Miroslav Novotný
On a class of languages

Archivum Mathematicum, Vol. 6 (1970), No. 3, 155--170

Persistent URL: http://dml.cz/dmlcz/104719

ON A CLASS OF LANGUAGES

MIROSLAV NOVOTNÝ, Brno

(Received May 6, 1969)

INTRODUCTION

In [1] I defined, for every language, configurations of order 1, 2, ... By means of these configurations we can define the so called generalized configurational grammar for every given language. This generalized configurational grammar generates the given language. A language is called finitely characterizable if the set of its strings which contain no configuration is finite and if the set of its so called simple configurations is finite, i.e. if its generalized configurational grammar is a grammar. I studied the class of all finitely characterizable languages and compared them with the well-known classes of languages of the Chomsky's classification.

Two other definitions of a configuration appeared in the literature before ([2], [3]) and similar theories of configurational grammars and finitely characterizable languages were constructed ([3], [4]). Thus, we have three different possibilities for the definition of a configuration and each of them can be formulated either in a weak form or in a strong one. The definition of a strong configuration of order 1 is the same in all these theories[1]).

In the present paper we study languages for which the set of strings which contain no strong configurations of order 1 is finite and for which the set of all simple strong configurations of order 1 is finite. Such languages will be called languages of strong depth 1. These languages form a subclass of the class of all context-free languages. The class of all context-free languages can be built up on the basis of the class of all languages of strong depth 1: Every context-free language is the intersection of a so called full language and a trace of a language of strong depth 1; the trace of a given language is defined to be a language which we obtain by cancelling all symbols which do not belong to a given set in all strings of the given language. In this way we have obtained a new characterization of context-free languages.

[1]) The terminology of these papers does not coincide. In [1], "configuration" stands for "weak configuration", in [2], [3] "configuration" stands for "strong configuration". In [4], "configuration" stands for "weak configuration" and the definition of a strong configuration corresponds to the definition of a strong configuration introduced in the present paper.

1. GENERALIZED GRAMMARS AND GRAMMARS

If V is a set, we denote by V^* the free monoid over V, i.e. the set of all finite sequences of elements of V in which the operation of concatenation is defined; we suppose that the empty sequence Λ is an element of V^*, too. We identify one-element-sequences with elements of V; thus, we have $V \subseteq V^*$ and, for every pair of non-negative integers k and l, $k \leqq l$, and for $x_k, x_{k+1}, \ldots, x_l \in V$, we write $x_k x_{k+1} \ldots x_l$ or $\prod_{i=k}^{l} x_i$ instead of $(x_k, x_{k+1}, \ldots, x_l)$. It is advantageous to define $\prod_{i=k}^{l} x_i = \Lambda$ if $0 \leqq l < k$ are integers. Thus, we use the symbol $x_k x_{k+1} \ldots x_l = \prod_{i=k}^{l} x_i$ for all pairs of non-negative integers k, l.

The elements of V are called symbols, the elements of V^* strings.

We put $|\Lambda| = 0$. If $x \in V^*$, $x = x_1 x_2 \ldots x_n$ where n is a natural number and $x_i \in V$ for $i = 1, 2, \ldots, n$, then we put $|x| = n$.

If n is a natural number and $A_i \subseteq V^*$ for $i = 1, 2, \ldots, n$, then we denote by $A_1 A_2 \ldots A_n$ the set

$$\{a_1 a_2 \ldots a_n; \quad a_i \in A_i, \qquad i = 1, 2, \ldots, n\}.$$

1.1. Definition. Let V, U be sets, let f be a mapping of the set V into U^*. We put $f_*(\Lambda) = \Lambda$; if $x = x_1 x_2 \ldots x_n$ where n is a natural number and $x_i \in V$ for $i = 1, 2, \ldots, n$, then we put $f_*(x) = f(x_1) f(x_2) \ldots \ldots f(x_n)$. For $S \subseteq V^*$, we define $f_*(S) = \{f_*(x); x \in S\}$.

1.2. Remark. If V, U are sets and f a mapping of V into U^*, then $f_*(xy) = f_*(x) f_*(y)$ for every $x, y \in V^*$.

1.3. Definition. Let V be a set, $L \subseteq V^*$; then the pair (V, L) is called a *language*.

1.4. Definition. Let V be a set. Then the language (V, V^*) is called *full*.

1.5. Definition. Let (V, L), (U, M) be languages. Then the language $(V \cap U, L \cap M)$ is called the *intersection of the languages* (V, L), (U, M).

1.6. Definition. Let V, V_T, S, R be sets with the properties $V_T \subseteq V$, $S \subseteq V^*$, $R \subseteq V^* \times V^*$. Then the quadruple $G = \langle V, V_T, S, R\rangle$ is called a *generalized grammar*.

