Karel Čulík

On some transformations in context-free grammars and languages

Czechoslovak Mathematical Journal, Vol. 17 (1967), No. 2, 278–311

Persistent URL: http://dml.cz/dmlcz/100776

Czechoslovak Mathematical Journal, 17 (92) 1967, Praha

ON SOME TRANSFORMATIONS IN CONTEXT-FREE GRAMMARS AND LANGUAGES[1])

KAREL ČULÍK, Praha

(Received April 18, 1966)

INTRODUCTION

The only aim of this paper is to describe a class of mathematically well defined mappings which meet probably all the linguistic requirements for (singulary) transformations. The main requirement expressed by N. CHOMSKY is that the transformations are mappings the domains and ranges of which are sets of phrase-markers. There are different mathematical definitions of a context-free grammar and of a phrase-marker determined by it, but the term phrase-marker is very often used by linguists independently of any context-free grammar and without any mathematical definition at all.

To be able to make any use of the main linguistic requirement concerning the transformations it is necessary to introduce a new definition of a phrase-marker (or of a structural description) which is independent of any context-free grammar and which is acceptable to linguists (i.e. which can be justified empirically). Both these conditions can easily be satisfied by a definition which is on the one part equivalent to that which depends on the notion of context-free grammar (as it was introduced originally by N. Chomsky, e.g. [1, 2]) but which does not make any use of the term rule (only the terms terminal and nonterminal symbols or vocabularies are necessary), i.e. a definition using the terms of graph theory only (see [3, 4]). Using this mathematical definition of the phrase-marker the reasoning leading to the considered class of mappings seems to be very natural. A mathematical definition of a transformation follows almost immediately from simple purely mathematical requirements.

Concerning the linguistic requirements for transformations the following explanation of the situation can be of some use.

[1]) The main part of this paper was lectured at the summer seminar (June, July, August 1965) on mathematical linguistics in M.I.T. under the leadership of N. Chomsky. During this seminar there were fruitful discussions with many other participants, especially with E. Bach, M. Halle, J. J. Katz et all.

In a traditional grammar of a natural language the one-to-one correspondences (in fact, there are some exceptions) between the sets of the active sentences and of the passive ones or between the declarative sentences and the questions etc., are very well known and very often used in teaching that language. Correspondences of that type were called transformations by Z. HARRIS [5]. In teaching e.g. English, the active sentences are considered simpler than the corresponding passive ones and therefore they are taught sooner. Only after a grammar for active sentences is built are the passive sentences introduced or described as a result of the active-passive transformation which is applied to the active sentences. Finally a grammar for the passive sentences is also built or the original grammar is extended to that case.

It is well known that in the traditional grammar the set L of all English sentences can be divided into special subsets $L_1, L_2, \ldots, L_n$ (e.g. L_1 is the set of active sentences, L_2 of passive ones, L_3 of declarative ones, L_4 of questions etc.) such that some one-to-one correspondences $C_{i,j}$ between L_i and L_j can be discovered as an important result of linguistic research (e.g. $C_{1,2}$ is the active-passive correspondence and $C_{3,4}$ the declarative-question one etc.). To avoid any misunderstanding, the correspondences $C_{i,j}$ will be called *descriptive transformations*.

Now from a pure mathematical point of view if we have the subsets $L_1, L_2 \ldots, L_n$ and the descriptive transformations $C_{1,2}, C_{3,4}, \ldots,$ it is possible to define new subsets $L'_1, L'_2, \ldots, L'_m$ by the condition that L'_i is non void and is an intersection of some of the original subsets $L_1, L_2, \ldots, L_n$ such that always either $L'_i \cap L_j = \emptyset$ or $L'_i \cap L_j = L'_i$ (e.g. L'_1 contains active declarative sentences, L'_2 active questions, L'_3 passive declarative sentences, L'_4 passive questions etc.). In general the subsets L'_i contain less elements than the subsets L_j but, on the other side the number m of the subsets L'_i is greater than the number n of L_j. Furthermore the partial transformations $C'_{i,j}$ of $C_{i,}$ are also determined between L'_i and L'_j such that $L'_i \cap L_i \neq \emptyset \neq L'_j \cap L_j$ (e.g. $C'_{1,2}$ is the transformation which assigns the corresponding question to the given declarative sentences but only if both are active sentences, $C'_{1,3}$ is the active-passive transformation but only for the declarative sentences, etc.). Finally, it is always possible to compose the transformations $C'_{i,j}$ and $C'_{j,k}$. The composed function $C'_{j,k}C'_{i,j}$ is a new transformation between L'_i and L'_k (e.g. $C'_{3,4}C'_{1,3}$ assigns the corresponding passive question belonging to L'_4 to each active declarative sentence from L'_1, because $C'_{1,3}$ assigns the passive form from L'_3 to the given active declarative sentence from L'_1 and then $C'_{3,4}$ assigns the corresponding question from L'_4).

In this situation one can look for a single distinguished subset L'_i or (perhaps construct a new subset L'_0 containing some artificial sentences) and some distinguished transformations such that, using them and using the composition of them, all the subsets $L'_1, L'_2, \ldots, L'_m$ can be generated. And this fact was stressed by N. Chomsky and used in his generative transformational grammars.

On the other hand, it was expressly required by N. Chomsky that the transformations in his sense – they will be called *structural transformations* here – concern the whole phrase-markers (or structural descriptions) of the sentences and not only the sentences themselves as it is the case when using descriptive transformations. Of course the notion of a structural transformation is more powerful than that of a descriptive one and each structural transformation defined for phrase-markers induces the corresponding descriptive one defined only for the sentence which is the terminal string of the considered phrase-marker.

In the transformational grammar of English N. Chomsky assumes that a set of phrase-markers $\mathfrak{P}_0$ of the "kernel sentences" (or later of the "basic strings") L'_0 is given and also some structural transformations $T'_1, T'_2, \ldots, T'_m$ are given together with certain rules of composition of them such that using the allowed composed transformations $T'_{i_1} T'_{i_2} \ldots T'_{i_k}$ to the elements of $\mathfrak{P}_0$, all the phrase-markers of all English sentences can be generated. If we denote by C'_i the descriptive transformation corresponding to the structural transformation T'_i and if we use the same rules of composition of C'_i as of T'_i then by some composed transformation $C'_{i_1} C'_{i_2} \ldots C'_{i_k}$ applied to L'_0 we have to obtain L'_1, by some different composed transformation applied to L'_0 we obtain L'_2, etc. It is clear that the original subsets $L_1, L_2, \ldots, L_n$ are also the unions of some L'_i and the original transformations $C_{i,j}$ are also determined by some $C'_{i,j}$, so we are back in the starting situation in traditional grammar.

By that the connection between the transformations in the Harris's sense and in the Chomsky's sense is clarified, of course only in the simplest case concerning so called *singulary transformations* (i.e. to a single sentence or to a single phrase-marker again a single sentence or single phrase-marker should always be assigned; the possibility of constructing from two or more sentences or phrase-markers only a single one is not considered here).

Finally there is an assumption about $\mathfrak{P}_0$ expressed by N. Chomsky, namely that $\mathfrak{P}_0$ is determined by a context-free grammar G_0, but not similar assumption is made with regard to the sets of derived phrase-markers $\mathfrak{P}'_1, \mathfrak{P}'_2, \ldots, \mathfrak{P}'_m$ of the sentences belonging to $L'_1, L'_2, \ldots, L'_m$ resp. Let e.g. $\mathfrak{P}'_1 = T'_{0,1}(\mathfrak{P}'_0)$ where $T'_{0,1} = T'_{i_1} T'_{i_2}, \ldots, T'_{i_k}$ is a certain allowed composed transformation. As $\mathfrak{P}'_0$ is a finite set, $\mathfrak{P}'_1$ is also finite and therefore we can assume that $\mathfrak{P}'_1$ is determined by a context-free grammar G_1. Under these assumptions the transformation $T'_{0,1}$ can be considered as a correspondence between two context-free grammars G_0 and G_1. And that is just the starting point for a general definition of a (singular) transformation introduced in this paper.

1. PHRASE-MARKERS AND MARKERS

Let us introduce – only in order to clarify the terms used here – the following conventions: the term *phrase-marker* is always dependent on a context-free grammar as it is stressed in its informal definition given by N. Chomsky [1] where the starting point is a derivation of the considered grammar and where the familiar diagram having the form of a labeled tree is drawn step by step for all the used rules; the term *structural description* is always independent of any context-free grammar but it corresponds again to the diagrams having the form of a labeled tree (e.g. the diagrams of the deep structure of "basic strings") where – probably – the edges express the immediate constituent relationship.

The confusing fact is that the same picture, the diagram of a labelled tree can be considered once as a phrase-marker and at some other time as a structural description. If we wanted to distinguish these two interpretations of the same labelled tree we could do it in the following way: in a phrase-marker we should mark out all the particular rules used in it by a closed dotted line around the corresponding vertices as shown in Fig. 1b), where evidently the rules used are as follows:

1. ⟨VP⟩ ::= ⟨Verb⟩ ⟨NP⟩
2. ⟨Verb⟩ ::= ⟨V⟩ ⟨Prt⟩
3. ⟨V⟩ ::= turn,
4. ⟨Prt⟩ ::= out,
5. ⟨NP⟩ ::= ⟨Determ⟩ ⟨N⟩,
6. ⟨Determ⟩ ::= ⟨Quant⟩ ⟨Art⟩,
7. ⟨Quant⟩ ::= some of,
8. ⟨Art⟩ ::= the,
9. ⟨N⟩ ::= lights.

Thus a) is a structural description but b) is a phrase-marker.

Unfortunately there is an exact mathematical definition neither for the notion of the phrase-marker nor for the structural description and therefore nothing that now follows can be proved, it can be only justified empirically by linguists or resulted by an example.

It is clear how to pass on from a phrase-marker to a "corresponding" structural description and also how to mark out certain "rules" in a structural description such that it becomes a "corresponding" phrase-marker. What we are claiming is an assumption – probably acceptable to all linguists – that the classes of all structural descriptions and of all phrase-markers are identical, i.e. it is of no importance to distinguish between the notions of a phrase-marker and a structural decription.

To be able to make any use of this assumption at a mathematical level it is necessary to have an exact mathematical notion underlying the two notions of phrase-marker and structural description and also being acceptable to all the linguists, i.e. each linguistically correct phrase-marker or structural description always satisfies the underlying definition and conversely each diagram satisfying the underlying definition is a phrase-marker or a corresponding structural description.

If we look at the diagrams in Fig. 1 it is clear that very important information is contained in the placing of the particular symbols on the paper. One part of this information is expressed by the drawn lines and by placing the relation "up – down" and the second part is not expressed by any lines at all but only by the relation "left-right". The first part of the information concerns a binary relation σ which is the *immediate constitutent relationship* and the second part concerns another binary relation ϱ which is the *immediate ordering relationship* (a part of it concerning the terminal symbols is the immediate word-order relationship and the remaining part concerns the order of non-terminal symbols which in the strings belong to the certain derivations).

Fig. 2a) shows an underlying structural description of a basic string and b) is the corresponding diagram where also the σ-lines but dotted are drawn and both the ϱ-

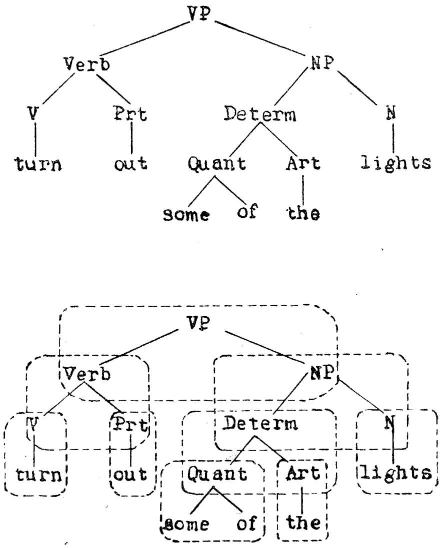

Fig. 1a, b.

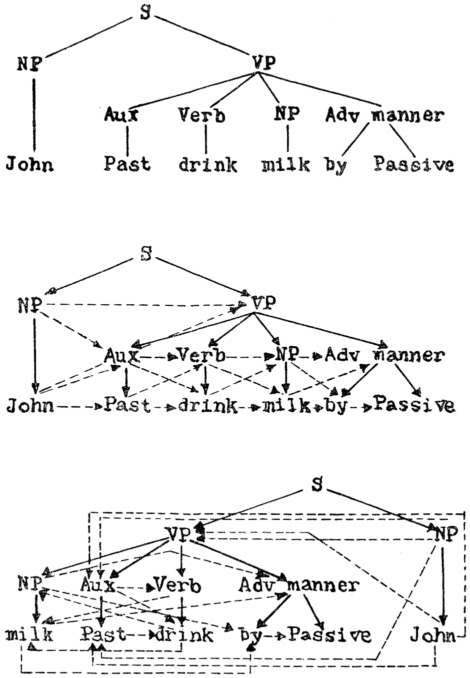

Fig. 2a, b, c.

and σ-lines are directed by making the use of the directions "up – down" and "left – right" respectively. The most important difference between a) and b) is that in b) nothing depends on the placing of the particular symbols on the paper because all the necessary information is contained in both types of arrows (that type of a description is necessary for any rigorous handling, e.g. for a machine handling). Thus b) and c) express the same phrase-marker or the same structural description as a), but evidently both are rather complicated.

There is another important fact in the structural description a) in Fig. 2, namely that the non-terminal symbol NP is used in two different places. This is the reason why it is necessary to distinguish the tree or generally the graph structure of a structural description on the one hand and the labelling of the nodes of it by some symbols on the other hand. Then a labelling of nodes is a function defined on the set of nodes of the underlying graph. The values of this function are the terminal or non-terminal symbols.

Now the required definition underlying all the linguistic structural descriptions or phrase-markers can be formulated as follows: a structural description (or a phrase-marker) over the terminal or non-terminal vocabulary V_T or V_N resp. (and over the context-free rules in $\mathfrak{R}$, i.e. over a contex-free grammar $G = \langle V_T, V_N, \mathfrak{R} \rangle$) is a finite *double graph with labelled vertices* $P = \langle R, \varrho, \sigma, f \rangle$, where R is a finite set of vertices, $\varrho \subset R \times R$ and $\sigma \subset R \times R$ are the binary relations the elements of which are called *edges* and $f \subset R \times V$, $V = V_T \cup V_N$ is a function called *labelling*, such that the following conditions are satisfied

(a) ϱ is a *rooted tree relation*, i.e. there is a distinguished vertex $r \in R$ called the *root* such that for each $x \in R$, $x \neq r$ there is exactly one *path* in ϱ from r to x (i.e. a sequence $(v_0, v_1, \ldots, v_n)$ of vertices from R such that $(v_{i-1}, v_i) \in \varrho$ for each $i = 1, 2, \ldots, n$ and $v_0 = r$, $v_n = x$);

(b) σ is an arbitrary *atransitive* and *acyclic relation*, i.e. if $(v_0, v_1, \ldots, v_n)$ is a path in σ then $(v_0, v_n) \notin \sigma$ for each $n > 1$ and $(v_n, v_0) \notin \sigma$ for each $n \geqq 1$, which is *connected*; the two relations ϱ and σ are connected by the following conditions

(c) if $x, y \in R$ and $x \neq y$ then there always occurs precisely one of the following possibilities: there is a path either in ϱ from x to y or in σ from x to y or in ϱ from y to x or in σ from y to x;

(d) if $(x, y) \in \sigma$ and $(x, x)'$, $(y, y)' \in \varrho$ then (x', y'), (x, y'), $(x', y) \in T\sigma$ (where $T\sigma$ is the *transitive closure* of σ); and the labelling f satisfies

(e) if $x \in R$ and $f(x) \in V_T$ then x is an *end vertex*, i.e. there is no vertex y such that $(x, y) \in \varrho$ ($f(r)$ is a distinguished symbol – usually $S \in V_N$); and if we are considering a phrase-marker over $G = \langle V_T, V_N \mathfrak{R} \rangle$ then

(f) if $x \in R$ and $Q = \{y \in R; (x, y) \in \varrho\} \neq \emptyset$ and $Q = \{y_1, y_2, \ldots, y_k\}$ such that $(y_i, y_{i+1}) \in \sigma$ for each $i = 1, 2, \ldots, k - 1$ then $(f(x) :: = f(y_1) f(y_2) \ldots f(y_k)) \in \mathfrak{R}$.

