for example, the maximum number of different phrases in Π with identical roots.) Hence, among $V(\sigma_0)$, $V(\sigma_1), \ldots$, $V(\sigma_{h_1})$, we have at most gk distinct sets. Since f is continuous from $\mathbf{T}_0$ to $\mathbf{T}_{\Pi'}$, if $V(\sigma_i) = V(\sigma_j)$, we have $V_{\Pi'}(\sigma_i') = V_{\Pi'}(\sigma_j')$, and hence, among $V_{\Pi'}(\sigma_0')$, $V_{\Pi'}(\sigma_1'), \ldots, V_{\Pi'}(\sigma_{h_1}')$, there are at most gk distinct sets.

Consequently, we have $h_1 \leqslant gk$, i.e., $h \leqslant gkH_1'$.

(2) From $(\tau) < (\sigma)$, it follows $(\tau')_{\Pi'} < (\sigma')_{\Pi'}$, as f is continuous from $\mathbf{T}_0$ to $\mathbf{T}_{\Pi'}$. Since f^{-1} is continuous from $\mathbf{T}_{\Pi'}$ to $\mathbf{T}_{\Pi}$, we have $(\tau)_{\Pi} < (\sigma)_{\Pi}$. From the assumption $|(\tau) \backslash (\sigma)| = k$, we have $|(\tau)_{\Pi} \backslash (\sigma)_{\Pi}| \leqslant k + H$, where $H = \mathrm{Max}_{\pi \in \Pi} |\pi|$. From (1), we have some k' such that $|(\tau')_{\Pi'} \backslash (\sigma')_{\Pi'}| < k'$.

THEOREM 7. *Let G and G' be principal standard regular grammars and let f be a one-to-one map from $K = K(G)$ onto $K' = K(G')$. Let Π_0 and Π_0' be the pruning sets of G and G', respectively, consisting of the phrases of the form $P(A(x))$. Assume that there exist pruning sets Π and Π' of G and G' containing Π_0 and Π_0', respectively, such that f is continuous from $\mathbf{T}_0$ to $\mathbf{T}_{\Pi'}$ and f^{-1} is continuous from $\mathbf{T}_{\Pi'}$ to $\mathbf{T}_{\Pi}$ (where $\mathbf{T}_0 = \mathbf{T}_{\Pi_0}(K)$ as in the preceding Lemma). Then f is weakly structurally continuous from K to K'.*

As in the proof of the preceding lemma, we write (σ), $V(\sigma)$, etc., instead of $(\sigma)_{\Pi_0}$, $V_{\Pi_0}(\sigma)$, etc. Let $\overline{\Pi}$ be an arbitrary pruning set of G. Take an arbitrary τ in K. Since G is principal, there exists τ_1 such that $(\tau_1) = (\tau)_{\overline{\Pi}}$. Put $\tau_1' = f(\tau_1)$. Since f is continuous from $\mathbf{T}_0$ to $\mathbf{T}_{\Pi'}$, we have $f(V(\tau_1)) \subset V_{\Pi'}(\tau_1')$.

Take an arbitrary σ in $V_{\overline{\Pi}}(\tau)$ and put $\sigma' = f(\sigma)$. From $(\tau)_{\overline{\Pi}} < (\sigma)_{\overline{\Pi}}$ and $(\tau_1) = (\tau)_{\overline{\Pi}}$, it follows that $(\tau_1) < (\sigma)$, i.e., $\sigma \in V(\tau_1)$, and hence, $\sigma' \in V_{\Pi'}(\tau_1')$, i.e., $(\tau_1')_{\Pi'} < (\sigma')_{\Pi'}$.