1.7. Definition. Let $G = \langle V, V_T, S, R\rangle$ be a generalized grammar. We write, for $x, y \in V^*$, $x \to y$ (G) instead of $(x, y) \in R$. For $x, y \in V^*$, we write $x \Rightarrow y$ (G) iff there exist such strings $u, v, t, z \in V^*$ that $x = utv$, $uzv = y$, $t \to z$ (G). For $x, y \in V^*$ we write $x \overset{*}{\Rightarrow} y$ (G) iff there exists a non-negative integer p and some strings $t_0, t_1, \ldots, t_p$ of V^* such that $x = t_0$, $t_p = y$ and $t_{i-1} \Rightarrow t_i(G)$ for $i = 1, 2, \ldots, p$. The sequence $t_0, t_1, \ldots, t_p$ is called an *x-derivation* of y in G. We put $\mathscr{L}(G) =$

$= \{x; x \in V_T^*$ and there exists some $s \in S$ with the property $s \overset{*}{\Rightarrow} x\,(G)\}$. The language $(V_T, \mathscr{L}(G))$ is called the *language generated by G*.

1.8 Definition. Let $G = \langle V, V_T, S, R\rangle$ be a generalized grammar. If $V = V_T$, then this generalized grammar is called a *special generalized grammar*. We write $\langle V, S, R\rangle$ instead of $\langle V, V, S, R\rangle$ if $\langle V, V, S, R\rangle$ is a special generalized grammar.

1.9. Definition. Let $G = \langle V, V_T, S, R\rangle$ be a generalized grammar. Then G is called a *grammar* iff the sets V, S, R are finite.

1.10. Remark. From the above definitions it is clear what by a special grammar is meant.

1.11. Definition. Let $G = \langle V, V_T, S, R\rangle$ be a grammar with the following properties: (1) There exists such an element $\sigma \in V - V_T$ that $S = \{\sigma\}$. (2) For each $(x, y) \in R$ it holds true $x \in V - V_T$. Then G is called a *context-free grammar in the usual sense.*

1.12. Definition. Let $G = \langle V, S, R\rangle$ be such a special grammar that $(x, y) \in R$ implies $x \in V$. Then G is called a *special context-free grammar.*

1.13. Lemma. *Let $G = \langle V, S, R\rangle$ be a special context-free grammar, n a natural number, $a_1, a_2, \ldots, a_n$ elements of V, $y \in V^*$ such a string that $a_1a_2 \ldots a_n \overset{*}{\Rightarrow} y\,(G)$. Then there exist such strings $y_1, y_2, \ldots, y_n$ in V^* that $a_i \overset{*}{\Rightarrow} y_i(G)$ for $i = 1, 2, \ldots, n$ and $y_1y_2 \ldots y_n = y$.*

This lemma is well known.

1.14. Definition. Let (V, L) be a language. This language is called a *special context-free language* iff there exists a special context-free grammar generating (V, L).

1.15. Definition. A language is called *context free* iff it is the intersection of a special context-free language and a full language.

1.16. Remark. Our definition of a context-free language differs only formally from the usual one.[2])

Usually, a context-free language is defined as a language generated by a context-free grammar $G = \langle V, V_T, \{\sigma\}, R\rangle$ in the usual sense. If we put $H = \langle V, \{\sigma\}, R\rangle$ then H is a special context-free grammar with the property $(V_T, \mathscr{L}(G)) = (V \cap V_T, \mathscr{L}(H) \cap V_T^*)$. Thus, the context-free language in the usual sense $(V_T, \mathscr{L}(G))$ is the intersection of the special context-free language $(V, \mathscr{L}(H))$ and the full language (V_T, V_T^*) and is context-free in our sense.

If (V, L) is a context-free language in our sense, then there exist a special context-free language (U, M) and a full language (W, W^*) such that $V = U \cap W$ and $L = M \cap W^*$. Clearly, we can suppose $W \subseteq U$, thus $V = W \subseteq U$. According to 1.14 there exists a special context-free grammar $G = \langle U, S\ R\rangle$ such that $\mathscr{L}(G) = M$. We take a set U' which is equivalent to U and a bijection b of U onto U'; we

[2]) See [5], p. 10.

suppose $U \cap U' = \emptyset$. Let σ be such an element that $\sigma \notin U \cup U'$. We put $R' = \{(\sigma, b_*(s)); s \in S\} \cup \{(b(x), b_*(y)); (x,y) \in R\} \cup \{(b(x), x); x \in U\}$ and we define $H = \langle U \cup U' \cup \{\sigma\}, V, \{\sigma\}, R'\rangle$. Clearly, H is a context-free grammar in the usual sense and $\mathscr{L}(H) = \mathscr{L}(G) \cap V^* = M \cap V^* = L$. Thus, the language generated by H is (V, L) and (V, L) is a context-free language in the usual sense.

in **1.17. Remark.** Let $G = \langle V, V_T, \{\sigma\}, R\rangle$ be a context-free grammar t the usual sense. Then we can suppose, without loss of generality, hat G has the following property:

(P) If $(x, y) \in R$, then there exist such strings $u, v \in V^*$ that $\sigma \overset{*}{\Rightarrow} uxv$ (G).