This definition was introduced in [3, 4] on the base of an analysis of phrase-markers or structural descriptions, only the conditions were formulated using the transitive closures $T\varrho$ and $T\sigma$ of both relations (namely then $T\varrho$ and $T\sigma$ are special *partially ordering relations* and $T\varrho \cup T\sigma$ is a *full ordering relation*).

It is clear that it is very difficult to work with such a complicated formulation. Therefore in [6] another definition of phrase-marker was introduced (of course equivalent to the previous one in the sense that there is a one-to-one correspondence between the formations in the first and in the second sense) which is dependent on the rules of a context-free grammar, i.e. the condition (f) is used in full. To avoid any possible confusions we shall speak only about the markers over the given context-free grammar instead about the phrase-markers.

A *proper marker* over a context-free grammar $G = \langle V_T, V_N, \mathfrak{R} \rangle$ is a rooted tree with labelled vertices and edges $M = \langle A, r, B, \varphi, \psi \rangle$; A is a finite set of vertices, $r \in A$ is a root, $B \subset A \times A$ is the set of edges, $\varphi \subset A \times \mathfrak{R}$ is a function called labelling of vertices and $\psi \subset B \times \{1, 2, 3, \ldots\}$ is a function called labelling of edges, such that the following conditions are satisfied:

(A) $\langle A, r, B \rangle$ is a rooted tree with the root r, i.e. r is a distinguished vertex and for each vertex $a \in A$, $a \neq r$ there is exactly one path in B from r to a;

(B) if $(a, b) \in B$ and $\varphi(a) = (X_0 ::= X_1 X_2 \ldots X_m) \in \mathfrak{R}$, $\varphi(b) = (Y_0 ::= Y_1 Y_2 \ldots \ldots Y_n) \in \mathfrak{R}$ where $X_i \in V$ for $0 \leqq i \leqq m$, $Y_j \in V$ for $0 \leqq j \leqq n$ then there exists an integer k such that $1 \leqq k \leqq m$, $X_k = Y_0$ and $\psi(a, b) = k$.

(C) is $b \neq c$ then $\psi(a, b) \neq \psi(a, c)$.

An *improper marker* is each particular symbol from V. An improper marker has no vertices. Finally a *disconnected marker* is any finite sequence of the length $l > 1$ of proper or improper markers over G.

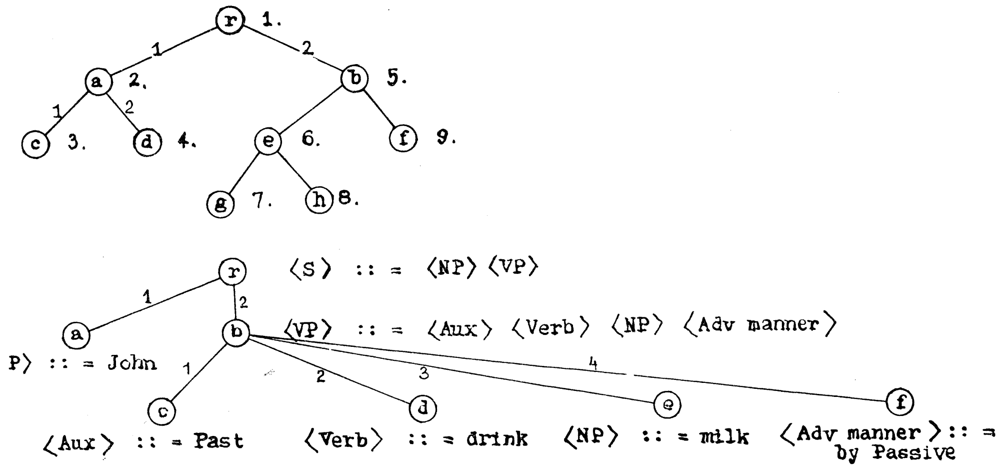

Fig. 3a, b.

In Fig. 3a) is a marker corresponding to the phrase-marker in Fig. 1b) and in Fig. 3b) is a marker which corresponds to Fig. 2. In the first case the labelling of vertices uses only the numbers of the particular rules instead of the rules themselves.

Again it is very easy to see how to pass from a phrase-marker to the corresponding marker (of course the underlying rooted tree is determined only in regard to the usual graph-theoretical isomorphism) and conversely from the markers to the phrase-markers.

What is important to stress here is that – briefly speaking – to the two non-isomorphic phrase-markers two isomorphic markers can correspond, because the isomorphism concerning the underlying rooted tree by markers is weaker than the isomorphism concerning the underlying double graphs by phrase-markers. In fact the ordering relation σ is expressed in a marker only by the labelling of edges ψ.

2. LINGUISTIC REQUIREMENTS CONCERNING TRANSFORMATIONAL GRAMMARS

The first requirement expressed bp N. Chomsky is that each transformation T is or is to be a mapping the domain and the range of which are the sets $\mathfrak{P}$ and $\mathfrak{P}^*$ of phrase-markers resp. In fact the range $\mathfrak{P}^*$ is not given in advance or explicitly but it has to be determined uniquely by the transformation T and by its domain $\mathfrak{P}$, i. e. $\mathfrak{P}^* = \{P^*;$ there is $P \in \mathfrak{P}$ such that $T(P) = P^*\}$. That means that a transformation has to be a constructive mapping containing always a complete procedure for obtaining P from P.

The second requirement expressed by N. Chomsky concerns the determination of the domain $\mathfrak{P}$ of T. The existence of a broad and unspecified class of phrase-markers (e. g. the cless ok all possible phrase markers over the all possible context-free grammars) is assumed and $\mathfrak{P}$ is a subset of it which is fully characterised by the structure index of T. Namely, if a is an arbitrary string over $V = V_T \cup V_N$, where obviously in V_T and V_N there are symbols used in the phrase-markers from $\mathfrak{P}$, then there is a uniquely determined set $\mathfrak{P}_a$ of phrase-markers such that a is their structure index. If we use the notion of th phrase-marker introduced in the previous section then $a - A_1A_2 \ldots A_n$, where $A \in V$ for $1 \leqq i \leqq n$ is a structural index of a phrase-marker $P = \langle R, \varrho, \sigma, f\rangle$ if there are vertices $r_1, r_2, \ldots, r_n \in R$ such that $r_i, r_{i+1}) \in \sigma$ for $i = 1, 2, \ldots, n-1$, i.e. $(r_1, r_2, \ldots, r_n)$ is a path in σ aue it is not possible to extend it, ane if $f_|^i(r_i) = A_i$ for each $i = 1, 2, \ldots, n$.

The third requirement expressed by N. Chomsky concerns the so-called elementary transformation T_{el1} which underlies T. If $t = L(P)$, where $L(P)$ denotes the *last string* of P, i.e. concatenation of all symbols by which the end vertices are labeled in the order determined by σ, then by the structural index $a = A_1A_2 \ldots A_n$ of P a sequence of strings $(t_1, t_2, \ldots, t_n)$ is uniquely determined such that $t_1t_2 \ldots t_n = t$ and that t_i is traceable in P to the vertex labeled by A_i for each $i = 1, 2, \ldots, n$, i.e. using the previous notation we can say that each vertex r_i determines one sub-phrase-marker P_i of P the root of which is r_i (in all details we can put $R_i = \{x \in R;$ either $x = r_i$ or $x \neq r_i$ but there is a path from r_i to x in $\varrho\}$, $\varrho_i = \varrho \cap (R_i \times R_i)$, $\sigma_i = \sigma \cap (R_i \times R_i)$, $f_i = f/R_i)$ and $t_i = L(P_i)$ for each $i = 1, 2, \ldots, n$. Now the requirement is that $L(P^*) = \sigma_1\sigma_2 \ldots \sigma_n$ where $\sigma_i = T_{el}\ (i;\ t_1, t_2, \ldots, t_n)$ for each $i = 1, 2, \ldots, n$. Besides that if $P' \in \mathfrak{P}$, $L(P') = t'$ and P'_i with the root r'_i are the corresponding sub-phrase-markers such that $L(P'_i) = t'_i$ and $f'(r'_i) = A_i$ for each $i = 1, 2, \ldots, n$, then each $\sigma'_i = T_{el}(i;\ t'_i, t'_2, \ldots, t'_n)$ is formed from σ_i by replacing t_j by t'_j for each $j = 1, 2, \ldots, n$. That means that σ_i has to be a string having the following form: $\sigma_i = x_0^{(i)}t_1x_1^{(i)}t_2 \ldots t_nx_n^{(i)}$, where $x_j^{(i)}$ is a terminal string over V_T for each i, $1 \leqq i \leqq n$ and each j, $0 \leqq j \leqq n$ (otherwise it would not be possible to substitute t'_j in σ_i for each $j = 1, 2, \ldots, n$). But obviously it should not be required $L(P^*) = x_0^{(1)}t_1x_1^{(1)}t_2 \ldots t_nx_n^{(1)}x_0^{(2)}t_1x_1^{(2)}t_2 \ldots t_nx_n^{(2)} \ldots x_0^{(n)}t_1 \ldots t_nx_n^{(n)}$ and therefore the requirement concerning T_{el} has to be reformulated in accordance with the effect which was to be realized, namely the effect of ruling out the possibility of applying transformations to particular strings of actually occuring words (or morphems) or in other words the effect of avoiding an arbitrary pairing off of sentences. That means that here the complete characterisation which mappings are transformations is required.

The fourth requirement expressed by N. Chomsky concerns the resulting effect that T has on the terms $t_1, t_2, \ldots, t_n$. N. Chomsky requires that, for instance, T may have effect of deleting or permuting certain terms, of substituting one for another, or adding a constant string in a fixed place and so on.

The fifth requirement expressed by N. Chomsky is picking out the importance of the relation "t_i is an A_i" in the phrase-marker P, i.e. the relation "t_i is traceable to A_i" or in other words "there is a sub-phrase-marker P_i of P such that $L(P_i) = t_i$ and $F(P_i) = A_i$", where $F(P)$ is the *first string* of P. By that he only reminds us that the graph-structure of P has to be changed very carefully in order to obtain the new graph-structure of P^*.

All five requirements concern only the so called singular transformations (such that apply to a single phrase-marker and not to pairs or generally to n-tuples of phrase-markers). If we review all of them we find out that there is no description at all of any procedure how to get P^* from P.

Using some examples how the singular transformations are determined by the linguist working in the area of transformational grammars the following scheme can be deduced: if T is a transformation and $a = A_1A_2 \ldots A_n$ its structural index then the following transformational rule is used $A_1A_2 \ldots A_n \Rightarrow a_0A_1^* a_1 A_2^* \ldots A_n^* a_n$ where $(i_1, i_2, \ldots, i_n)$ is a permutation of $\{1, 2, \ldots, n\}$, a_j are strings over V_T and either $A_j^* = A_{i_j}$ or $A_j^* = I$ (I is the null-string in the free semigroup over V).

If we are thinking about the phrase-marker P in Fig. 1 the transformational rule can be ⟨V⟩ ⟨Prt⟩ ⟨NP⟩ ⇒ ⟨V⟩ ⟨NP⟩ ⟨Prt⟩, i.e. using the usual notation $A_1 =$ = ⟨V⟩, A_2 = ⟨Prt⟩, A_3 = ⟨NP⟩, i.e. $n = 3$ and $i_1 = 1$, $i_2 = 3$, $i_3 = 2$, $A_j^* = A_{i_j}$ for $j = 1, 2, 3$ and $a_j = I$ for $j = 0, 1, 2, 3$. Then the $T(P) = P^*$ is to be the phrase-marker drawn in Fig. 4a).

Similarly if P is phrase-marker in Fig. 2 the transformational rule can be as follows: $n = 6$, A_1 = ⟨NP⟩, A_2 = ⟨Aux⟩, A_3 = ⟨Verb⟩, A_4 = ⟨NP⟩, A_5 = by, $A_6 =$ = Passive and $i_1 = 4$, $i_2 = 2$, $i_3 = 6$, $i_4 = 3$, $i_5 = 5$, $i_6 = 1$, $A_j^* = A_{i_j}$ for $j =$ $= 1, 2, \ldots, 6$ and a_2 = be, $a_j = I$ for $j \neq 2$, $j = 0, 1, \ldots, 6$. Then the $T(P) = P^*$ is to be a phrase-marker in Fig. 4b).

In the first case in P^* it is possible to find the corresponding structural index ⟨V⟩ ⟨NP⟩ ⟨Prt⟩ and in the second case ⟨NP⟩ ⟨Aux⟩ be Passive ⟨Verb⟩ by ⟨NP⟩, both are in fact the right sides of the used transformational rules. In the first case – because this transformation should be a permutation only – the subphrase-markers of P^* below the structural index are well defined but not above it and in the second case neither below nor above the structural index of P^* there is determined what subphrase-markers should be – because of adding "be".

As the last requirement it can be stressed that the transformations are to be positively one-to-one mappings. Thus to a transformation T there allways exists its inverse transformation T^{-1}. This is an important requirement in regard to the semantical questions, namely that by using a transformation no information can be lost.

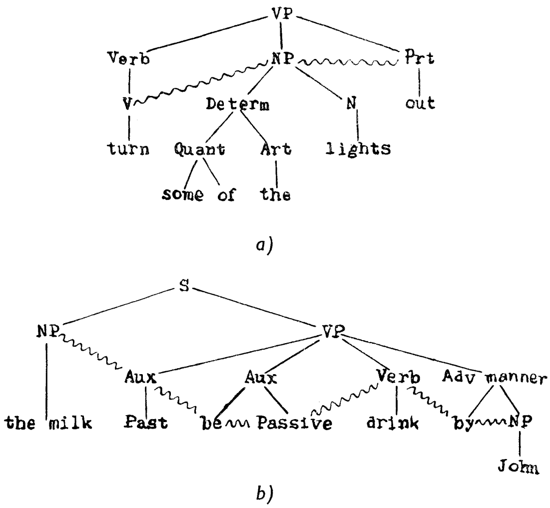

Fig. 4a, b.