Note that $(\tau_1) < (\tau)$ and $|(\tau_1) \backslash (\tau)| \leqslant \overline{H}$, where $\overline{H} = \mathrm{Max}_{\pi \in \overline{\Pi}} |\pi|$. By (2) of the preceding lemma, there exists $\overline{K}$ such that $|(\tau_1')_{\Pi'} \backslash (\tau')_{\Pi'}| < \overline{K}$, and such that $\overline{K}$ is independent of the choice of τ and τ_1. Define a pruning set $\overline{\Pi}'$ of G' as follows: $\overline{\Pi}' = \{\bar{\pi}'; |(\bar{\pi}')_{\Pi'}| \leqslant \overline{K}\}$. Then $(\tau')_{\overline{\Pi}'} < (\tau_1')_{\Pi'}$.

Combining the results of the preceding two paragraphs, we have $(\tau')_{\overline{\Pi}'} < (\tau_1')_{\Pi'} < (\sigma')_{\Pi'}$. Now consider $(\sigma')_{\overline{\Pi}'}$. It may happen that

$(\tau')_{\bar{\Pi}'} \not< (\sigma')_{\bar{\Pi}'}$, but if so, we must have $|(\tau')_{\bar{\Pi}'}\backslash(\sigma')| \leqslant \mathrm{Max}_{\bar{\pi}' \in \Pi'} |\bar{\pi}'|$. Hence, at most, finitely many such σ' can exist. That is, except for at most a finite number of elements, $f(V_{\bar{\Pi}}(\tau))$ is contained in $V_{\bar{\Pi}'}(\tau)$. Consequently, f is continuous from $\mathbf{T}_{\bar{\Pi}}{}^*$ to $\mathbf{T}^*_{\bar{\Pi}'}$.

COROLLARY. *Let G, G', Π_0, and Π_0' be as in the theorem, and assume f is a homeomorphism from $\mathbf{T}_{\Pi_0}$ onto $\mathbf{T}_{\Pi_0'}$. Then f is a weak structural homeomorphism of $K = K(G)$ onto $K' = K(G')$.*

The following three examples show that the preceding corollary is, in a sense, the best we can expect.

EXAMPLE 1. Define regular grammars $\bar{G}$ and G by the 'following rules:

$$\begin{aligned}
\bar{G}:\quad & S \to AS\\
& S \to A\\
& A \to a\\
G:\quad & S \to AP\\
& S \to A\\
& P \to AS\\
& P \to A\\
& A \to a.
\end{aligned}$$

(Strictly speaking, $\bar{G}$ and G are not *standard* as we defined the term in Section 1, since the preterminal A appears on the right-hand side of more than one rule. But we can easily standardize them, and the essential points of this example are not affected. This observation also applies to three examples that follow.)