Indeed, if $G = \langle V, V_T, \{\sigma\}, R\rangle$ is a context-free grammar that has not the property (P), then there exists a pair $(x, y) \in R$ such that there exist no strings $u, v \in V^*$ with the property $\sigma \overset{*}{\Rightarrow} uxv$ (G). Clearly, $\mathscr{L}(G) = \mathscr{L}(\langle V, V_T, \{\sigma\}, R - \{(x, y)\}\rangle)$. After a finite number of such steps we obtain a context-free grammar $H = \langle V, V_T, \{\sigma\}, R_1\rangle$ with the property (P) such that $\mathscr{L}(H) = \mathscr{L}(G)$. Thus G and H generate the same language. (Compare [5], p. 19, Lemma 1.4.2.)

2. TRACES OF LANGUAGES AND GENERALIZED GRAMMARS

2.1. Definition. Let V, U be sets. For each $v \in V$ we put $\mathbf{t}^U(v) = v$ if $v \in U$ and $\mathbf{t}^U(v) = \Lambda$ if $v \in V - U$. According to 1.1 we define the mapping $\mathbf{t}^U_*$ of V^* into U^*. If $x \in V^*$ is a string, then $\mathbf{t}^U_*(x)$ is called the *trace of x in U^**.

2.2. Lemma. *Let V, U be sets. Then the mapping $\mathbf{t}^U_*$ has the following properties:*

(A) *For each $x, y \in V^*$ it holds true $\mathbf{t}^U_*(xy) = \mathbf{t}^U_*(x)\, \mathbf{t}^U_*(y)$.*

(B) *If $\mathbf{t}^U_*(u) = x'y'$ for some $u \in V^*$, $x', y' \in U^*$, then there exist such strings $x, y \in V^*$ that $\mathbf{t}^U_*(x) = x'$, $\mathbf{t}^U_*(y) = y'$, $xy = u$.*

(C) *For each $x \in V^*$ we have $\mathbf{t}^U_*(\mathbf{t}^U_*(x)) = \mathbf{t}^U_*(x)$.*

Proof. 1. (A) is a special case of 1.2.

2. Let us have $\mathbf{t}^U_*(u) = x'y'$ for some $u \in V^*$, $x', y' \in U^*$. If $u = \Lambda$, then we take $x = \Lambda = y$ and we have $\mathbf{t}^U_*(x) = \Lambda = x'$, $\mathbf{t}^U_*(y) = \Lambda = y'$, $xy = \Lambda = u$. Let us suppose $u \neq \Lambda$; then there exist a natural number n and some elements $u_1, u_2, \ldots, u_n$ of V such that $u = u_1u_2 \ldots u_n$. Thus, $\mathbf{t}^U(u_1)\, \mathbf{t}^U(u_2) \ldots \mathbf{t}^U(u_n) = \mathbf{t}^U_*(u) = x'y'$. Thus, such a natural number m, $1 \leqq m \leqq n + 1$, exists that $\mathbf{t}^U(u_1) \ldots \mathbf{t}^U(u_{m-1}) = x'$, $\mathbf{t}^U(u_m) \ldots \mathbf{t}^U(u_n) = y'$. We put $x = u_1 \ldots u_{m-1}$, $y = u_m \ldots u_n$. Clearly, $\mathbf{t}^U_*(x) = x'$, $\mathbf{t}^U_*(y) = y'$ and $xy = u$. Thus, (B) holds true.

3. If $x = \Lambda$, then $\mathbf{t}^U_*(\mathbf{t}^U_*(x)) = \mathbf{t}^U_*(\mathbf{t}^U_*(\Lambda)) = \mathbf{t}^U_*(\Lambda) = \mathbf{t}^U_*(x)$. Let us suppose $x \in V^*$, $x \neq \Lambda$. Then there exist such a natural number n and such elements $x_1, x_2, \ldots, x_n \in V$ that $x = x_1x_2 \ldots x_n$. Therefore, $\mathbf{t}^U_*(x) =$