3. LINGUISTIC REQUIREMENTS CONCERNING TRADITIONAL TRANSFORMATIONS

In regard to a traditional grammar of a natural language we can suppose that all the sentences of that language contained in the set L are strings over the terminal vocabulary V_T, where V_T contains all the basic forms of all word-forms, all suffixes and also a special symbol denoting space (we do not go into details and consider e.g. the irregular verbs, where in the word-forms "goes" and "went" there is no common root or stem, etc.). Further we can suppose that V_T is divided into subsets according to the traditional *grammatical categories* as substantives, adjectives, verbs, numerals on the one hand and into the other grammatical or logical categories as pronouns, articles, auxiliary verbs, prepositions, conjunctions and all other phrases having pure logical character as negation, "if ... then", "there exists", "for each" etc. and of course what remains, i.e. different suffixes, on the other hand. The first type of categories can be called *proper* and a set of all terminal symbols belonging to any proper category can be called *proper (terminal) vocabulary* and denoted by V_P. The second type of categories can be called *auxiliary* and $V_A = V_T - V_P$ can be called *auxiliary (terminal) vocabulary*. Obviously the words having full lexical meaning belong to V_P while the words of a grammatical or logical character belong to V_A

(again here we are not going into all details about the meaning because we do not consider a determination of place and time and different possible modalities etc.).

Besides that we can generally suppose that we have more specialized proper categories than the above mentioned ones, e.g. we have the categories for intransitive verbs, transitive and double transitive ones etc. and let $C_1, C_2, \ldots, C_p$ be the subsets of V_P corresponding to the particular proper categories $\langle C_1 \rangle, \langle C_2 \rangle, \ldots, \langle C_p \rangle$ (further it is possible to assign to each category $\langle C_i \rangle$ certain *logical characteristics* expressing whether the words belonging to C_i express some individual objects or classes of them, or properties of them, or relations between them, or properties of properties etc. which in other words says what type of a logical predicate the words are or can be). Now we can define the variables ξ_i for each $i = 1, 2, \ldots, p$ as the new symbols denoting arbitrary elements of C_i. In fact it would not be necessary to introduce these new symbols ξ_i because we could use the names $\langle C_i \rangle$ to be same purposes as it is done in the ALGOL 60 language description [7].

Now in regard to the division of V_T into V_P and V_A each string q from L (or each string over the vocabulary V_T) can be uniquely expressed in a *canonical form* $q = q_0 Q_1 q_1 \ldots Q_k q_k$ where $Q_i \in C_{j_i} \subset V_P$ for each $i = 1, 2, \ldots, k$ and q_i are strings (may be null-string I) over the auxiliary vocabulary V_A for each $i = 0, 1, \ldots, k$. Here k can be called *canonical length* of q. Further the string $q_0 \xi_{j_1} q_1 \ldots \xi_{j_k} q_k$, where ξ_{j_i} is a variable corresponding to the category $\langle C_{j_i} \rangle$ determined by the condition $Q_i \in C_{j_i}$ for each $i = 1, 2, \ldots, k$, can be called a *sentence-form of* q and the strings q_i or symbols ξ_{j_i} are called the *constants* or *variables of that form* resp. It is clear that such a sentence-form $q_0 \xi_{j1} q_1 \ldots \xi_{j_k} q_k$ is uniquely determined by a sequence of constants $(q_0, q_1, \ldots, q_k)$ and by a sequence of categories $(\langle C_{j_1} \rangle, \langle C_{j_2} \rangle, \ldots, \langle C_{j_k} \rangle)$ or only by their indices $(j_1, j_2, \ldots, j_k)$ if a fixed ordering of them is assumed.

If we are thinking about the subsets $L_1, L_2, \ldots, L_n$ of the considered set L such that any the traditional transformation $T_{i,j}$ between L_i and L_j may be obtained, then the following requirement concerning the subsets L_i has to be satisfied: if $q \in L_i$ and $q = q_0 Q_1 q_1 \ldots Q_k q_k$ is a canonical expression then also $q_0 Q_1^* q_1 Q_2^* \ldots Q_k^* q_k \in L_i$ for each Q_h^* such that Q_h^* and Q_h belong to the same category C_{j_h} for each $h = 1, 2, \ldots, k$. In other words if F is the set of all sentence-forms determined by the set L of sentences and if in the similar way the sentence-forms F_i correspond to the subsets L_i for $i = 1, 2, \ldots, n$ then the previous requirement can be expressed as a condition $F_i \cap F_j = \emptyset$ for $i \neq j$ and $i, j = 1, 2, \ldots, n$.

The further requirement can be formulated as a mathematical definition of a traditional transformation (only a singular one!) $T_{i,j}$ between L_i and L_j in the following way: $T_{i,j}$ is a one-to-one mapping from L_i onto L_j such that if $q = q_0 Q_1 q_1 \ldots Q_k q_k \in L_i$ then $T_{i,j}(q) = q_0^* Q_1^* q_1^* \ldots Q_l^* q_l$ and a) $l = k$, b) there is a mapping $f_{i,j}$ which assigns a sequence $(s_0, s_1, \ldots, s_h) \in F_j$ to an arbitrary sequence $(r_0, r_1, \ldots, r_h) \in F_i$ of the strings (may be also null-strings) over V_A for each $k = 0, 1, 2, \ldots$ such that $f_{i,j}(q_0, q_1, \ldots, q_k) = (q_0^*, q_1^*, \ldots, q_k^*)$, c) to each sequence $(q_0, q_1, \ldots, q_k)$ there corresponds a permutation π_q of the integers $\{1, 2, \ldots, k\}$ such that $Q_h^* = Q_{\pi_q^{-1}(h)}$ for each $h = 1, 2, \ldots, k$.

In a very special case it can happen that every sentence belonging to L_i has a fixed canonical length k. Then $f_{i,j}$ corresponds to the Chomsky's elementary transformation and it is always possible to consider instead of $f_{i,j}$ a sequence of mappings $f^{(0)}, f^{(1)}, \ldots, \ldots, f^{(k)}$ such that $f^{(h)}(q_0, q_1, \ldots, q_k) = q_h^*$ for each $h = 0, 1, 2, \ldots, k$. However, one cannot make any use of that because it does not touch the important question how to determine $f_{i,j}$ in another manner than by an enumeration, which is of course a non-effective method in a general case.

The mapping $f_{i,j}$ is the most important part of a transformation because – in fact – the transformations are essentially the mappings rather of sentence-forms than of the sentences themselves.

But also in a general case it is here quite clear what is a constant and what a variable (i.e. a category or the corresponding non-terminal symbol called the metalinguistic variable in ALGOL 60 language) and also all the possible effects of the transformation are determined exactly. Namely the proper symbols (or the variables if we are thinking at the level of sentence-forms) can be permuted or – which is the same – one can be substituted for another which is determined by the permutation π. It is also clear that a deleting or adding of a symbol always concerns the constants, i.e. the auxiliary terminal strings only and that it is determined by the mapping $f_{i,j}$ which is as general as possible. E.g. $T_{i,j}$ will be a pure permutation if $f_{i,j}(q_0, q_1, \ldots, q_k) = (q_0, q_1, \ldots, q_k)$ for each sequence $(q_0, q_1, \ldots, q_k)$, i.e. if $f_{i,j}$ is an identity mapping; there will be an effect of deleting or adding if for $f_{i,j}(q_0, q_1, \ldots, q_k) = (q_0^*, q_1^*, \ldots, q_k^*)$ there is an index h such that $q_h \neq I$ and $q_h^* = I$ or on the contrary $q_h = I$ but $q_h^* \neq I$ etc.

But it is important to mention here that these properties of permuting, deleting or adding are not the properties of the mapping $f_{i,j}$ or even $T_{i,j}$ in general but that they are only locally depending on each particular choice of a sequence $(q_0, q_1, \ldots, q_k)$, i.e. the effects can be different for different sequences under the condition that we have not in our mind some special cases (e.g. if L_i and L_j are very small subsets containing only the sentences belonging to one single sentence-form; then the domain of $f_{i,j}$ is a single sequence etc.).

With regard to the transformations which are mappings of phrase-markers the two previous requirements concern the last strings of them only. They determine a very strong correspondence between the proper terminal symbols (namely they do not change, only permute and therefore in the corresponding last strings it has to be the same number of such symbols) but very weak correspondence between the auxiliary terminal strings (there are almost no conditions concerning f besides that the length of sequences has to be preserved).

Further it can be assumed that in each phrase-marker each proper terminal symbol Q is dominated by the corresponding category $\langle C_i \rangle$, i.e. $Q \in C_i$ and all the categories are the non-terminal symbols of an underlying context-free grammar (and obviously $\langle C_i \rangle ::= Q_i$ are some of the rules). According to that assumption there is a very strong correspondence between the categories used in the corresponding phrase-markers.

The last requirement concerning the non-terminal symbols in the corresponding phrase-markers contains also the previous assumption about the proper categories. If V_N and V_N^* are all the nonterminal symbols used in the phrase-marker of the domain and range of a transformation T resp. then by T a correspondence between V_N and V_N^* should be determined that it might be possible to interpret and understand the new phrase-markers using the "is a" relation from the original ones.

That correspondences between V_N and V_N^* of two context-free grammars G and G^* have a great importance in all the linguistic questions where two languages or two grammars have to be considered and compared simultaneously as it is by the translation or in comparative linguistics or when constructing a grammar from several partial grammars for special parts of a language etc. Here a question is touched about the linguistic categories which are independent on the particular languages (see Chomsky [8]).

4. HOMOMORPHISMS OF GRAMMARS

Let $G = \langle V_T, V_N, \mathfrak{R}, S\rangle$ be a context-free grammar, $S \in V_N$, $V_T \cap V_N = \emptyset$, $\mathfrak{R} \subset V_N \times V^\infty$ where $V = V_T \cup V_N$ and V^∞ denotes a free semigroup of strings over V, $G(S)$ denotes the language generated by G from S etc. E.g. e is an identity symbol (with respect to the operation of concatenation), i.e. $xe = ex = x$ for each string $x \in V^\infty$, but $e \notin V^\infty$.

G is said to be *canonical* if the terminal vocabulary V_T is divided into two parts $V_P \neq \emptyset$ and V_A called proper and auxiliary vocabulary resp. (i.e. $V_T = V_P \cup V_A$, $V_P \cap V_A = \emptyset$) in such a way that the following condition is satisfied

(1) if $(x ::= y) \in \mathfrak{R}$ then y has to contain either a nonterminal or a proper symbol.

Then the union $V_C = V_N \cup V_P$ is called the *canonical vocabulary of* G. Separately we shall also extend the notion of the canonical grammar to the case $V_P = \emptyset$, namely each context-free grammar will be called canonical without the necessity to satisfy (1). It will be seen that even in this special case of canonical grammars all further notions will have good sense.

In a canonical grammar G each rule $w_A \in \mathfrak{R}$ or $w_B \in \mathfrak{R}$ (where the capitals A, B are used as indices only) can be expressed in a unique way in the following canonical form

(2) $w_A = (A_0 ::= a_0 A_1 a_1 \ldots A_k a_k)$ or $w_B = (B_0 ::= b_0 B_1 b_1 \ldots B_l b_l)$, where $A_i \in V_C$ and $a_i \in V_A^\infty$ or $a_i = e$ for each $i = 0, 1, \ldots, k$ (and similarly for w_B).

The integer k determines a number of occurences of canonical symbols at the right side of the rule w_A and is said to be the *canonical length of* w_A. The number of occurences of proper symbols in a string x is said to be the *proper length of* x. Obviously an arbitrary string x over V can also be expressed in its canonical form $x_0 X_1 x_1 \ldots X_k x_k = x$, where $X_i \in V_C$ for each $i = 1, 2, \ldots, k$ and $x_i \in V_A^\infty$ or $x_i = e$ for each $i = 0, 1, \ldots, k$. The canonical form is also determined in the case $V_C = V_N$.

The most important characteristic concerning the recursive structure of a context-free grammar G and therefore also the language generated by it is the following ternary relation

(3) $\mathfrak{G}_{\mathfrak{R}} = \{(w_A, w_B, i);\ w_A, w_B \in \mathfrak{R},\ B_0 = A_i \text{ and } 1 \leqq i \leqq k,$ where k is the canonical length of $w_A\}$

which can be called the *relation of applicability of rules*, because the triple (w_A, w_B, i) means that the rule w_B can be used to the i-th canonical symbol (which has to be a nonterminal one) of the right side of the rule w_A. Again it is clear that this notion has an exact meaning in the context-free grammars too.

Now let $G^* = \langle V_P^*, V_A^*, V_N^*, \mathfrak{R}^*, S^* \rangle$ be another canonical grammar and let us assume that either $V_P \neq \emptyset \neq V_P^*$ or $V_P = \emptyset = V_P^*$.

An (*usual*) *homomorphism* of $\mathfrak{R}$ into $\mathfrak{R}^*$ is a mapping Φ of $\mathfrak{R}$ into or onto $\mathfrak{R}^*$ preserving the canonical and proper lengths of rules and preserving the relation of applicability of rules too, i.e. Φ satisfies the following two conditions

(4) if $w \in \mathfrak{R}$ then $\Phi(w) \in \mathfrak{R}^*$ and the right side of $\Phi(w)$ has the same canonical and proper length as the w has,

(5) $(w_A, w_B, i) \in \mathfrak{G}_{\mathfrak{R}}$ if and only if $(\Phi(w_A), \Phi(w_B), i) \in \mathfrak{G}_{\mathfrak{R}^*}$ i.e. $B_0 = A_i$ if and only if $B_0^* = A_i^*$ where we denote the canonical forms as follows: $\Phi(w_A) = (A_0^* ::= a_0^* A_1^* a_1^* \ldots A_k^* a_k^*)$ and $\Phi(w_B) = (B_0^* ::= b_0^* B_1^* b_1^* \ldots B_l^* b_l^*)$.

It is necessary to lay stress on the fact that in (4) and (5) there is no requirement concerning the auxiliary vocabulary or, in the case $V_P = \emptyset = V_P^*$, no requirement concerning the terminal vocabulary at all. At the first sight this is the main difference between our definition of homomorphism and that which was introduced by M. P. Schützenberger [11] and which concerns the vocabularies instead of sets of rules as it is here. Later on, however, some connections between these two different concepts of homomorphism will be established.

As we want to be able to change the word-order of sentences or to compare the sentences distinguished from each other by the ordering of their elements only, it is necessary to introduce a more general notion of homomorphism as follows.

Let $\pi_A, \pi_B, \ldots, \pi_Z$ be permutations assigned to all the particular rules $w_A, w_B, \ldots, \ldots, w_Z$ from $\mathfrak{R}$ such that each π_X permutes the set of integers $\{1, 2, \ldots, k\}$ if and only if k is the canonical length of w_X. A mapping Φ of $\mathfrak{R}$ into or onto $\mathfrak{R}^*$ together with these prescribed permutations $\pi_A, \pi_B, \ldots, \pi_Z$ is said to be the *permutational homomorphism* if Φ satisfies (4) and

(6) $(w_A, w_B, i) \in \mathfrak{G}_{\mathfrak{R}}$ if and only if $(\Phi(w_A), \Phi(w_B), \pi_A(i)) \in \mathfrak{G}_{\mathfrak{R}*}$ i.e. $B_0 = A_i$ if and only if $B_0^* = A^*_{\pi_A(i)}$.