G and G' both generate the same string language $L = \{a^n; n \geqslant 1\}$ unambiguously. The tree sentences of $\bar{G}$ and G whose terminal string sentences are a^n will be called $\bar{\tau}_n$ and τ_n, respectively. Let $\bar{f}$ be the function from $\bar{K} = K(\bar{G})$ onto $K = K(G)$ defined by $\bar{f}(\bar{\tau}_n) = \tau_n$. Then $\bar{K}$, K, and $\bar{f}$ satisfy the conditions of the corollary to Theorem 7. For, if we let $\bar{\Pi}_0 = \{S(A(a))\}$ and $\Pi_0 = \{P(A(a)), S(A(a))\}$, then $\bar{\tau}_n <_{\bar{\Pi}_0} \bar{\tau}_m$ if and only if $n \leqslant m$, and $\tau_n <_{\Pi_0} \tau_m$ if and only if $n \leqslant m$. Hence, according to the corollary, $\bar{f}$ is a weak structural homeomorphism. But $\bar{f}$ is not a strong structural homeomorphism. To see this, let $\Pi = \{S(A(a)), P(A(a)), S(A(a), P(A(a)))\}$. The paths of G whose origin is S and whose endpoint is not the final state (hence, is either S or P) can be identified by their lengths; so call θ_n the path of length n. (For example, $\theta_0 = (S)$, $\theta_1 = (S, P)$, $\theta_2 = (S, P, S)$.) Then, for each k, $\theta(\tau_{2k})_\Pi = \theta_{2k-2}$, $\theta(\tau_{2k+1})_\Pi = \theta_{2k}$. Hence, $\tau_{2k} <_\Pi \tau_{2k-1}$. Now, let $\bar{\Pi}$ be an arbitrary pruning set of $\bar{G}$. We shall see that f^{-1} cannot be continuous from $\mathbf{T}_\Pi$ to $\mathbf{T}_{\bar{\Pi}}$. Again we can identify the paths of $\bar{G}$ from the initial state to a nonfinal state (i.e., in fact, from S to S) by their lengths; $\bar{\theta}_n$ will denote the path of length n. Let π_0 be the phrase of $\bar{\Pi}$ of maximum height. We may assume that π_0 is a nonpreterminal phrase; otherwise, $\mathbf{T}_{\bar{\Pi}}$ is discrete and f^{-1} cannot be continuous from $\mathbf{T}_\Pi$ to $\mathbf{T}_{\bar{\Pi}}$. Then, π_0 is a sentence of $\bar{G}$ and we may put $\pi_0 = \bar{\tau}_{h+1}$ for some $h \geqslant 0$. Let j be an arbitrary number such that $j \geqslant h$. Then the length of $\theta(\bar{\tau}_j)_{\bar{\Pi}}$ is $j - h$, i.e., $\theta(\bar{\tau}_j)_{\bar{\Pi}} = \bar{\theta}_{j-h}$. Consequently, if we take k such that $2k \geqslant h$, we have $\bar{\tau}_{2k} \not<_{\bar{\Pi}} \bar{\tau}_{2k-1}$. Comparing this with the relation $\tau_{2k} <_\Pi \tau_{2k-1}$ established earlier, we conclude that f^{-1} is not continuous from $\mathbf{T}_\Pi$ to $\mathbf{T}_{\bar{\Pi}}$ at τ_{2k}.

EXAMPLE 2. Let G' be defined by the rules:

$$S \to AS$$
$$S \to B$$
$$A \to aa$$
$$B \to a$$
$$B \to aa$$

Like $\bar{G}$ and G, G' also generates the same string language $L = \{a^n; n \geqslant 1\}$ unambiguously. Call τ_n' the tree sentence of G' whose terminal string sentence is a^n. Define the function f from K onto $K' = K(G')$ by $f(\tau_n) = \tau_n'$, and define the function f' from K' onto $\bar{K}$ by $f'(\tau_n') = \bar{\tau}_n$. f is a strong structural homeomorphism. Indeed, if we had substituted the rule $B \to aac$, with a new terminal letter c, for the rule $B \to aa$ in the definition of G', G' would then be constructed from G as in the corollary of Theorem 2, and f would be the structural homeomorphism from K onto K' constructed in the same corollary. But out G' and the G' redefined in this way are trivially shown to be structurally homeomorphic by indentifying the phrase $B(aa)$ of the one with the phrase $B(aac)$ of the other. (The structural homeomorphism of G and G' by f can, of course, be easily confirmed directly.) Consequently, from Example 1, it follows that f' is a weak structural homeomorphism (but not a strong structural homeomorphism). Put $\Pi_0' = \{S(B(aa)), S(B(a))\}$. Then f' is not continuous from $\mathbf{T}_{\Pi_0'}$ to $\mathbf{T}_{\bar{\Pi}_0}$. For we have $\tau_{2k}' <_{\Pi_0'} \tau_{2k-1}'$ but not $\bar{\tau}_{2k} <_{\bar{\Pi}_0} \bar{\tau}_{2k-1}$.