$= \mathbf{t}^U(x_1)\, \mathbf{t}^U(x_2) \ldots \mathbf{t}^U(x_n)$ and $\mathbf{t}^U_*(\mathbf{t}^U_*(x)) = \mathbf{t}^U_*(\mathbf{t}^U(x_1))\, \mathbf{t}^U_*(\mathbf{t}^U(x_2)) \ldots \mathbf{t}^U_*(\mathbf{t}^U(x_n))$ according to (A). If $\mathbf{t}^U(x_i) = \Lambda$, then $\mathbf{t}^U_*(\mathbf{t}^U(x_i)) = \Lambda = \mathbf{t}^U(x_i)$, if $\mathbf{t}^U(x_i) = x_i$, then $\mathbf{t}^U_*(\mathbf{t}^U(x_i)) = \mathbf{t}^U_*(x_i) = \mathbf{t}^U(x_i)$. Thus, $\mathbf{t}^U_*(\mathbf{t}^U_*(x)) = \mathbf{t}^U(x_1)\, \mathbf{t}^U(x_2) \ldots \mathbf{t}^U(x_n) = \mathbf{t}^U_*(x)$.

We have proved (C).

2.3. Definition. Let (V, L) be a language, U a set. We put $\mathbf{t}^U_*(L) = \{\mathbf{t}^U_*(x);\ x \in L\}$; the language $(U, \mathbf{t}^U_*(L))$ is called the *trace of the language* (V, L) *in* U^*.

2.4. Definition. Let $G = \langle V, S, R\rangle$ be a special generalized grammar, U a set. We put $S_1 = \mathbf{t}^U_*(S)$, $R_1 = \{(\mathbf{t}^U_*(x), \mathbf{t}^U_*(y));\ (x, y) \in R\}$. Then the special generalized grammar $\langle U, S_1, R_1\rangle$ is called the *trace of the special generalized grammar* G *in* U^*.

2.5. Theorem. *Let $G = \langle V, S, R\rangle$ be a special generalized grammar, U a set with the property that $(x, y) \in R$ implies $x \in U$. Let $H = \langle U, S_1, R_1\rangle$ be the trace of G in U^*. Then $(U, \mathscr{L}(H))$ is the trace of the language $(V, \mathscr{L}(G))$ in U^*.*

Proof. 1. Let $x \in \mathscr{L}(G)$. Then there exist an element $s \in S$, a non-negative integer p and such elements $s_0, s_1, \ldots, s_p$ in V^* that $s = s_0$, $s_p = x$ and $s_{i-1} \Rightarrow s_i(G)$ for $i = 1, 2, \ldots, p$. We prove by induction with respect to i that $\mathbf{t}^U_*(s_i) \in \mathscr{L}(H)$ for $i = 0, 1, \ldots, p$.

We denote by $C(n)$ the following assertion: $\mathbf{t}^U_*(s_n) \in \mathscr{L}(H)$.

Then $C(0)$ holds true trivially as we have $s_0 \in S$ and $\mathbf{t}^U_*(s_0) \in \mathbf{t}^U_*(S) = S_1 \subseteq \mathscr{L}(H)$.

Let m be an integer, $0 < m \leqq p$. We prove that $C(m - 1)$ implies $C(m)$. Indeed, $C(m - 1)$ means $\mathbf{t}^U_*(s_{m-1}) \in \mathscr{L}(H)$. We have $s_{m-1} \Rightarrow s_m(G)$, i.e. there exist some strings $u, v, t, z \in V^*$ with the properties $s_{m-1} = utv$, $uzv = s_m$, $t \to z\ (G)$. According to 2.2 (A) we have $\mathbf{t}^U_*(s_{m-1}) = \mathbf{t}^U_*(u)\, \mathbf{t}^U_*(t)\, \mathbf{t}^U_*(v)$, $\mathbf{t}^U_*(s_m) = \mathbf{t}^U_*(u)\, \mathbf{t}^U_*(z)\, \mathbf{t}^U_*(v)$ and according to 2.4 it holds true $(\mathbf{t}^U_*(t), \mathbf{t}^U_*(z)) \in R_1$. Thus, $\mathbf{t}^U_*(s_{m-1}) \Rightarrow \mathbf{t}^U_*(s_m)\ (H)$ and $\mathbf{t}^U_*(s_m) \in \mathscr{L}(H)$ which is $C(m)$.

It follows that $C(m)$ holds true for $m = 0, 1, 2, \ldots, p$. Especially, we have $\mathbf{t}^U_*(x) = \mathbf{t}^U_*(s_p) \in \mathscr{L}(H)$. Thus we have proved that $\mathbf{t}^U_*(\mathscr{L}(G)) \subseteq \mathscr{L}(H)$.

2. Let us suppose $x' \in \mathscr{L}(H)$. Then there exist an element $s' \in S_1 = \mathbf{t}^U_*(S)$, a non-negative integer p and some elements $s'_0, s'_1, \ldots, s'_p$ in U^* such that $s' = s'_0$, $s'_p = x'$ and $s'_{i-1} \Rightarrow s'_i\ (H)$ for $i = 1, 2, \ldots, p$. By induction with respect to i we prove that there exist such elements $s_0, s_1, \ldots, s_p$ in $\mathscr{L}(G)$ that $\mathbf{t}^U_*(s_i) = s'_i$ for $i = 0, 1, \ldots, p$.