If all the prescribed permutations π_A are identical (i.e. $\pi_A(i) = i$ for each i) then the permutational homorphism is a usual one. Evidently under these assumptions condition (6) reduces to condition (5).

Theorem 1. *Let $\mathfrak{R}$ satisfy the following requirement: Each nonterminal symbol is contained at the right side of a rule and also at the left side of a rule. Now a mapping Φ of $\mathfrak{R}$ into $\mathfrak{R}^*$ satisfying (4) is a permutational homomorphism with respect to the prescribed permutations $\pi_A, \pi_B, \ldots, \pi_Z$ if and only if the following two conditions are satisfied: let us denote $w_A = (A_0 ::= a_0 A_1 a_1 \ldots A_k a_k)$ and $\Phi(w_A) = w_A^* = (a_0^* A_1^* a_1^* \ldots A_k^* a_k^*)$ the canonical forms; then*

(7) *$A_i \in V_N$ if and only if $A^*_{\pi_A(i)} \in V_N^*$ for each $i = 1, 2, \ldots, k$ and*

(8) *there is a mapping τ_N of V_N into V_N^* such that if $A_i \in V_N$ then $A^*_{\pi_A(i)} = \tau(A_i)$ for each $i = 1, 2, \ldots, k$ and $\tau_N(A_0) = A_0^*$ for each $w_A \in \mathfrak{R}$.*

Proof. First of all let us suppose that Φ is a permutational homomorphism with respect to the given permutations $\pi_A, \pi_B, \ldots, \pi_Z$ and let us prove (7) and (8).

If $A_i \in V_N$ for some i, $1 \leqq i \leqq k$ and some $w_A \in \mathfrak{R}$ then there is – according to our requirement concerning $\mathfrak{R}$ – $w_B \in \mathfrak{R}$ such that $B_0 = A_i$ and therefore by (6) $B_0^* = A^*_{\pi_A(i)}$, i.e. $A^*_{\pi_A(i)} \in V_N^*$. If $A_i \in V_P$ for some i, $1 \leqq i \leqq k$ and if it were $A^*_{\pi(i)} \in V_N^*$ then w_A and w_A^* would have different proper lengths which is a contradiction to (4). Therefore Φ satisfies (7).

Further let us consider the set of all couples, $(A_i, A^*_{\pi_A(i)})$ corresponding to all rules $w_A \in \mathfrak{R}$, $w_A = (A_0 ::= a_0 A_1 a_1 \ldots A_k a_k)$ and to all integers $i = 1, 2, \ldots, k$ such that $A_i \in V_N$. We add to this set the couples (A_0, A_0^*) corresponding to all rules $w_A \in \mathfrak{R}$. Formally we assume that each permutation π_A is extended as follows: $\pi_A(0) = 0$. If this set of couples were not a mapping there would be two different couples $(A_i, A^*_{\pi_A(i)})$ and $(B_j, B^*_{\pi_B(j)})$ such that $A_i = B_j$ but $A^*_{\pi_A(i)} \neq B^*_{\pi_B(j)}$. We need to deduce a contradiction from this assumption. There are the following possibilities:

(i) $i \geqq 1, j \geqq 1$; according to the requirement about $\mathfrak{R}$ there exists a rule $w_X \in \mathfrak{R}$ such that $X_0 = A_i = B_j$ where $w_X = (X_0 ::= x_0 X_1 x_1 \ldots X_h x_h)$. Then $(w_A, w_X, i) \in \mathfrak{G}_{\mathfrak{R}}$ and $(w_B, w_X, j) \in \mathfrak{G}_{\mathfrak{R}}$ and therefore by (6) $X_0^* = A^*_{\pi_A(i)}$ and $X_0^* = B^*_{\pi_B(j)}$, i.e. $A^*_{\pi_A(i)} = B^*_{\pi_B(j)}$ which is the required contradiction.

(ii) $i = 0, j \geqq 1$ (and similarly $i \geqq 1, j = 0$); now the first considered couple is (A_0, A_0^*) – according to the extension of π_A – and therefore $A_0 = B_j$ and by (6) $A_0^* = B^*_{\pi_B(j)}$. On the other side $A_0^* = A^*_{\pi_A(0)}$ i.e. $A^*_{\pi_A(0)} = B^*_{\pi_B(j)}$ which is again a contradiction.

(iii) $i = 0, j = 0$; using the other part of the requirement about $\mathfrak{R}$ a rule $w_X \in \mathfrak{R}$ such that $X_t = A_0 = B_0$ for some t, $1 \leqq t \leqq h$ has to exist, where $w_X =$

$= (X_0 :: = x_0 X_1 x_1 \dots X_h x_h)$. Then by (6) $A_0^* = X_{\pi_X(t)}^*$ and $B_0^* = X_{\pi_X(t)}^*$ and therefore $A_{\pi_A(0)}^* = A_0^* = B_0^* = B_{\pi_B(0)}^*$ which is also a contradition.

We have proved that the considered set of couples $(A_i, A_{\pi(i)}^*)$ is (or it determines) a mapping τ_N of V_N into V_N^* and we can therefore write $\tau_N(A_i) = A_{\pi(i)}^*$ always when $A_i \in V_n$ which means that Φ satisfies the condition (8).

If conversely Φ satisfies (4), (7) and (8) then for all rules $w_A, w_B \in \mathfrak{R}$ such that $B_0 = A_i$, where $i \geqq 1$ the following must be valid: $B_0^* = B_{\pi_B(0)}^* = \tau_N(B_0) = \tau_N(A_i) = A_{\pi_A(i)}^*$ which is the condition (6).

The requirement concerning $\mathfrak{R}$ in the theorem 1 is not very strong because every canonical grammar G can be reduced in such a way that the new grammar satisfies this requirement. For this reason there is a very small loss of generality if we mainly consider the homomorphisms such that the mapping τ from Theorem 1 really exists. Moreover we shall consider the homomorphisms Φ such that the assumed mapping τ_N can be extended to the proper vocabulary V_P too, i.e. that the following condition is valid

(9) There is a mapping τ_P of V_P into V_P^* such that if $A_i \in V_P$ then $A_{\pi_A(i)}^* = \tau_P(A_i)$ for each $i = 1, 2, \dots, k$.

If we define the mapping τ of V_C into V_C^* by the conditions $\tau|_{V_N} = \tau_N$ and $\tau|_{V_P} = \tau_P$ then the conditions (8) and (9) can be expressed in another form

(10) $\Phi(A_0 ::= a_0 A_1 a_1 \dots A_k a_k) = (\tau(A_0) :: = a_0^* \, \tau(A_{\pi_A^{-1}(1)}) \, a_1^* \dots \tau(A_{\pi_A^{-1}(k)}) \, a_k^*)$ for each $(A_0 ::= a_0 A_1 a_1 \dots A_k a_k) \in \mathfrak{R}$.

Corollary 1. *Let a canonical grammar* $G = \langle V_P, V_A, V_N, \mathfrak{R}, S \rangle$ *be given and let the sets* V_P^*, V_A^*, V_N^* *be prescribed in such a way that they are mutually disjoint and* $S^* \in V_N^*$. *If* τ_N *is an arbitrary mapping of* V_N *into* V_N^* *and* τ_P *of* V_P *into* V_P^* *and if to each rule* $w_A \in \mathfrak{R}_A$ *having the canonical form* $A_0 :: = a_0 A_1 a_1 \dots A_k a_k$ *a permutation* π_A *of* $\{1, 2, \dots, k\}$ *and a sequence* $(a_0^*, a_1^*, \dots, a_k^*)$ *is prescribed such that* $a_i^* \in V_A^{*\infty}$ *or* $a_i^* = e$ *for each* $i = 0, 1, \dots, k$ *then* $\mathfrak{R}^* = \{(\tau(A_0) :: = a_0^*(A_{\pi_A^{-1}(1)}) \, a_1^* \dots (A_{\pi_A^{-1}(k)}) \, a_k^*); w_A \in \mathfrak{R}\}$ *is the homomorphic image of* $\mathfrak{R}$ *with respect to the permutations* $\pi_A, \pi_B, \dots, \pi_Z$ *and* $G^* = \langle V_P^*, V_A^*, V_N^*, \mathfrak{R}^*, S^* \rangle$ *is a canonical grammar.*

The proof is obvious.

A *permutational homomorphism of a canonical grammar* $G = \langle V_P, V_A, V_N, \mathfrak{R}, S \rangle$ into or onto another canonical grammar $G^* = \langle V_P^*, V_A^*, V_N^*, \mathfrak{R}^*, S^* \rangle$ is a permutational homomorphism Φ of $\mathfrak{R}$ into or onto $\mathfrak{R}$ such that there exists a mapping τ of V_C into V_C^* satisfying the condition $\tau(S) = S^*$.

5. WELL TRANSFORMATIONS

From a pure mathematical point of view it seems to be natural and useful to look for a subclass of transformations of phrase-markers such that all the other transformations could be generated by some type of composition of them. Besides that this subclass should contain the transformations which are as simple as possible.

As the transformations should be mappings of phrase-markers the simplest transformations have to be those which map the simplest phrase-markers.

The simplest phrase-markers as the labelled trees are those containing only one vertex and no edge, but from these — without edges — it is not possible to compose more complicated phrase-markers. Therefore we have to choose the simple phrase-markers containing at least one edge.

From the theoretical point of view it is sufficient to choose only one type of these phrase-markers, namely that one which contains two vertices (one of them being the root) connected by an edge, because each phrase-marker can be composed of them using the identification of vertices only.

But there are very good reasons that we allow also a little more complicated trees containing in general more than one edge but having a very special form, namely that all their edges are connected with the root. Exactly these simple rooted trees — or the double graphs from Sec. 1 — with a labelling of their vertices, correspond to the particular rules of the context-free grammars as it was shown in Fig. 1. And this fact is a crucial point in regard to the next definitions of transformation, because essentially the transformations of phrase-markers will be determined by some mappings of rules, namely by the homomorphisms investigated in the previous Sec. 4.

The following structural transformations are said to be well-transformations because they are essentially the same mappings as the well-translations introduced and studied in [6, 9]. In the following definition the term of marker instead of phrase-marker is used because it is easier and clearer to describe by them the mappings of rules (actually only the labellings will be changed).

Let be given the canonical grammars $G = \langle V_P, V_A, V_N, \mathfrak{R}, S\rangle$ and $G^* = \langle V_P^*, V_A^*, V_N^*, \mathfrak{R}^*, S^*\rangle$, a permutational homomorphism Φ of G onto G^* with the prescribed permutations $\pi_{w_1}, \pi_{w_2}, \ldots, \pi_{w_n}$ corresponding to the rules $w_1, w_2, \ldots, w_n$ from $\mathfrak{R}$ resp. Further let $\mathfrak{M}$ and $\mathfrak{M}^*$ be the sets of all proper markers in G and G^* resp. (i.e. they are proper markers over $\mathfrak{R}$ and $\mathfrak{R}^*$).

Now a *well transformation* T of G into or onto G^* determined by the homomorphism Φ is a mapping of $\mathfrak{M}$ into $\mathfrak{M}^*$ such that if we denote $M = \langle A, r, B, \varphi, \psi\rangle \in \mathfrak{M}$ and $T(M) = M^* = \langle A^*, r^*, B^*, \varphi^*, \psi^*\rangle$ the following conditions are satisfied:

(11) there is an isomorphism ι of the rooted tree $\langle A, r, B\rangle$ onto $\langle A^*, r^*, B^*\rangle$, i.e. ι is a one-to-one mapping of A onto A^* such that $\iota(r) = r^*$ and $(a, b) \in B$ if and only if $(\iota(a), \iota(b)) \in B^*$; in other words we can suppose $A^* = A$, $r^* = r$, $B^* = B$;

(12) if $a \in A$ then $\varphi^*(\iota(a)) = \Phi(\varphi(a))$;

(13) if $(a, b) \in B$ then $\psi^*(\iota(a), \iota(b)) = \pi_{\varphi(a)}(\psi(a, b))$.

It is necessary to note that in canonical grammars we use a little modified notion of marker in the following sense: in the condition (B) of Sec. 1 the labelling ψ concerns the canonical forms of rules, i.e. if $\psi(a, b) = k$ then $\varphi(a) = (X_0 ::= x_0X_1x_1 \ldots \ldots X_mx_m)$ and $\varphi(b) = (Y_0 ::= y_0Y_1y_1 \ldots Y_ny_n)$ and $Y_0 = X_k$, where X_k is the k-th canonical symbol (either nonterminal or proper) and not the k-th symbol at the right side of the rule $\varphi(a)$.

Lemma 1. *The well transformation T determined by a homomorphism Φ of G onto G^* is a mapping of $\mathfrak{M}$ onto $\mathfrak{M}^*$.*

Proof. Let $M^* = \langle A^*, r^*, B^*, \varphi^*, \psi^* \rangle$ be an arbitrary marker such that $M^* \in \mathfrak{M}^*$ and let us define another marker $M = \langle A, r, B, \varphi, \psi \rangle$ as follows: $A = A^*$, $r = r^*$, $B = B^*$; as Φ mapps $\mathfrak{R}$ onto $\mathfrak{R}^*$, $\Phi^{-1}(w^*) \neq \emptyset$ for each $w^* \in \mathfrak{R}^*$ and therefore one can choose a rule $w_k \in \Phi^{-1}(w_k^*)$ for each $w_k^* \in \mathfrak{R}^*$ and then put $\varphi(a) = w_k$ for each $a \in A$ and for each $w_k^* = \varphi^*(a)$. Finally one defines $\psi(a, b) = \pi_{w_k}^{-1}(\psi^*(a, b))$ where $w_k = \varphi(a)$.

Now we have to show that M satisfies (A), (B) and (C) from Sec. 1. Condition (A) is evidently satisfied and also condition (C) immediately follows by the fact that M^* is a marker and $\pi_{w_k}^{-1}$ has to be a permutation again. As to condition (B) we want to prove that if $(a, b) \in B$ and $\varphi(a) = (X_0 ::= x_0X_1x_1 \ldots X_kx_k)$, $\varphi(b) = (Y_0 ::= y_0Y_1y_1 \ldots \ldots Y_ly_l)$ then $X_{\psi(a,b)} = Y_0$, i.e. $(\varphi(a), \varphi(b), \psi(a, b)) \in \mathfrak{G}_{\mathfrak{R}}$. By these assumptions and by the fact that M^* is a marker it follows that $(a, b) \in B^*$ and $(\varphi^*(a), \varphi^*(b), \psi^*(a, b)) \in \in \mathfrak{G}_{\mathfrak{R}^*}$ where $\varphi^*(a) = (X_0^* ::= x_0^*X_1^*x_1^* \ldots X_k^*x_k^*)$ and $\varphi^*(b) = (Y_0^* ::= y_0^*Y_1^*y_1^* \ldots \ldots Y_l^*y_l^*)$, i.e. $X^*_{\psi(a,b)} = Y_0^*$. By condition (6) the required condition follows immediately.

Finally obviously $M \in \mathfrak{M}$ and it is clear that $T(M) = M^*$.