EXAMPLE 3. Let $\hat{G}$ be defined by the rules:

$$\begin{aligned} &S \to AS \\ &S \to AP \\ &S \to A \\ &P \to BP \\ &P \to B \\ &A \to a \\ &B \to b \end{aligned}$$

$\hat{G}$ generates the string language $\hat{L} = \{a^m b^n, m \geqslant 1, n \geqslant 0\}$ unambiguously. Let $\tau_{m,n}$ be the tree sentence generated by $\hat{G}$ whose terminal string is $a^m b^n$. Put $\hat{\Pi}_0 = \{S(A(a)), P(B(b))\}$. Then, $\tau_{m,n} <_{\hat{\Pi}_0} \tau_{m',n'}$ if and only if either (1) $n = 0$, $m \leqslant m'$, or else (2) $n \neq 0$, $m = m'$, $n \leqslant n'$. Let a function $\hat{f}$ from $\hat{K} = K(\hat{G})$ onto $\hat{K}$ be defined as follows: $\hat{f}(\tau_{m,n}) = \tau_{m,n}$ if $n = 0$ or $n > m$; $\hat{f}(\tau_{m,1}) = \tau_{m,m}$; $\hat{f}(\tau_{m,n}) = \tau_{m,n-1}$ if $1 < n \leqslant m$. Now assume $\tau_{m,n} <_{\hat{\Pi}_0} \tau_{m',n'}$. Then $\hat{f}(\tau_{m,1}) <_{\hat{\Pi}_0} \hat{f}(\tau_{m',m'})$, unless $m = m'$, and $n = 1$. If $m = m'$, $n = 1$, then $\hat{f}(\tau_{m,1}) \nless_{\hat{\Pi}_0} \hat{f}(\tau_{m,n'})$ for $n' = 2, \ldots, m$. It follows that $\hat{f}$ is not automorphic with respect to $\mathbf{T}_{\hat{\Pi}_0}$, but it is automorphic with respect to $\mathbf{T}^*_{\hat{\Pi}_0}$. We now define $\hat{\Pi} = \{S(A(a), P(B(b))), S(A(a)), P(B(b))\}$. For each $m \geqslant 1$, $V_{\hat{\Pi}}(\tau_{m,1}) = \{\tau_{m',n'}; m' \geqslant m\}$. Let $\hat{\Pi}'$ be an arbitrary pruning set of $\hat{G}$, and let h be the maximum of the heights of the phrases in $\hat{\Pi}'$. Then, if $m \geqslant h$, $V_{\hat{\Pi}'}(\hat{f}(\tau_{m,1})) = V_{\hat{\Pi}'}(\tau_{m,n}) \subset \{\tau_{m,n}; n \geqslant 1\}$. It follows that $\hat{f}$ is not continuous from $\mathbf{T}_{\hat{\Pi}}{}^*$ to $\mathbf{T}^*_{\hat{\Pi}'}$ at $\tau_{m,1}$.

We cannot strengthen the conclusion of the corollary of Theorem 7 by replacing "weak structural homeomorphism" by "strong structural homeomorphism" (Example 1), nor can we weaken the premise by replacing "homeomorphism from $\mathbf{T}_{\Pi_0}$ onto $\mathbf{T}_{\Pi_0'}$" by "homeomorphism from $\mathbf{T}^*_{\hat{\Pi}_0}$ onto $\mathbf{T}^*_{\hat{\Pi}_0'}$" (Example 3). Furthermore, the condition stated as the premise of the corollary is not a necessary condition for the conclusion (Example 2).

EXAMPLE 4. Let $\bar{G}$, G and G' be as in Examples 1 and 2. We shall consider the languages obtained from $K = K(G)$ and $K' = K(G')$ by "substituting" $\bar{K} = K(\bar{G})$ for the terminal symbol a. More precisely, let us define $\tilde{G}$ and $\tilde{G}'$ as follows:

$$\tilde{G}: \quad \begin{aligned} &S \to AP \\ &S \to A \\ &P \to AS \\ &P \to A \end{aligned}$$

$$\begin{array}{ll}
 & A \to \bar{S} \\
 & \bar{S} \to \bar{A}\bar{S} \\
 & \bar{S} \to \bar{A} \\
 & \bar{A} \to a \\
\tilde{G}': & S \to AS \\
 & S \to B \\
 & A \to \bar{S}\bar{S} \\
 & B \to \bar{S} \\
 & B \to \bar{S}\bar{S} \\
 & \bar{S} \to \bar{A}\bar{S} \\
 & \bar{S} \to \bar{A} \\
 & \bar{A} \to a.
\end{array}$$