By $D(n)$ we denote the following assertion: There exists such an element $s_n \in \mathscr{L}(G)$ that $\mathbf{t}^U_*(s_n) = s'_n$.

Then $D(0)$ holds true trivially as $s'_0 \in \mathbf{t}^U_*(S)$ which implies the existence of an element $s_0 \in S \subseteq \mathscr{L}(G)$ with the property $s'_0 = \mathbf{t}^U_*(s_0)$.

Let m be an integer, $0 < m \leqq p$. We prove that $D(m - 1)$ implies

$D(m)$. Indeed, $D(m-1)$ means the existence of an element $s_{m-1} \in \mathscr{L}(G)$ with the property $\mathbf{t}^U_*(s_{m-1}) = s'_{m-1}$. We have $s'_{m-1} \Rightarrow s'_m(H)$, i.e. there exist such strings $u', v', t', z' \in U^*$ that $s'_{m-1} = u't'v'$, $u'z'v' = s'_m$, $t' \to z'(H)$. It follows $(t', z') \in R_1$. Thus, some strings $t, z \in V^*$ exist with the properties $(t, z) \in R$, $\mathbf{t}^U_*(t) = t'$, $\mathbf{t}^U_*(z) = z'$. From $(t, z) \in R$ it follows $t \in U$ which implies $t' = \mathbf{t}^U_*(t) = t$.

According to 2.2 (B) such strings $u, v, w \in V^*$ exist that $\mathbf{t}^U_*(u) = u'$, $\mathbf{t}^U_*(w) = t'$, $\mathbf{t}^U_*(v) = v'$, $uwv = s_{m-1}$. We have $\mathbf{t}^U_*(w) = t' = t \in U$ which implies the existence of strings $a, b \in V^*$ such that $w = atb$. We have $\mathbf{t}^U_*(w) = \mathbf{t}^U_*(atb) = \mathbf{t}^U_*(a)\,\mathbf{t}^U_*(t)\,\mathbf{t}^U_*(b) = \mathbf{t}^U_*(a)\,\mathbf{t}^U_*(\mathbf{t}^U_*(w))\,\mathbf{t}^U_*(b) = \mathbf{t}^U_*(a)\mathbf{t}^U_*(w)\mathbf{t}^U_*(b)$ according to 2.2 (A), (C) which implies $\mathbf{t}^U_*(a) = \Lambda = \mathbf{t}^U_*(b)$. We put $u_1 = ua$, $v_1 = bv$. Then $s_{m-1} = uwv = uatbv = u_1tv_1$. We put $s_m = u_1zv_1 = uazbv$. Thus, $s_{m-1} \Rightarrow s_m(G)$ and $s_m \in \mathscr{L}(G)$. Further we have $\mathbf{t}^U_*(s_m) = \mathbf{t}^U_*(u)\,\mathbf{t}^U_*(a)\,\mathbf{t}^U_*(z)\,\mathbf{t}^U_*(b)\,\mathbf{t}^U_*(v) = u'z'v' = s'_m$.

Thus, $D(m-1)$ implies $D(m)$ if $0 < m \leqq p$.

It follows that $D(m)$ holds true for $m = 0, 1, 2, \ldots, p$. Especially, there exists an element $s_p \in \mathscr{L}(G)$ with the property $\mathbf{t}^U_*(s_p) = s'_p = x'$.

Thus, we have proved $\mathscr{L}(H) \subseteq \mathbf{t}^U_*(\mathscr{L}(G))$.

3. We have $\mathscr{L}(H) = \mathbf{t}^U_*(\mathscr{L}(G))$. It follows that $(U, \mathscr{L}(H))$ is the trace of $(V, \mathscr{L}(G))$ in U^*.

3. LANGUAGES OF STRONG DEPTH 1

In the following definitions we denote by (V, L) an arbitrary language.

3.1. Definition. For $x \in V^*$ we put $x\nu(V, L)$ iff there exist some strings $u, v \in V^*$ with the property $uxv \in L$.