Theorem 2. *If T is well transformation determined by a permutational homomorphism Φ such that there exists a mapping τ of V_C into V_c^* and if $M \in \mathfrak{M}$ then*

1) *the strings $L(M)$ and $L(T(M))$ have the same canonical and proper length and*

2) *if $x_0X_1x_1 \ldots X_kx_k$ and $x_0^*X_1^*x_1^* \ldots X_k^*x_k^*$ are the canonical forms of $L(M)$ and $L(T(M))$ resp. then there exists a permutation Π of $\{1, 2, \ldots, k\}$ such that $X_{\Pi(i)}^* = = \tau(X_i)$ for each $i = 1, 2, \ldots, k$.*

Proof. We shall use an induction with respect to the integer $n = \text{card } A$ where $M = \langle A, r, B, \varphi^*, \psi \rangle$ is a proper marker over $\mathfrak{R}$. Let us denote $T(M) = M^* = = \langle A^*, r^*, B^*, \varphi^*, \psi^* \rangle$ where $A^* = A$, $r^* = r$ and $B^* = B$ which is allowed according to (11).

If $n = 1$ then $A = \{r\}$ and $\varphi(r) = (X_0 ::= x_0X_1x_1 \ldots X_kx_k)$ and $\varphi^*(r) = = (x_0^* ::= x_0^*X_1^*x_1^* \ldots X_k^*x_k^*)$; further $\Phi(\varphi(r)) = \varphi^*(r)$ and therefore by (4) $L(M)$ and $L(M^*)$ have the same canonical and proper length, i.e. the requirement 1) is satisfied. On the other side if $\pi_{\varphi(r)}$ is prescribed permutation corresponding to the rule $\varphi(r)$ then by (10) $X_{\pi_{\varphi(r)}(i)} = \tau(X_i)$ for each $i = 1, 2, \ldots, k$ and therefore we can put $\Pi = \pi_{\varphi(r)}$, i.e. the requirement 2) is satisfied.

Now let be $n > 1$. Then there exists the end vertex $b \in A$ of the rooted tree $\langle A, r, B \rangle$ such that $b \neq r$ (i.e. there is no vertex $c \in A$ such that $(b, c) \in B$). Let a be the unique vertex such that $a \in A$ and $(a, b) \in B$. If we define $\tilde{A} = A - \{b\}$, $\tilde{r} = r$, $\tilde{B} = B - \{(a, b)\}$, $\tilde{\varphi} = \varphi|_{\tilde{A}}$ and $\tilde{\psi} = \psi|_{\tilde{B}}$ then $\tilde{M} = \langle \tilde{A}, \tilde{r}, \tilde{B}, \tilde{\varphi}, \tilde{\psi} \rangle$ is a proper marker over $\mathfrak{N}$ again, i.e. $\tilde{M} \in \mathfrak{M}$, and therefore we can denote $T(\tilde{M}) = \tilde{M}^* = \langle \tilde{A}^*, \tilde{r}^*, \tilde{B}^*, \tilde{\varphi}^*, \tilde{\psi}^* \rangle$. As card $\tilde{A} < n$ we can use the following inductive assumptions; 1) the canonical and proper length of $L(\tilde{M})$ and $L(\tilde{M}^*)$ is the same, i.e. we can denote the canonical forms of $L(\tilde{M})$ and of $L(\tilde{M}^*)$ by $z_0 Z_1 z_1 \dots Z_m z_m$ and $z_0^* Z_1^* z_1^* \dots Z_m^* z_m^*$ resp.; 2) there exists a permutation $\tilde{\Pi}$ of $\{1, 2, \dots, m\}$ such that $Z^*_{\tilde{\Pi}(i)} = \tau(Z_i)$ for each $i = 1, 2, \dots, m$.

Further let us denote $\varphi(b) = (Y_0 ::= y_0 Y_1 y_1 \dots Y_l y_l)$ and $\varphi(a) = (W_0 ::= w_0 W_1 w_1 \dots W_h w_h)$. As $\varphi^*(b) = \Phi(\varphi(b))$ and $\varphi^*(a) = \Phi(\varphi(a))$ we can by (4) denote $\varphi^*(b) = (Y_0^* ::= y_0^* Y_1^* y_1^* \dots Y_l^* y_l^*)$ and $\varphi^*(a) = (W_0^* ::= w_0^* W_1^* w_1^* \dots W_h^* w_h^*)$ and we know that the proper length of $\varphi^*(b)$ and $\varphi(b)$ and similarly of $\varphi^*(a)$ and $\varphi(a)$ is also the same. Let π_Y and π_W be a permutation prescribed by Φ to the rules $\varphi(b)$ and $\varphi(a)$ resp. Then by (10) $Y^*_{\pi_y(i)} = \tau(Y_i)$ for each $i = 1, 2, \dots, l$.

Now according to the definition of the marker $Y_0 = W_{\psi(a,b)}$ and $Y_0^* = W^*_{\psi^*(a,b)}$ in M and M^* resp. On the other side the canonical symbol $W_{\psi(a,b)}$ has to appear in $L(\tilde{M})$ as a canonical symbol Z_p for some p, $1 \leqq p \leqq m$ (this fact is clarified by introduction of a mapping ν, see the end of this section and Lemma 2 and Theorem 3). Similarly $W^*_{\psi(a,b)}$ has to appear in $L(\tilde{M}^*)$ as Z_q^* for some q, $1 \leqq q \leqq m$. Thus $L(M) = z_0 Z_1 z_1 \dots z_{p-1} y_0 Y_1 y_1 \dots Y_l y_l z_p Z_{p+1} \dots Z_m z_m$ and similarly $L(M^*) = z_0^* Z_1^* z_1^* \dots z_{q-1}^* y_0^* Y_1^* y_1^* \dots Y_l^* y_l^* z_q^* Z_{q+1}^* \dots Z_m^* z_m^*$.

By these expressions and by the previous inductive assumptions it is clear that $L(M)$ and $L(M^*)$ have the same canonical and proper length which proves 1).

Further as $L(M)$ and $L(M^*)$ have the canonical expressions $x_0 X_1 x_1 \dots X_k x_k$ and $x_0^* X_1^* x_1^* \dots X_k^* x_k^*$ the following equalities hold: $k = m + l - 1$; $X_i = Z_i$ for $1 \leqq i < p$, $X_i = Y_{i-p+1}$ for $p \leqq i < p + l$, $X_i = Z_{i-l+1}$ for $p + l \leqq i \leqq k$ and similarly $X_j^* = Z_j^*$ for $1 \leqq j < q$, $X_j^* = Y^*_{j-q+1}$ for $q \leqq j < q + l$ and $X_j^* = Z^*_{j-l+1}$ for $q + l \leqq j \leqq k$.

Now we can define the permutation Π of $\{1, 2, \dots, m + l - 1\}$ by a special type of composition of the permutations $\tilde{\Pi}$ of $\{1, 2, \dots, m\}$ and π_Y of $\{1, 2, \dots, l\}$. This composition depends on the prescribed integers p and q and can be determined as follows:

if $1 \leqq i < p$ and $1 \leqq \tilde{\Pi}(i) < q$ then $\Pi(i) = \tilde{\Pi}(i)$

if $1 \leqq i < p$ and $q \leqq \tilde{\Pi}(i) \leqq m$ then $\Pi(i) = \tilde{\Pi}(i) + l$

if $p \leqq i < p + l$ then $\Pi(i) = \pi_Y(i - p + 1) + q - 1$

if $p + l \leqq i \leqq m + l - 1$ and $1 \leqq \tilde{\Pi}(i - l) < q$ then $\Pi(i) = \tilde{\Pi}(i - l)$

if $p + l \leqq i \leqq m + l - 1$ and $q \leqq \tilde{\Pi}(i - l) \leqq m$ then $\Pi(i) = \tilde{\Pi}(i - l) + l$.

Using the inductive assumptions we obtain immediately that $X^*_{\Pi(i)} = \tau(X_i)$ for each $i = 1, 2, \ldots, k$, which proves 2) and accomplishes our proof.

The following part of this section is intended to clarify the statement 2) and the proof of Theorem 2. Therefore we shall not prove some results in this direction.

Let $M = \langle A, r, B, \varphi, \psi \rangle$ be a proper marker over $\mathfrak{R}$ and let $x_0 X_1 x_1 \ldots X_k x_k$ be the canonical form of $L(M)$. If we take an arbitrary vertex $a \in A$ and if we denote the canonical form of the rule $\varphi(a)$ as $(Y_0 ::= y_0 Y_1 y_1 \ldots Y_l y_l)$ then it can happen that there exists an integer n, $1 \leqq n \leqq l$, and such that there does not exist a vertex $b \in A$ such that $(a, b) \in B$ and $\psi(a, b) = n$. We can say that the *n-th canonical symbol corresponding to the vertex a is free* and we shall express this fact by the pair $[a, n]$.

Now there is exactly one path in $\langle A, r, B \rangle$ starting in r and ending in a; let us write it as a sequence $(r = b_0, b_1, \ldots, b_m = a)$ (i.e. $(b_{i-1}, b_i) \in B$ for each $i = 1, 2, \ldots, m$). By this path one uniquely determines the following sequence of integers $(k_1, k_2, \ldots, k_m)$ defined by the condition $k_i = \psi(b_{i-1}, b_i)$ for each $i = 1, 2, \ldots, m$. If $m = 0$, i.e. if $a = r$, we put instead of the sequence $(k_1, \ldots, k_m)$ the single number 0. In this way to each vertex $a \in A$ there belongs its *value* $v(a) = (k_1, k_2, \ldots, k_m)$ and we can suppose that the set of values or the set of all vertices is fully ordered by the lexicographical ordering (i.e. according to the first difference from left to right) $\leqq$. Obviously two different vertices have two different values.

Further the marker $M' = \langle A', r', B', \varphi', \psi' \rangle$ is uniquely determined where $A' = \{c \in A;$ either $v(c) \leqq v(a)$ or $v(c) = (h_1, h_2, \ldots, h_m)$ and $h_i = k_i$ for $i = 1, 2, \ldots, m$, where $v(a) = (k_1, k_2, \ldots, k_m)\}$; $r' = r$; $B' = B \cap (A' \times A')$; $\varphi'(c) = \varphi(c)$ for $c \neq b_i$, where $i = 0, 1, \ldots, m$ and $(b_0, b_1, \ldots, b_m)$ is the path connecting r and a; if $c = b_i$ for some i, $0 \leqq i < m$ and if $(Z_0 ::= z_0 Z_1 z_1 \ldots Z_h z_h)$ is the canonical form of $\varphi(c)$ then $\varphi'(c)$ has the canonical form $(Z_0 ::= z_0 Z_1 z_1 \ldots z_{k_{i+1}-1} Z_{k_{i+1}})$ where $k_{i+1} = \psi(b_i, b_{i+1}) \leqq h$; if $c = a$ then $\varphi'(a) = (Y_0 ::= y_0 Y_1 y_1 \ldots y_{n-1} Y_n)$; at last $\psi'(c, d) = \psi(c, d)$ for each $(c, d) \in B'$.

By all the previous constructions there is defined a mapping ν of all the pairs $[a, n]$ of M into the set of all integers $\{1, 2, \ldots\}$ if we put $\nu[a, n] = p$, where p is the canonical length of $L(M')$.

Lemma 2. *Let $M = \langle A, r, B, \varphi, \psi \rangle$ be a proper marker over the $\mathfrak{R}$, where $G = \langle V_P, V_A, V_N, \mathfrak{R}, S \rangle$ is a canonical grammar. Then the mapping ν (belonging to M) is one-to-one mapping onto the set $\{1, 2, \ldots, k\}$ where k is the canonical length of $L(M)$ and $Y_n = X_{\nu[a,n]}$ for each pair $[a, n]$, where the canonical forms of $\varphi(a)$ and $L(M)$ are $(Y_0 ::= y_0 Y_1 y_1 \ldots Y_l y_l)$ and $x_0 X_1 x_1 \ldots X_k x_k$ respectively.*

Theorem 3. (Continuation of Theorem 2.) *Using the mappings ν in M and ν^* in M^* the permutation Π of 2) can be determined as follows*

$$(14) \qquad \Pi(\nu[a, n]) = \nu^*[a, \pi_{\varphi(a)}(n)]$$

for each pair $[a, n]$ in M where $\pi_{\varphi(a)}$ is a permutation corresponding to the rule $\varphi(a)$.

A language generated by a canonical grammar is said to be *canonical*. Let $G(S)$ and $G^*(S^*)$ be the canonical languages generated by the canonical grammars $G = \langle V_P, V_A, V_N, \mathfrak{R}, S\rangle$ and $G^* = \langle V_P^*, V_A^*, V_N^*, \mathfrak{R}^*, S^*\rangle$ resp.

The sequences $(x_0, x_1, \ldots, x_k)$ such that $x_i \in V_A^\infty$ or $x_i = e$ and $x_0X_1x_1 \ldots X_kx_k \in G(S)$ where $X_i \in V_P$ for each $i = 0, 1, \ldots, k$ are said to be forms of $G(S)$. Let us denote the set of all forms of $G(S)$ by the symbol $\mathfrak{F}_G$.

A mapping t of $G(S)$ into or onto $G^*(S^*)$ is said to be a *descriptive transformation* if there exists a mapping σ of V_P into V_P^* and if there exists a permutation π_X of $\{1, 2, \ldots, k\}$ corresponding to the string $x \in G(S)$ with the proper length k such that the following condition is satisfied:

(15) if $X_0X_1x_1 \ldots X_kx_k$ is the canonical form of x and if $x_0^*X_1^*x_1^* \ldots X_l^*x_l^*$ is the canonical form of $x^* = t(x)$ then $k = l$ and $X_{\pi_x(i)}^* = \sigma(X_i)$ for each $i = 1, 2, \ldots, k$.

A descriptive transformation t is said to be a *descriptive form transformation* if there exists a mapping f of $\mathfrak{F}_G$ into $\mathfrak{F}_G^*$ such that f preserves the length of forms and

(16) if $x_0X_1x_1 \ldots X_kx_k$ is the canonical form of x and if $x_0^*X_1^*x_1^* \ldots X_k^*x_k^*$ is the canonical form of $t(x) = x^*$ then $(x_0^*, x_1^*, \ldots, x_k^*) = f(x_0, x_1, \ldots, x_k)$ and

(17) if the permutations π_X and π_Y correspond to the canonical forms $x_0X_1x_1 \ldots \ldots X_kx_k \in G(S)$ and $y_0Y_1y_1 \ldots Y_ly_l$ $G(S)$ such that $k = l$ and $x_i = y_i$ for each $i = 0, 1, \ldots, k$, then $\pi_X = \pi_Y$.

Let us note that for these notions the requirement $V_P \neq \emptyset$ is essential.

We say that the descriptive transformation t of $G(S)$ into $G^*(S^*)$ is *induced* by the well transformation T of $\mathfrak{M}$ into $\mathfrak{M}^*$ if

(18) $L(T(M)) = t(L(M))$ for each $M \in \mathfrak{M}_G$,
where $\mathfrak{M}_G = \{M \in \mathfrak{M};\ F(M) = S \text{ and } L(M) \in V_T^\infty\}$.

Let us note that if $M \in \mathfrak{M}_G$ then $L(M) \in G(S)$ but it can happen that $L(M) \in G(S)$ but $M \notin \mathfrak{M}_G$.