We shall use the notation defined in Examples 1 and 2. Let $\tilde{f}$ be the function from $\tilde{K} = K(\tilde{G})$ onto $\tilde{K}' = K(\tilde{G}')$ which is obtained by expanding by means of the identity map of $\bar{K}$ the function f from K onto K' defined in Example 2. More precisely, if $\tilde{\tau}' = \tilde{f}(\tilde{\tau})$, then $\tilde{\tau}$ and $\tilde{\tau}'$ contains the same number of $\bar{S}$'s and if we match the occurrences of $\bar{S}$ in $\tilde{\tau}$ and $\tilde{\tau}'$ from left to right, these corresponding occurrences of $\bar{S}$ dominate identical trees (i.e., the same sentence of $\bar{K}$). Now we shall see that $\tilde{f}$ is not a structural homeomorphism. Let $\bar{\Pi} = \{P(A(\bar{S}(\bar{A}(a)))), \bar{S}(\bar{A}(a))\}$. We denote by $\tau_n(\bar{\tau}_{i_1}, \bar{\tau}_{i_2}, \ldots, \bar{\tau}_{i_n})$ the sentence of $\tilde{K}$ obtained from the sentence τ_n of K by replacing the n occurrences of a in τ_n by sentences $\bar{\tau}_{i_1}, \bar{\tau}_{i_2}, \ldots, \bar{\tau}_{i_n}$, from left to right. We define similarly $\tau_n'(\bar{\tau}_{i_1}, \bar{\tau}_{i_2}, \ldots, \bar{\tau}_{i_n})$, which denotes a sentence of $\tilde{K}'$. For some k, $k > 0$, consider a sentence $\tilde{\tau}$ of $\tilde{K}$ defined as $\tilde{\tau} = \tau_{2k}(\bar{\tau}_{i_1}, \bar{\tau}_{i_2}, \ldots, \bar{\tau}_{i_{2k}})$, where $\bar{\tau}_{i_1}, \bar{\tau}_{i_2}, \ldots, \bar{\tau}_{i_{2k-1}}$ are arbitrary sentences of $\bar{K}$, and $\bar{\tau}_{i_{2k}} = \bar{S}(\bar{A}(a))$. For any $n \geqslant 2k$, consider $\tau_n(\bar{\tau}_{i_1}, \bar{\tau}_{i_2}, \ldots, \bar{\tau}_{i_{2k}}, \bar{\tau}_{i_{2k+1}}, \ldots, \bar{\tau}_{i_n})$, where $\bar{\tau}_{i_{2k+1}}, \ldots, \bar{\tau}_{i_n}$ are arbitrary sentences of $\bar{K}$. We use $\tilde{\tau}_n$ as the generic name for such a sentence of $\tilde{K}$. Then for any n, $n \geqslant 2k$, $V_{\tilde{\Pi}}(\tilde{\tau})$ contains all the $\tilde{\tau}_n$'s. Now let $\tilde{\Pi}'$ be an arbitrary pruning set of $\tilde{K}'$. For $\tilde{f}$ to be continuous from $\mathbf{T}_{\tilde{\Pi}}$ to $\mathbf{T}_{\tilde{\Pi}'}$ at $\tilde{\tau}$, it is necessary that $V_{\tilde{\Pi}'}(\tilde{f}(\tilde{\tau}))$ contain all the $\tilde{f}(\tilde{\tau}_n)$'s. It follows that the branch $S(B(\bar{\tau}_{i_{2k-1}}, \bar{\tau}_{i_{2k}}))$ must be pruned from $\tilde{f}(\tilde{\tau})$ by $\tilde{\Pi}'$. But since $\tilde{\tau}_{i_{2k-1}}$ is arbitrary, for some $\tilde{\tau}$ such pruning is impossible, as $\tilde{\Pi}'$ is finite. Hence, $\tilde{f}$ cannot be structurally continuous and cannot be a structural homeomorphism.