3.2. Definition. For $x, y \in V^*$ we put $x > y\ (V, L)$ iff $uxv \in L$ implies $uyv \in L$ for every $u, v \in V^*$.

3.3. Definition. For $x, y \in V^*$ we put $x \equiv y\ (V, L)$ iff $x > y\ (V, L)$ and $y > x(V, L)$.[3])

3.4. Definition. Let us suppose $x, y \in V^*$. The string x is called a *strong configuration with the result* y iff the following conditions are fulfilled: $x\nu(V, L)$, $x \equiv y\ (V, L)$, $1 = |y| < |x|$. By $C(V, L)$ we denote the set of all strong configurations of the language (V, L) and we put $A(V, L) = L - V^*C(V, L)\,V^*$. Further we put $E(V, L) = \{(y, x);\ x \in C(V, L),\ y \text{ a result of } x\}$.[4])

---

[3]) ν is a unary relation, $>$, $\equiv$ are binary relations on V^* which depend on (V, L); thus, we ought to have denoted them by $\nu_{(V,L)}$, $>_{(V,L)}$, $\equiv_{(V,L)}$, respectively. The symbols are complicated from the typographical point of view. We have thus preferred a simpler but less consequent way of notation in 3.1, 3.2 and 3.3.

[4]) We have defined strong configurations of order 1 if using the terminology of [4]. As we deal only with these strong configurations of order 1 here we call them—for the sake of brevity—strong configurations.

3.5. Definition. Let $x \in C(V, L)$. Then x is called a *simple* strong configuration iff, for each strings $u, v \in V^*$, $x' \in C(V, L)$, the condition $x = ux'v$ implies $u = \Lambda = v$. We denote by $P(V, L)$ the set of all simple strong configurations of the language (V, L), by $F(V, L)$ the set of all ordered pairs (y, x) where $x \in P(V, L)$ and y is a result of x.

3.6. Lemma. *Let (V, L) be a language, $x \in C(V, L)$. Then there exist such strings $u, v \in V^*$, $x' \in P(V, L)$ that $x = ux'v$.*

Proof. There exist such strings $u, v \in V^*$, $y \in C(V, L)$ that $x = uyv$ (it suffices to put $u = \Lambda = v$, $y = x$). We take such strings u, v, y with this property, for which $|y|$ is minimal. Clearly, $y \in P(V, L)$.

3.7. Definition. Let (V,L) be a language. We put $K(V,L) = \langle V, A(V, L), F(V, L)\rangle$. The triple $\langle V, A(V, L), F(V, L)\rangle$ is called the *generalized strong configurational grammar of depth 1.*

3.8. Theorem. *Let (V, L) be a language, $K(V, L)$ its generalized strong configurational grammar of depth* 1. *Then (V, L) is the language generated by $K(V, L)$.*

Proof. 1. By induction with respect to $|x|$ we prove that $x \in L$ implies $x \in \mathscr{L}(K(V, L))$.

By $E(n)$ we denote the following assertion: If $x \in L$ and $|x| = n$, then $x \in \mathscr{L}(K(V, L))$.

Then $E(0)$ holds true as $x \in L$, $|x| = 0$ implies $x \in A(V, L) \subseteq \subseteq \mathscr{L}(K(V, L))$.

Let $m > 0$ be an integer. We prove that the validity of $E(0), E(1), \ldots, E(m-1)$ implies the validity of $E(m)$.

Indeed, let us have $x \in L$, $|x| = m$.

If $x \in A(V, L)$, then $x \in \mathscr{L}(K(V, L))$.

Let us suppose $x \notin A(V, L)$. Then $x \in V^*C(V, L)\, V^*$. According to 3.6 we can suppose the existence of such strings $u, v \in V^*$, $z \in P(V, L)$ that $x = uzv$. Let t be a result of z. Then $x \in L$, $z \equiv t\ (V, L)$ imply $utv \in L$. We have $|t| < |z|$ which implies $|utv| < |x| = m$. As $E(0), E(1), \ldots, E(m-1)$ are valid, then $utv \in \mathscr{L}(K(V, L))$. Thus, such a string $s \in A(V, L)$ exists that $s \stackrel{*}{\Rightarrow} utv\ (K(V, L))$. We have $(t, z) \in F(V, L)$, i.e. $t \rightarrow z\ (K(V, L))$. Therefore, $utv \Rightarrow uzv\ (K(V, L))$ and $s \stackrel{*}{\Rightarrow} uzv\ (K(V, L))$, $x = uzv$. Thus, $x \in \mathscr{L}(K(V, L))$.

We have proved that the validity of $E(0), E(1), \ldots, E(m-1)$ implies the validity of $E(m)$. Thus, $E(m)$ holds true for $m = 0, 1, 2, \ldots$.

Therefore, $L \subseteq \mathscr{L}(K(V, L))$.

2. We prove by induction with respect to $|x|$ that $x \in \mathscr{L}(K(V, L))$ implies $x \in L$.

By $F(n)$ we denote the following assertion: If $x \in \mathscr{L}(K(V, L))$ and $|x| = n$, then $x \in L$.

Then $F(0)$ holds true as $x \in \mathscr{L}(K(V, L))$, $|x| = 0$ implies $x \in A(V, L) \subseteq \subseteq L$.