Theorem 4. *If the well transformation T is determined by a permutational homomorphism Φ of a canonical grammar G into G^* then T induces a descriptive transformation t of $G(S)$ into $G^*(S^*)$ if and only if the following is valid*

(19) *if $L(M_1) = L(M_2)$ then $L(T(M_1)) = L(T(M_2))$ for all $M_1, M_2 \in \mathfrak{M}_G$.*

If T satisfies (19) and Φ maps $\mathfrak{R}$ onto $\mathfrak{R}^*$ then t maps $G(S)$ onto $G^*(S^*)$.

Proof. If $M \in \mathfrak{M}_G$ then $L(M) \in G(S)$, $F(M) = S$ and $L(M) \in V_T^\infty$ which means that the proper and canonical length of $L(M)$ are the same. Now by the definition of the homomorphism Φ of the grammar G into G^* (i.e. not only of $\mathfrak{N}$ into $\mathfrak{N}^*$; see the end of Sec. 4) there exists the mapping τ of V_C into V_C^* such that $\tau(S) = S^*$ and therefore by (10) $F(T(M)) = S^*$. Further in virtue of the part 1) of Theorem 2 $L(M)$ and $L(T(M))$ have the same proper and canonical lengths and therefore $L(T(M)) \in V_T^{*\infty}$. Thus we have showed that $T(M) \in \mathfrak{M}_G^*$ and therefore also $L(T(M)) \in G^*(S^*)$. Now it is clear that by T a binary relation is determined containing all the pairs $(L(M), L(T(M))) \in G(S) \times G^*(S^*)$ and (19) is a necessary and sufficient condition when this relation is a function. The remaining part of this Theorem follows immediately by Lemma 1.

Corollary 2. *Let T be a well transformation determined by a permutational homomorphism Φ of G into G^*. If G is not ambiguous or if $V_A = \emptyset$ and Φ is not permutational, then the condition* (19) *is satisfied.*

Proof is obvious.

The basic problem concerning the descriptive transformations is the problem of the construction of a well transformation which induces the given descriptive transformation. The canonical grammars generating the given canonical languages either are given fixed or they are to be chosen suitably such that they admit a permutational homomorphism determining the required well transformation. Obviously this problem need not always have a solution.

7. DECOMPOSITION TRANSFORMATIONS

As there can be many different (eventually permutational) homomorphisms of G into G^* it is possible to introduce more general structural transformations than the well ones are. For this purpose we shall use some decompositions of markers into submarkers such that each of these submarkers will be mapped by a different homomorphism (or by a different well transformation).

Therefore first of all let us suppose that a *procedure $\mathcal{D}$ how to determine a decomposition* $\mathcal{D}(M)$ of a marker $M \in \mathfrak{M}$ is given. It can be allowed that the procedure $\mathcal{D}$ be not applicable to all markers from $\mathfrak{M}$ but only to some subset. This subset will be denoted by $\mathfrak{M}_{\mathcal{D}}$.

A *decomposition* $\overline{M}$ of a proper marker $M = \langle A, r, B, \varphi, \psi \rangle$ is a set of proper markers $M_1, M_2, \ldots, M_n$ which satisfy the following conditions:

(20) M_i is a submarker of M for each $i = 1, 2, \ldots, n$, i.e. if $M_i = \langle A_i, r_i, B_i, \varphi_i, \psi_i \rangle$ then $A_i \subset A$; $B_i = B \cap (A_i \times A_i)$;

to each vertex $a_i \in A_i$ there is a path in M from r to a_i containing r_i; $\varphi_i = \varphi|_{A_i}$; $\psi_i = \psi|_{B_i}$ and further

(21) $\bigcup_{i=1}^{n} A_i = A$ and $A_i \cap A_j = \emptyset$ for each $i, j = 1, 2, \ldots, n$ where $i \neq j$.

A decomposition of an unproper marker is only this unproper marker itself and a decomposition of a disconnected marker is a sequence of the decompositions of its components.

Lemma 3. *There is a one-to-one correspondence between the set of all decompositions of a proper marker and the set of all subsets of its vertices containing the root. Namely, the elements of the given subset of vertices are the prescribed roots of the corresponding submarkers.*

Proof. Let a decomposition $\overline{M} = \{M_1, M_2, \ldots, M_n\}$ be given. Then $r_1, r_2, \ldots, r_n$ is the required subset of vertices because according to (21) there has to be an index j such that $r_j = r$, where r is the root of the considered proper marker M.

If on the contrary the required subset $\{r_1, r_2, \ldots, r_n\} \subset A$ is given and if e.g. $r_j = r$, then the subsets A_i for each $i = 1, 2, \ldots, n$ can be defined as follows: using the paths from r to r_i we take a vertex r_i such that its path has the maximum length and we define $A_i = \{a_i \in A$; there is a path in M from r to a_i which contains $r_i\}$; it is clear that by the set of vertices $A - A_i$ (as far as $A - A_i \neq \emptyset$) again a proper submarker of M is determined and therefore the described construction may be repeated. Obviously the obtained subsets $A_1, A_2, \ldots, A_n$ satisfy (21) and if we define $B_i = B \cap (A_i \times A_i)$, $\varphi_i = \varphi|_{A_i}$, $\psi_i = \psi|_{B_i}$ then $M_i = \langle A_i, r_i, B_i, \varphi_i, \psi_i \rangle$ is a submarker of M which satisfies (20). Therefore $\overline{M} = \{M_1, M_2, \ldots, M_n\}$ is a decomposition of M with the prescribed set of roots $\{r_1, r_2, \ldots, r_n\}$.

Using this lemma a decomposition procedure of a marker can be determined as a procedure determining some subset of vertices of the given marker.

By a decomposition $\overline{M} = \{M_1, M_2, \ldots, M_n\}$ of a proper marker $M = \langle A, r, B, \varphi, \psi \rangle$ the following *factor-marker* $M_0 = \langle A_0, r_0, B_0, \varphi_0, \psi_0 \rangle$ of M can be defined: $A_0 = \{r_1, r_2, \ldots, r_n\}$ (but it would also be possible to put $A_0 = \{M_1, M_2, \ldots, M_n\}$ because that is unimportant with respect to an isomorphism of markers); $r_0 = r$ (by Lemma 3 r_j has to be such that $r_j = r$); $B_0\{(r_i, r_j)$; there is $a_i \in A_i$ such that $(a_i, r_j) \in B$ and $i \neq j$, where $i, j = 1, 2, \ldots, n\}$; $\varphi_0(r_i) = (F(M_i) :: = L(M_i))$ for each $i = 1, 2, \ldots, n$ and the integers $\psi_0(r_i, r_j)$ for all j such that $(r_i, r_j) \in B_0$ are determined in a non arithmetical way as follows: we take all the paths in M_i which connect the root r_i with an arbitrary end vertex $a_i \in A_i$; by each end vertex a_i its value $v(a_i)$ is determined (see the end of Sec. 5), these values are ordered lexicographically and if in the k-th place of this ordering is $v(a_i)$ such that $(a_i, r_j) \in B$ then we put $\psi_0(r_i, r_j) = k$.

We will not prove that the factor-marker of a proper marker is really a marker. Evidently the factor-marker M_0 is not a marker over the set of rules $\mathfrak{R}$ as the original marker M is but it is easy to construct new rules from the rules of $\mathfrak{R}$ over which the factor-marker is defined. A factor-marker of a disconnected marker is a sequence of factor-markers of its components.

Now let us introduce further assumptions concerning the decomposition procedure $\mathcal{D}$. We shall say that *$\mathcal{D}$ has the length n* if $\mathcal{D}(M) = \{M_1, M_2, \ldots, M_n\}$ for each

$M \subset \mathfrak{M}_{\mathscr{D}}$, i.e. if each decomposition consists of n submarkers. Further we shall assume that an *ordering* of each decomposition $\mathscr{D}(M) = \{M_1, M_2, \ldots, M_n\}$ can be established. This ordered decomposition will be denoted by $[M_1, M_2, \ldots, M_n]$ and the corresponding procedure $\mathscr{D}$ will be called the *ordered decomposition procedure* (there are many different possibilities how to order the submarkers $M_1, M_2, \ldots, M_n$ or – by the lemma – the vertices of M).

A sequence of homomorphisms $\Phi_1, \Phi_2, \ldots, \Phi_n$ of G into G^* (whose corresponding mappings of V_C into V_C^* are $\tau_1, \tau_2, \ldots, \tau_n$) is *compatible* with the ordered decomposition procedure $\mathscr{D}$ of the length n if the following condition is satisfied:

(22) $\tau_i(F(\varphi_0(r_j))) = \tau_j(F(\varphi_0(r_j)))$ for each $(r_i, r_j) \in B_0$ and for each $M \in \mathfrak{M}_{\mathscr{D}}$ where M_0 is a factor-marker of M.

If we denote $\varphi_0(r_j) = (x :: = y)$ then – a single rule can always be considered as a proper marker – $F(\varphi_0(r_j)) = x$ and the condition (22) requires $\tau_i(x) = \tau_j(x)$.

Now a *decomposition transformation* T of G into G^* determined by an ordered decomposition procedure $\mathscr{D}$ of the length n and by a compatible sequence of homomorphisms $\Phi_1, \Phi_2, \ldots, \Phi_n$ of G into G^* is a mapping of $\mathfrak{M}_{\mathscr{D}}$ into $\mathfrak{M}^*$ such that if we denote $M = \langle A, r, B, \varphi, \psi \rangle$, $T(M) = M^* = \langle A^*, r^*, B^*, \varphi^*, \psi^* \rangle$ and $\mathscr{D}(M) = [M_1, M_2, \ldots, M_n]$ the following conditions are satisfied:

(11′) $A^* = A$; $r^* = r$; $B^* = B$;

(12′) if $a \in A_i$ then $\varphi^*(a) = \Phi_i(\varphi(a))$ for each $i = 1, 2, \ldots, n$;

(13′) if $(a, b) \in B$ $a \in A_i$ then $\psi^*(a, b) = \pi^{(i)}_{\varphi(a)}(\psi(a, b))$ for each $i = 1, 2, \ldots, n$, where $\pi^{(i)}_{\varphi(a)}$ is the permutation belonging to the rule $\varphi(a)$ in the permutational homomorphism Φ_i.

Obviously the previous notation $M_i = \langle A_i, r_i, B_i, \varphi_i, \psi_i \rangle$ for each $i = 1, 2, \ldots, n$ e.t.c. is assumed.

It is clear that each well transformation is a decomposition because we can always choose a trivial decomposition procedure of the length 1 which is of course ordered.

On the other hand everybody expects that after the necessary modifications Theorems 2 and 3 remain to be valid for the decomposition transformations, but we shall not give complete proofs here (they can be given in a strong analogy to the proofs of Theorems 2 and 3).

Let us assume that a decomposition transformation T of G into G^* is given and that $\mathscr{D}$ is its ordered decomposition procedure of the length n. If $\mathscr{D}(M) = [M_1, M_2, \ldots, M_n]$ where $M \in \mathfrak{M}_{\mathscr{D}}$ and $M_i = \langle A_i, r_i, B_i, \varphi_i, \psi_i \rangle$ for each $i = 1, 2, \ldots, n$ and if $T(M) = M^*$ then by (11′) and by Lemma 3 a decomposition $\overline{M}^* = [M_1^*, M_2^*, \ldots, M_n^*]$ of M^* is determined (the prescribed roots of submarkers for $\overline{M}^*$ are the same as for M^*, i.e. $\{r_1, r_2, \ldots, r_n\}$). Thus another ordered decomposition procedure $\mathscr{D}$ is determined putting $\mathscr{D}^*(M^*) = \overline{M}^*$ for each $M^* \in \mathfrak{M}^*_{\mathscr{D}^*} = \{M^* \in \mathfrak{M}^*;$ there is $M \in \mathfrak{M}$ such that $M^* = T(M)\}$.

If $M_0 = \langle A_0, r_0, B_0, \varphi_0, \psi_0 \rangle$ and $M_0^* = \langle A_0^*, r_0^*, B_0^*, \varphi_0^*, \psi_0^* \rangle$ is a factor-marker of M and M^* resp. corresponding to the decomposition $\mathscr{D}(M)$ and $\mathscr{D}^*(M^*)$ resp. then it is easy to see that by T another well transformation T_0 of the factor-markers is determined. This factor transformation T_0, i.e. the underlying homomorphism and the permutations corresponding to the particular rules, can be explicitly determined by Theorems 2 and 3 (but it is necessary to give a complete definition of the modified grammar and the set of rules concerning the factor-marker; this has not been done here).

Remark. There is another possibility how to describe the situation by decomposition transformations. Instead of introducing factor-markers one can introduce rather different algebraic structure $\tilde{M} = \langle \tilde{A}, \tilde{r}, \tilde{B}, \tilde{\varphi}, \tilde{\psi} \rangle$ corresponding to the decomposition $\overline{M} = \{M_1, M_2, \dots, M_n\}$ which is not a marker. Here $\langle \tilde{A}, \tilde{r}, \tilde{B} \rangle$ is the same rooted tree as in the factor-marker only φ does not assign particular rules but the whole submarkers to the vertices r_i, M_i and ψ is again the same as in the factor-marker. Obviously this definition has a recurrent character.

Now we shall introduce a special type of decomposition transformations which are very close to the singulary transformations described by Chomsky, Bach a.o. using the notions of structural index and transformational rule.

Each string $L(M)$ where M is a marker in the canonical grammar G such that $F(M) = S$ can be called a *structural index in G.* In fact the most important cases occur if $L(M)$ contains also other elements than the terminal symbols.

A submarker $M' = \langle A', r', B', \varphi', \psi' \rangle$ of a proper marker $M = \langle A, r, B, \varphi, \psi \rangle$ is said to be *main* or *secondary* if $r' = r$ or $r' \neq r$ resp. It is clear that in a decomposition of a proper marker exactly one of the submarkers must be main. M' is said to be an *end submarker* of M if each vertex $a' \in A'$ which is an end vertex in M' is an end vertex in M too (let us remind that a' is an end vertex in M' if there is no $b' \in A'$ such that $(a', b') \in B'$). In a decomposition there must be at least one end submarker and at least as many end submarkers as is the difference between canonical and proper length of the last string belonging to its main submarker.

A specialization of our decomposition transformation depends on the used decomposition procedure $\mathscr{D}$. We shall be concerned with the decompositions such that each their secondary submarker must be an end submarker, i.e. in other words the factor-markers of these decompositions are extremely simple because they contain only the root and the end vertices. Such decompositions will be called *index decompositions.*

Lemma 4. *Let M' be a main submarker of a proper marker M. There exists exactly one index decomposition of M containing M'.*

Proof. If $M = \langle A, r, B, \varphi, \psi \rangle$ and $M' = \langle A', r', B', \varphi', \psi' \rangle$ then either $A' = A$, i.e. $M' = M$, which is a trivial case or $A' \neq A$, i.e. $A - A' \neq \emptyset$. In this last case $\langle A - A', B \cap ((A - A') \times (A - A')) \rangle$ is a directed graph the connected components of which are the rooted trees $\langle A_i, r_i, B_i \rangle$ for $i = 1, 2, \ldots, n$. If we define $\varphi_i = \varphi|_{A_i}$ and $\psi_i = \psi|_{B_i}$ for each $i = 1, 2, \ldots, n$ then $M_i = \langle A_i, r_i, B_i, \varphi_i, \psi_i \rangle$ are submarkers of M. It is easy to show that M_i is an end submarker of M for each $i = 1, 2, \ldots, n$ and therefore $\{M', M_1, M_2, \ldots, M_n\}$ is an index decomposition of M containing M'.