Assume now that $\tilde{\tau}$ is such that the branch $S(B(\bar{\tau}_{i_{2k-1}}, \bar{\tau}_{i_{2k}})$ is not pruned from $\tilde{f}(\tilde{\tau})$ by $\tilde{\Pi}'$. Then for all n, $n > 2k$, $\tilde{f}(\tilde{\tau}_n)$ are all not in $V_{\tilde{\Pi}'}(\tilde{f}(\tilde{\tau}))$. Furthermore, consider $\tau_n(\bar{\tau}_{j_1}, \bar{\tau}_{j_2}, \ldots, \bar{\tau}_{j_{2k}}, \bar{\tau}_{j_{2k+1}}, \ldots, \bar{\tau}_{j_n})$, where $2k < n$ and $i_1 \leqslant j_1$, $i_2 \leqslant j_2, \ldots, i_{2k} \leqslant j_{2k}$, $1 \leqslant j_{2k+1}, \ldots, 1 \leqslant j_n$; let $\tilde{\sigma}_n$ be the generic name for such a sentence. Then $\tilde{\sigma}_n$ is also in $V_{\tilde{\Pi}}(\tilde{\tau})$, and $\tilde{f}(\tilde{\sigma}_n)$ is not in $V_{\tilde{\Pi}'}(\tilde{f}(\tilde{\tau}))$. If $\tilde{f}$ is continuous at $\tilde{\tau}$ from $\mathbf{T}_{\tilde{\Pi}}{}^*$ to $\mathbf{T}^*_{\tilde{\Pi}'}$, there must be a finite union of the closures of partial sentences that contains all the $\tilde{\sigma}_n$'s. For some $\tilde{\sigma}_n$'s to be contained in the closure of some partial sentence, it is necessary that they share some phrase at some node which is not almost terminal. Thus the minimum of the distances of a terminal node of $\tilde{\sigma}_n$ from the root must be bounded. But the minimum of the distance of a terminal node of $\tilde{\sigma}_n$ from its root is not bounded as $\tilde{\sigma}_n$ varies. From this it follows that no finite union of the closures of partial sentenens can contain all the $\tilde{\sigma}_n$'s. Hence, $\tilde{f}$ is not continuous at $\tilde{\tau}$ from $\mathbf{T}_{\tilde{\Pi}}{}^*$ to $\overset{*}{\mathbf{T}}_{\tilde{\Pi}'}$. Since $\tilde{\Pi}'$ is arbitrary, this shows that $\tilde{f}$ is not structurally *-continuous, and not a weak structural homeomorphism.

We can conclude that in a certain sense the operation of "substitution" on context-free languages preserves neither strong nor weak structural homeomorphism.

RECEIVED: July 1, 1972; REVISED: August 1, 1975

REFERENCES

CHOMSKY, N. (1959), On certain formal properties of grammars, *Inform. Contr.* **2**, 137–167.

HOPCROFT, J. E. AND ULLMAN, J. D. (1969), "Formal Languages and their Relation to Automata," Addison–Wesley, Reading, Mass.

KURODA, S.-Y. (1972), "On Structural Similarity of Phrase-Structure Languages," Proceedings of IRIA Symposium on Theory of Automata, Languages, and Programming, Rocquencourt.

KURODA, S.-Y. (1973), Généralisation de la notion d'équivalence de grammaires, *in* "The Formal Analysis of Natural Languages" (M. Gross *et al.*, Eds.), Mouton, The Hague.