Let $m > 0$ be an integer. We prove that the validity of $F(0)$, $F(1)$, ..., $F(m - 1)$ implies the validity of $F(m)$.

Indeed, let us have $x \in \mathscr{L}(K(V, L))$, $|x| = m$.

If $x \in A(V, L)$, then $x \in L$.

Let us have $x \notin A(V, L)$. Then $x \in \mathscr{L}(K(V, L)) - A(V, L)$. Thus, such a natural number p and such strings $t_0, t_1, \ldots, t_p$ in V^* exist that $t_0 \in A(V, L)$, $t_p = x$ and $t_{i-1} \Rightarrow t_i\ (K(V, L))$ for $i = 1, 2, \ldots, p$. Especially we have $t_{p-1} \Rightarrow x\ (K(V, L))$ which means the existence of a pair $(t, z) \in \in F(V, L)$ and of such strings $u, v \in V^*$ that $t_{p-1} = utv$, $uzv = x$. It implies $|t| < |z|$ and $|t_{p-1}| = |utv| < |uzv| = |x| = m$. We have $t_{p-1} \in \mathscr{L}(K(V, L))$ which implies, as $F(0)$, $F(1)$, ..., $F(m - 1)$ are valid, $t_{p-1} \in L$. Further we have $t \equiv z\ (V, L)$ and $utv = t_{p-1} \in L$ which implies $x = uzv \in L$.

We have proved that the validity of $F(0)$, $F(1)$, ..., $F(m - 1)$ implies the validity of $F(m)$. Thus, $F(m)$ holds true for $m = 0, 1, 2, \ldots$.

Therefore, $\mathscr{L}(K(V, L)) \subseteq L$.

3. We have $\mathscr{L}(K(V, L)) = L$. Thus, the language generated by $K(V, L)$ is $(V, \mathscr{L}(K(V, L))) = (V, L)$.

3.9. Definition Let (V, L) be a language, U a set. This set is called *essential with respect to* (V, L) iff $(x, y) \in F(V, L)$ implies $x \in U$.

3.10. Definition. Let (V, L) be a language. This language is called a *language of strong depth 1* iff the sets V, $A(V, L)$, $P(V, L)$ are finite.

3.11. Lemma. *Let (V, L) be a language. Then (V, L) is a language of strong depth 1 iff the sets V, $A(V, L)$, $F(V, L)$ are finite.*

Proof. If the language (V, L) is a language of strong depth 1, then the set $P(V, L)$ is finite. As every $x \in P(V, L)$ has only a finite number of results y, which follows from the fact that $|y| = 1$ and that V is finite, the set $F(V, L)$ is finite. — If the sets V, $A(V, L)$, $F(V, L)$ are finite, then $P(V, L)$ is finite, too, and (V, L) is a language of strong depth 1.

3.12. Corollary. *Let (V, L) be a language. Then (V, L) is a language of strong depth* 1 *iff $K(V, L)$ is a special grammar.*

3.13. Theorem. *Every language of strong depth 1 is a special context-free language.*

Proof. If (V, L) is a language of strong depth 1, then $K(V, L)$ is a special grammar according to 3.12 and this grammar is context free. According to 3.8 (V, L) is generated by $K(V, L)$ and is therefore a special context-free language.

4. SOME PROPERTIES OF SPECIAL CONTEXT-FREE GRAMMARS

4.1. Lemma. *Let $G = \langle V, S, R\rangle$ be a special context-free grammar. Let Z_1, Z_2 be sets with the following properties: there exist a bijection f_1*

of R onto Z_1, a bijection f_2 of R onto Z_2 and the sets V, Z_1, Z_2 are mutually disjoint. We put $f_1(r) = [_r$, $f_2(r) =]_r$ for every $r \in R$, $V_1 = V \cup Z_1 \cup Z_2$, $R_1 = \{(x, [_r y]_r); (x, y) = r \in R\}$, $G_1 = \langle V_1, S, R_1\rangle$.[5]

Then the following assertions hold true:

(i) Let $s \in V_1$, $x \in V_1^*$ be such elements that $s \overset{*}{\Rightarrow} x\ (G_1)$. Then $|s| \leqq |x|$. If $|x| = 1$, then $s = x$; if *$|x| = 2$, then $s \in V$ and $s \to x\ (G_1)$.*

(ii) Let u, v, a, $b \in V_1^*$, $(t, z) \in R_1$, $(x, y) \in R_1$ be such elements that $uyv = azb$. Then either $u = azb_1$, $b_1b_2 = b$, $b_2 = yv$ for suitable strings $b_1, b_2 \in V_1^*$ or $uy = a_1$, $a_1a_2 = a$, $a_2zb = v$ for suitable strings a_1, $a_2 \in V_1^*$ or $u = a$, $y = z$, $v = b$.