If conversely $\{M', M_1, M_2, \ldots, M_n\}$ is an index decomposition of M and $\{M', M'_1, M'_2, \ldots, M'_m\}$ another index decomposition, then the equality of these decompositions follows by a simple induction with respect to n.

Lemma 5. *Let M' and M'' be the main submarkers of a proper marker M and let M be a marker over a not ambiguous canonical grammar G. If $L(M') = L(M'')$ then $M' = M''$ but this assertion need not be valid if G is ambiguous.*

The proof follows immediately by the definition of ambiguity of G.

Finally an ordered decomposition procedure $\mathscr{D}$ which is determined by a structural index w in an unambiguous canonical grammar G can be described as follows: $\mathfrak{M}_{\mathscr{D}} = \{M \in \mathfrak{M}$; there is a main submarker M' of M such that $L(M') = w\}$; if $M \in \mathfrak{M}_{\mathscr{D}}$ then by Lemma 4 there exists only one main submarker M' of M such that $L(M') = M$ and therefore by Lemma 5 there exists exactly one index decomposition $\overline{M}$ of M containing M' and we put $\mathscr{D}(M) = \overline{M}$. By Lemma 4 it follows that the length of $\mathscr{D}$ is equal to the difference of the canonical and proper length of w and it is clear that $\overline{M}$ can be ordered in a unique way using a one-to-one correspondence between all the end submarkers of $\overline{M}$ and all the end vertices of the factor-marker M_0 of M which is determined by $\overline{M}$ (this ordering corresponds to the ordering of the occurrences of nonterminal symbols in w). Such ordered decomposition procedure with a fixed length will be called an *index decomposition procedure*.

Now let us assume that a decomposition transformation T of a canonical grammar G into G^* is given, the underlying compatible sequence of homomorphisms of which is $\Phi_1, \Phi_2, \ldots, \Phi_m$ and the decomposition procedure of which is an index decomposition procedure $\mathscr{D}$ with the prescribed structural index $w = w_0 W_1 w_1 \ldots \ldots W_n w_n$ (this is the canonical form of w). It is clear that there are exactly $m - 1$ occurrences of nonterminal symbols among the canonical symbols $W_1, W_2, \ldots, W_n$, i.e. $m - 1 \leqq n$. If $M \in \mathfrak{M}_{\mathscr{D}}$, then $\mathscr{D}(M) = [M_1, M_2, \ldots, M_m]$ and we can assume that the first submarker M_1 is main and that the remaining submarkers M_i correspond (in the factor-marker) to the nonterminal occurrences $W_{j_{i-1}}$ for each $i = 2, 3, \ldots, m$ where $1 \leqq j_1 < j_2 < \ldots < j_{m-1} \leqq n$.

It is easy to show that the corresponding decomposition procedure $\mathscr{D}^*$ of $M^* = T(M)$ is again an index decomposition procedure because the main submarker M_1 is mapped by Φ_1 (Φ_1 is nothing else than a well transformation; obviously the possible permutations corresponding to the particular rules are always assumed) onto a main submarker M_1^* of M^* and from the fact $L(M_1) = w$ by Theorems 2 and 3 it follows that $L(M_1^*) = w_0^* W_1^* w_1^* \ldots W_n^* w_n^*$ and that there is a permutation Π, of $\{1, 2, \ldots, n\}$ satisfying the condition $W_{\Pi(i)}^* = \tau(W_i)$ for each $i = 1, 2, \ldots, n$ where τ is a mapping of V_C into V_C^* assumed by T.

Now a pair of structural indices (w, w^*) together with the permutation Π can be called a *transformational rule* of T. It is necessary to note here that a transformational rule need not determine the decomposition transformation T uniquely and therefore a transformational rule is not sufficient to determine T. The necessary additional information concerning the vertices and edges of new phrase-markers always given by the linguists (see e.g. Chomsky [10], Bach [12]) is here fully expressed by the underlying homomorphisms and their permutations.

It is possible to deduce the necessary and sufficient conditions for a pair (w, w^*) of structural indices in order that it may be a transformational rule.

In a special simple case when $\Phi_i = \Phi$ for each $i = 1, 2, \ldots, m$ the decomposition procedure $\mathscr{D}$ is not necessary and it is sufficient to consider a well transformation instead of that of decomposition. The prescribed structural index w remains to be necessary for the determination of the domain $\mathfrak{M}_{\mathscr{D}}$ of the well transformation T.

In order to clarify the value of the transformational rule we can say that by the transformational rule a partially descriptive transformation is prescribed (the descriptive transformation is used here in a broader sense which does not require only the terminal symbols to be concerned – the essential characteristic is that a descriptive transformation maps string onto strings but a structural transformation maps markers onto markers) and we are looking for a structural transformation by which the given descriptive transformation would be induced.

8. EXAMPLES

In this Section different examples are considered. Examples 1 and 2 concern the pure linguistic point of view but the further examples are more abstract and finally the last examples concern some pure mathematical questions, not completely solved here.

Example 1. The canonical grammars $G_{\text{act}} = \langle V_P, V_A, V_N, \mathfrak{R}, \langle S \rangle\rangle$ and $G^*_{\text{pas}} = \langle V^*_P, V^*_A, V^*_N, \mathfrak{R}^*, \langle S^* \rangle\rangle$ are determined as follows:

$V_P = \{$father, daughter, desk, ..., kill, see, like, ...$\}$,
$V^*_P = \{$father, daughter, desk, ..., killed, seen, liked, ...$\}$,
$V_A = \{$the, a, s, /$\}$, where / denotes space,
$V^*_A = \{$the, a, is, by, /$\}$,
$V_N = \{\langle \text{S} \rangle, \langle \text{NP} \rangle, \langle \text{VP} \rangle, \langle \text{N} \rangle, \langle \text{V}_{\text{act}} \rangle, \langle \text{V}_{\text{base}} \rangle\}$,
$V^*_N = \{\langle \text{S}^* \rangle, \langle \text{NP} \rangle, \langle \text{VP} \rangle, \langle \text{N} \rangle, \langle \text{V}_{\text{pas}} \rangle, \langle \text{V}_{\text{pret}} \rangle\}$,

$\mathfrak{R} = \{w_1, w_2, \ldots, w_{11}\}$ and $\mathfrak{R}^* = \{w^*_1, w^*_2, \ldots, w^*_{11}\}$ where the particular rules w_i or w^*_i are given in the i-th row of the following list in the left or right column resp.

1.	$\langle S\rangle$	$::= \langle NP\rangle\langle VP\rangle$	1.	$\langle S^*\rangle$	$::= \langle VP\rangle\langle NP\rangle$
2.	$\langle NP\rangle$	$::=$ /the $\langle N\rangle$	2.	$\langle NP\rangle$	$::=$ /the $\langle N\rangle$
3.	$\langle NP\rangle$	$::=$ /a$\langle N\rangle$	3.	$\langle NP\rangle$	$::=$ /a$\langle N\rangle$
4.	$\langle VP\rangle$	$::= \langle V_{act}\rangle\langle NP\rangle$	4.	$\langle VP\rangle$	$::= \langle NP\rangle\langle V_{pas}\rangle$
5.	$\langle V_{act}\rangle$	$::= \langle V_{base}\rangle$ s	5.	$\langle V_{pas}\rangle$	$::=$ /is $\langle V_{pret}\rangle$/by
6.	$\langle N\rangle$	$::=$ /father	6.	$\langle N\rangle$	$::=$ /father
7.	$\langle N\rangle$	$::=$ /daughter	7.	$\langle N\rangle$	$::=$ /daughter
8.	$\langle N\rangle$	$::=$ /desk	8.	$\langle N\rangle$	$::=$ /desk
⋮					
9.	$\langle V_{base}\rangle$	$::=$ /kill	9.	$\langle V_{pret}\rangle$	$::=$ /killed
10.	$\langle V_{base}\rangle$	$::=$ /see	10.	$\langle V_{pret}\rangle$	$::=$ /seen
11.	$\langle V_{base}\rangle$	$::=$ /like	11.	$\langle V_{pret}\rangle$	$::=$ /liked
⋮			⋮		

It is easy to see that the canonical language $G(S)$ and also $G^*(S^*)$ contains 108 different sentences. Therefore we shall not enumerate all the corresponding pairs of sentences in the well known active-passive descriptive transformation t of $G(S)$ onto $G^*(S^*)$. E.g. if $s =$ /the/father/kills/a/daughter then $s \in G(S)$ and $t(s) =$ = /a/daughter/is/killed/by/the/father $\in G^*(S^*)$ etc. The form of s is (/the, e, s/a, e) and of $t(s)$ is (/a, /is/, /by/the, e/).

All the sentences of $G(S)$ have the canonical length 3 and the required transformation π_s of $\{1, 2, 3\}$ is equal $\pi_s(1) = 3$, $\pi_s(2) = 2$ and $\pi_s(3) = 1$ for each $s \in G(s)$. The required mapping σ of V_P into V_P^* is determined as follows: σ(father) = father, σ(daughter) = daughter, σ(desk) = desk, ..., σ(kill) = killed, σ(see) = seen, σ(like) = = liked, ...

Both grammars G_{act} and G^*_{pas} are chosen in such a way that there exists a structural transformation T of G_{act} into G^*_{pas} which induces t.

The mapping τ is defined as follows: $\tau(x) = \sigma(x)$ for each $x \in V_P$ and $\tau(\langle S\rangle) = = \langle S^*\rangle$, $\tau(\langle NP\rangle) = \langle NP\rangle$, $\tau(\langle VP\rangle) = \langle VP\rangle$, $\tau(\langle N\rangle) = \langle N\rangle$, $\tau(\langle V_{act}\rangle) = \langle V_{pas}\rangle$, $\tau(\langle V_{base}\rangle) = \langle V_{pret}\rangle$. The homomorphism Φ of G onto G^* is defined as follows: $\Phi(w_i) = w_i^*$ for each $i = 1, 2, \ldots, 11$. Φ is permutational and there are nonidentical permutations for w_1 and w_4 only. They are prescribed as follows: $\pi_1(1) = 2$, $\pi_1(2) = = 1$; $\pi_4(1) = 2$, $\pi_4(2) = 1$. By τ and Φ a well transformation T is determined which induces t indeed.

In Fig. 5 there are phrase-markers P and P^* of the sentences s and $t(s)$ resp. which were chosen above.

The well transformation T can be considered as a decomposition transformation with respect to the index decomposition procedure $\mathscr{D}$ determined by the structural index $w = \langle NP\rangle\langle V_{act}\rangle\langle NP\rangle$ and with respect to the transformational rule (w, w^*) where $w^* = \langle NP\rangle\langle V_{pas}\rangle\langle NP\rangle$ together with the permutation Π of $\{1, 2, 3\}$ such that $\Pi(1) = 3$, $\Pi(2) = 2$, $\Pi(3) = 1$. Obviously it is necessary to put $\Phi_i = \Phi$ for each

$i = 1, 2, 3, 4$ (it is clear that $\mathscr{D}$ has the length 4). Here it is clear how the permutation Π is determined by the permutations π_1 and π_4.

In Fig. 5 the particular sub-phrase-markers of the index decompositions $\mathscr{D}(P)$ and $\mathscr{D}^*(P^*)$ are marked out by the dotted lines.

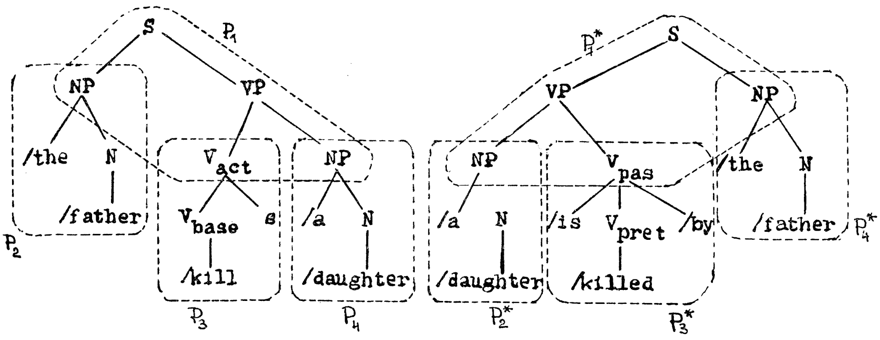

Fig. 5.

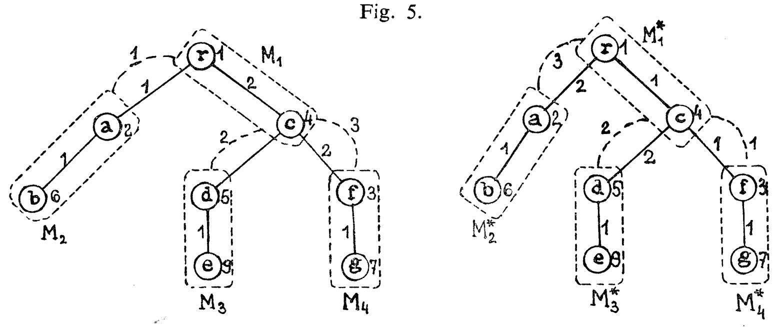

Fig. 6.

According to the previous definitions the markers and submarkers M, M^* and M_i, M_i^* are used instead of phrase-markers. These markers and their decompositions and the corresponding factor-markers M_0^* are shown in Fig. 6. If $M = \langle A, r, B, \varphi, \psi \rangle$ then $A = \{r, a, b, c, d, e, f, g\}$; $B = \{(r, a), (a, b), (r, c), (c, d), (d, e), (c, f), (f, g)\}$; $\varphi(r) = w_1$, $\varphi(a) = w_2$, $\varphi(b) = w_6$, $\varphi(c) = w_4$, $\varphi(d) = w_5$, $\varphi(e) = w_9$, $\varphi(f) = w_3$, $\varphi(g) = w_7$, and $\psi(r, a) = 1$, $\psi(a, b) = 1$ $\psi(r, c) = 2$, $\psi(c, d) = 1$, $\psi(d, e) = 1$, $\psi(c, f) = 2$, $\psi(f, g) = 1$ and similarly all the other submarkers M_i, M^*, M_i^* and also the factor-markers M_0 and M_0^* can be described by the enumeration of their different elements as it was done by M. Especially if $M^* = \langle A^*, r^*, B^*, \varphi^*, \psi^* \rangle$ where evidently $A^* = A$, $r^* = r$ and $B^* = B$, i.e. condition (11) is satisfied, then $\varphi^*(x) = \Phi(\varphi(x))$ for each $x \in A$, i.e. condition (12) is satisfied, because the labellings of the vertices in both markers M and M^* are the same. Finally $\psi^*(r, a) = 2 = \pi_1(\psi(r, a))$, $\psi^*(r, c) = 1 = \pi_1(\psi(r, c))$ and $\psi^*(c, d) = 2 = \pi_4(\psi(c, d))$, $\psi^*(c, f) = 1 = \pi_4(\psi(c, f))$ and in all other cases the prescribed permutations are identical, thus condition (13) is satisfied. Therefore $M^* = T(M)$.

One easily sees that $L(M_1) = \langle \mathrm{NP} \rangle \langle \mathrm{V_{act}} \rangle \langle \mathrm{NP} \rangle$ and $L(M_1^*) = \langle \mathrm{NP} \rangle \langle \mathrm{V_{pas}} \rangle \langle \mathrm{NP} \rangle$ and that Π for M_1 satisfies (14) when the mappings ν and ν^* are determined.

In this example it was not necessary to define $\Phi_i = \Phi$ for each $i = 1, 2, 3, 4$, i.e. to define Φ_i for each rule of $\mathfrak{R}$, because obviously e.g. it was sufficient to define Φ_1 only on the subset $\mathfrak{R}_1 \subset \mathfrak{R}$, where $\mathfrak{R}_1 = \{w_1, w_4\}$; Φ_2 only on the subset $\mathfrak{R}_2 = \{w_2, w_3, w_6, w_7, w_8\}$ etc. That means that Φ_i need not be necessarily homomorphisms of $\mathfrak{R}$ into $\mathfrak{R}^*$ but sometimes they can be homomorphisms of some special subsets $\mathfrak{R}_i$ of $\mathfrak{R}$. It is clear that the subset $\mathfrak{R}_i$ is always determined as the set of rules

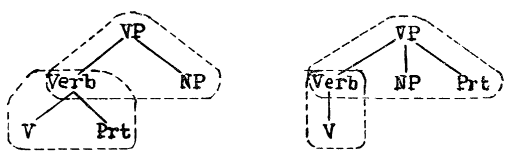

Fig. 7.

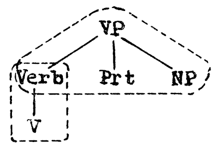

Fig. 8.

which are used in the submarker M_i for each marker $M \in \mathfrak{M}_{\mathscr{G}}$. This is a practical reason why it can be useful to consider a well transformation as a decomposition transformation and use the structural indices too.

There is another possibility how to define the canonical grammars generating the canonical languages $G(S)$ and $G^*(S^*)$. The following changes are necessary: the symbols "the" and "a" are transferred from V_A and V_A^* in V_P and V_P^*; a new non-terminal symbol $\langle \mathrm{Det} \rangle$ is added to V_N and V_N^* and instead of the rules w_2 and w_3, the rules $w_{12} = w_{12}^* = \langle \mathrm{NP} \rangle ::= \langle \mathrm{Det} \rangle \langle \mathrm{N} \rangle$, $w_{13} = w_{13}^* = \langle \mathrm{Det} \rangle ::= /\mathrm{the}$ and $w_{14} = w_{14}^* = \langle \mathrm{Det} \rangle ::= /\mathrm{a}$ are introduced.

These modified canonical grammars are weakly equivalent to the previous ones and again the former can be well transformed onto the latter. The forms of the considered particular sentences s and $t(s)$ are as follows: (e, e, e, s, e, e) and $(e, e, /is, /by, e, e)$. Therefore this descriptive transformation is a form transformation.

Example 2. Here particular examples are investigated of the singular transformations which are not well transformations and which are introduced by different linguists.

In Fig. 1a) and 4a) a pair of corresponding phrase-markers was shown in the simplest case when "the singulary transformation is a permutation". According to the basic definitions of well or decomposition transformations it is clear that the phrase-marker in Fig. 4a) can never be the image of Fig. 1a) in a well transformation. This impossibility is caused by the impossibility that the right side in Fig. 7 is the image of the left side in a well transformation, because the canonical lengths of the both rules in the left side are equal to 2 and on the right side one is equal to 3 and other the to 1.

There are no difficulties from the mathematical point of view to introduce more general mappings of phrase-markers than the well and decomposition transformations are. It is possible to decompose the phrase-marker (or the markers) into parts which need not be rules but some chosen types of sub-phrase markers and to prescribe directly the corresponding sub-phrase-marker (i.e. a new rooted tree and new labellings). The pair of phrase-markers in Fig. 7 can be an example of such general correspondence. We have not introduced these general transformations here because there is another way how to make possible to use the notion of decomposition transformation here.

Let us change the left side in Fig. 7 as it is shown in Fig. 8. Now it is clear that this modified sub-phrase marker can be mapped onto the right side of Fig. 7 by a well transformation. This modification means that in the underlying context-free grammar the rules ⟨VP⟩ : : = ⟨Verb⟩ ⟨NP⟩ and ⟨Verb⟩ : : = ⟨V⟩ ⟨Prt⟩ are substituted by new rules ⟨VP⟩ : : = ⟨Verb⟩ ⟨Prt⟩ ⟨NP⟩ and ⟨Verb⟩ : : = ⟨V⟩. We do not know any linguistic objections against this modification, because it seems to us that the linguistic interpretations of the nonterminal symbols ⟨VP⟩, ⟨Verb⟩ and ⟨V⟩ are not sufficiently distinguished and determined. Therefore ⟨Prt⟩ can be in "is a"-relation to ⟨VP⟩ directly instead of to ⟨Verb⟩.

Another pair of corresponding phrase-markers was shown in Fig. 2a) and 4b). Here too, some modifications as in the previous case would be necessary. We shall not analyse this case in detail because a possible handling with active—passive transformation is given in Example 1 and we do not believe that the shown phrase-markers have a definite and correct form. What we want to stress here is that again a convention concerning the rooted tree is silently assumed which must be added to the transformation rule in order to determine fully the new phrase-marker.

For this reason, in the underlying context-free grammar it would be more useful to use the rule ⟨S⟩ : : = ⟨NP⟩ ⟨VP⟩ ⟨NP⟩ (if we have in mind the transitive verbs) expressing that ⟨VP⟩ denotes a binary relation between the subject and the object, than the traditional rules ⟨S⟩ : : = ⟨NP⟩ ⟨VP⟩ and ⟨VP⟩ : : = ⟨Verb⟩ ⟨NP⟩. Similarly in the case of double transitive verbs the first rule should be ⟨S⟩ : : = ⟨NP⟩ ⟨VP⟩ ⟨NP⟩ ⟨NP⟩ etc. It must be considered why these oldest grammatical and logical concepts and schemas are to be used today when the logical analysis of language and the logic itself are developed more in detail.

Example 3. Let us consider two context-free grammars $G_i = \langle V_T, V_N, \mathfrak{R}_i, S\rangle$ and $G_i = \langle V_T^*, V_N^*, \mathfrak{R}_i^*, S^*\rangle$ (from the point of view of canonical grammars we suppose $V_A = V_T$, $V_P = \emptyset$). We shall say that G_i is *well* (or *decompositionally*) and *permutationally transformable onto* G_i^* if there exists a well (or decomposition) transformation T of G_i onto G_i^* which is determined by a non permutational (or permutational) homomorphism of $\mathfrak{R}_i$ onto $\mathfrak{R}_i^*$.

If $\mathfrak{R}_1 = \{$S : : = Sa, S : : = b$\}$ and $\mathfrak{R}_1^* = \{$S* : : aS*b, S* : : = c$\}$ where a, b, c are terminal symbols then G_1 is well transformable onto G_1^*.

In Fig. 9 two corresponding phrase-markers are shown which are evidently not isomorphic as the rooted trees (but the corresponding markers are isomorphic).

If $\mathfrak{R}_2 = \{S ::= AB,\ A ::= aA,\ A ::= b,\ B ::= c\}$ and $\mathfrak{R}_2^* = \{S^* ::= BA,$ $A :: aA,\ A ::= b,\ B ::= c\}$ where A, B are nonterminal and a, b, c are terminal symbols then G_2 is not well transformable onto G_2^* but G_2 is well permutationally transformable onto G_2^*. In Fig. 10 the two corresponding phrase-markers are shown.

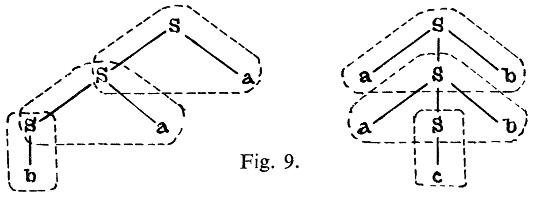

Fig. 9.

If $\mathfrak{R}_3 = \{S^* ::= S^*a,\ S^* ::= S^*c,\ S^* ::= b\}$ where a, b, c are terminal symbols then G_1 is not well transformable onto G_3 (because the cardinal number of $\mathfrak{R}_3$ is greater than that of $\mathfrak{R}_1$) but G_1 is decompositionally transformable onto G_3. The corresponding decomposition procedure $\mathscr{D}$ can be described as follows: the roots of the submarkers of $\mathscr{D}(M)$ are all the vertices of M which are labelled by the rule $S ::= Sa$ and in the known ordering of them the odd or even vertices will correspond to the rules $S^* ::= S^*a$ or $S^* ::= S^*c$ respectively.

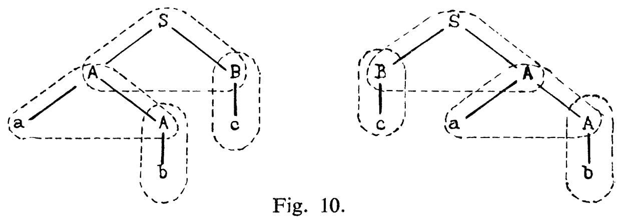

Fig. 10.

If $\mathfrak{R}_4 = \{S^* ::= S^*aR,\ S^* ::= S^*c,\ S^* ::= b,\ R ::= bR,\ R ::= c\}$ where R is nonterminal and a, b, c are terminal symbols then $\mathfrak{R}_1$ is not decompositionally and permutationally transformable onto $\mathfrak{R}_4$ (because in $\mathfrak{R}_4$ there is a rule having canonical length two but in $\mathfrak{R}_1$ each rule has canonical length at most one).

It is clear that with respect to the homomorphisms of the sets of context-free rules a classification of all context-free grammars is determined because a set $\mathfrak{R}$ of rules can be called *simple* if each homomorphism Φ of $\mathfrak{R}$ onto $\mathfrak{R}^*$ implies that Φ is an isomorphism. Thus to the same class belong the sets having the same homomorphism image. This classification induces a classification of context-free languages if we define that a context-free language L is *well* (or *decompositionally*) (and *permutationally) transformable onto* L^* if there are context-free grammars G and G^* generating L and L^* respectively such that G is well (or decompositionally) (and permutationally) transformable onto G^*.

Then $G_1(S) = \{ba^n;\ n \geqq 0\}$ is well transformable onto $G_1^*(S^*) = \{a^n cb^n;\ n \geqq 0\}$, but $G_1(S)$ is a regular event and $G_1^*(S^*)$ is not. Therefore the class of regular events is not preserved by the well transformations and to the same class of the mentioned classification the regular and also some nonregular languages will belong.

On the other hand, it is clear how to generalize the notion of homomorphism and then the notion of well transformation to the context-sensitive or quite general grammars and languages. Especially by the context-sensitive grammars the situation is very simple. Here it can be expected that these generalized transformations will transfer context-sensitive grammars on the context-free ones and that therefore they could be a suitable tool for study of context-sensitive grammars and languages.

Finally it is interesting to note that if we put $\mathfrak{R}_5 = \{\mathrm{S} :: = \mathrm{SaS}, \mathrm{S} :: = \mathrm{b}\}$ where a, b are terminal symbols then G_1 can not be well transformable onto G_5 but $G_1(S) = \{ba^n;\ n \geqq 0\}$ is well transformable onto $G_5(S) = \{b(ab)^n;\ n \geqq 0\}$ because this last language is generated by the following grammar $G_6 : \mathfrak{R}_6 = \{\mathrm{S} :: = \mathrm{Sab}, \mathrm{S} :: = \mathrm{b}\}$, where a, b are terminal symbols. Now it is clear that one of the two occurrences of the symbol S in the rule $\mathrm{S} :: = \mathrm{SaS}$ in $\mathfrak{R}_5$ is not important.

Bibliography

[1] *Chomsky N.:* On the notion "Rule of Grammar". Proc. Symp. Applied Math. *12* (1961).
[2] *Chomsky N.:* Three models for the description of language. I.R.E. Trans. *2* (1956), 113—124.
[3] *Čulík K.:* On the description of sentence structure (Czech.) Studies in applied linguistics of N. P. Andrejev. Prague 1963, 98—109.
[4] *Čulík K.:* Applications of graph theory to mathematical logic and linguistics. Theory of graphs and its applications, Publ. House of Czech. Academy of Sciences, Prague 1964, 13—20.
[5] *Harris Z.:* Co-occurrence and Transformation in Linguistic Structure, Language *33* (1957), 283—340.
[6] *Čulík K.:* Well-translatable grammars and ALGOL-like languages. Formal Language Description Languages for Computer Programming, North-Holland, Amsterdam 1965.
[7] *Backus J. W.* et all.: Report on the algorithmic language ALGOL 60. Num. Math. *2* (1960), 106—136.
[8] *Chomsky N.:* The logical basis of linguistic theory. Proc. of the 9th international congress of linguists, Monton, Haag 1964, 914—978.
[9] *Čulík K.:* Semantics and Translation of Grammars and ALGOL — like Languages. Prague, Cybernetics *1* (1965), 47—49.
[10] *Chomsky N.:* Chapters 11—13 in the Handbook of Mathematical Psychology, II., Wiley, New York 1963.
[11] *Chomsky N., Schützenberger M. P.:* The algebraic theory of context-free languages, Computer Programming and Formal Systems, North-Holland, Amsterdam 1963.
[12] *Bach E.:* An introduction to transformational grammars. Holt, Rinehart and Winston, 1964.
[13] *Chomsky N.:* Aspects of the theory of syntax. The M.I.T. Press, Cambridge 1965.

Author's address: Praha 1, Žitná 25, ČSSR (Matematický ústav ČSAV).

Резюме

О НЕКОТОРЫХ ТРАНСФОРМАЦИЯХ КОНТЕКСТНО-СВОБОДНЫХ ЯЗЫКОВ

КАРЕЛ ЧУЛИК (Karel Čulík), Прага

В статье вводится и изучается с учетом требований лингвистов, особенно Н. Чомского (N. Chomsky), предъявляемых к сингулярным грамматическим трансформациям — довольно общий класс отображений, называемых „хорошими трансформациями". Последние сопоставляют фразовым показателям любой грамматики типа 2 опять фразовые показатели другой такой грамматики, определенные однозначно некоторым гомоморфизмам певрой грамматики во вторую. Под гомоморфизмом следует понимать отображение сопостявляющее правилам грамматики опять правила, следовательно не отображение алфавитов (как его вводит M. P. Schützenberger).).

Специальные, строгие условия, налагаемые на гомоморфизм, касаются разбиения множества терминальных символов на собственные и вспомогательные и отношения применимости одного правила к другому. Фразовый показатель определяется математически как некоторый двойной граф; в статье используется равносильное понятие показателя. Показатель это корневое дерево, узлы которого помечены правилами и ребя натуральными числами. Естественным образом вводится понятие первого и последнего слова по отношению к данному показателю. Если выполнены некоторые дополнительные условия, индуцируется трансформация языка, порожденного в язык, порожденный второй. Особую важность приобретают перестановочные гомоморфизмы и с их помощью определенные трансформации. Они позволяют проводить некоторые фиксированные перестановки собственных, терминальных и вспомогательных символов в правилах. На примере показана возможность отобразить с помощью хорошей трансформации регулярное событие на нерегулярный язык. Изучаются некоторые дополнительные проблемы.