(iii) Let $u, v \in V_1^*$, $(x, y) \in R_1$, $s \in V_1$ be such elements that $s \overset{*}{\Rightarrow} uyv\ (G_1)$. Then $s \overset{*}{\Rightarrow} uxv\ (G_1)$.

(iv) Let $s \in V_1$, $x \in V_1^*$ be such elements that $s_0, s_1, \ldots, s_p$ is an s-derivation of x in G_1 with the property $p \geqq 1$. Then the first [last] symbols of s_1 and s_i are the same for $i = 1, 2, \ldots, p$.

(v) Let $s \in V_1$, $u, v, x, y \in V_1^*$ be such elements that $s \overset{*}{\Rightarrow} uxv\ (G_1)$, $s \overset{*}{\Rightarrow}$ $\overset{*}{\Rightarrow} uyv\ (G_1)$. If $u \neq \Lambda$ or $v \neq \Lambda$, then for every s-derivation $s_0, s_1, \ldots, s_p$ of uxv in G_1 and for every s-derivation $t_0, t_1, \ldots, t_q$ of uyv in G_1 with the properties $p \geqq 1$, $q \geqq 1$ we have $s_1 = t_1$.

(vi) If $s \in V_1$, $x, y \in V_1^*$ are such elements that $s \overset{*}{\Rightarrow} x\ (G_1)$ and $s \overset{*}{\Rightarrow} xy\ (G_1)$ then $y = \Lambda$.

(vi′) If $s \in V_1$, $x, y \in V_1^*$ are such elements that $s \overset{*}{\Rightarrow} x\ (G_1)$ and $s \overset{*}{\Rightarrow} yx\ (G_1)$ then $y = \Lambda$.

(vii) Let $u, v, y \in V_1^*$, $x \in V_1$, $s \in V_1$ be such elements that $s \overset{*}{\Rightarrow} uxv\ (G_1)$, $s \overset{*}{\Rightarrow} uyv\ (G_1)$, $|x| < |y|$. Then $x \overset{*}{\Rightarrow} y\ (G_1)$.

Proof of (*i*). It is clear that $(x, y) \in R_1$ implies $|y| \geqq 2$. It follows that $x \Rightarrow y\ (G_1)$ implies $|x| < |y|$. Thus, $s = s_0 \Rightarrow s_1 \Rightarrow s_2 \Rightarrow \ldots \Rightarrow s_p = x$ implies $|s| \leqq |x|$. Further on, $p > 0$ implies $|s_p| > 1$ and $p > 1$ implies $|s_p| > 2$. Therefore, $s \overset{*}{\Rightarrow} x\ (G_1)$, $|x| = 1$ implies $p = 0$ and $s = x$ and $s \overset{*}{\Rightarrow} x\ (G_1)$, $|x| = 2$ implies $p = 1$ and $s \to x\ (G_1)$.

Proof of (*ii*). If none of the first two assertions holds, then there exists an element $c \in V_1$ such that $uyv = uy_1cy_2v$, $azb = az_1cz_2b$, $uy_1 = az_1$, $y_2v = z_2b$ for suitable strings $y_1, y_2, z_1, z_2 \in V_1^*$. We have $y_1cy_2 = y$, $z_1cz_2 = z$. Clearly, $y_1 = \Lambda$ implies $c = [_r$ for a suitable $r \in R$ which implies $z_1 = \Lambda$ as $[_r \in V_1 - V$ and all symbols of z which are different from the first and the last one are in V. In a similar way we prove that $z_1 = \Lambda$ implies $y_1 = \Lambda$. The conditions $y_1 = \Lambda$, $z_1 = \Lambda$ are thus equivalent. Similarly, the conditions $y_2 = \Lambda$, $z_2 = \Lambda$ are equivalent, too.

(α) If $y_1 = \Lambda$, then $z_1 = \Lambda$ and $y_2 \neq \Lambda \neq z_2$ as $|y| \geqq 2$, $|z| \geqq 2$. From the equality $y_2v = z_2b$ and from the fact that $y_2 = w]_r$, $z_2 = w']_{r'}$ for suitable $w, w' \in V_1^*$, $r, r' \in R$ and from the fact that $]_r \in V_1 - V$,

[5]) Compare [5], p. 35, Exercice 1.6.8 and [6].

$]_{r'} \in V_1 - V$ we get $w = w'$, $r = r'$ which implies $y_2 = z_2$ and $v = b$. Thus we have $y = cy_2 = cz_2 = z$. Clearly, $u = a$.

(β) If $y_1 \neq \Lambda$, then $z_1 \neq \Lambda$ and $uy_1 = az_1$. It implies the existence of $w, w' \in V^*$, $r, r' \in R$ such that $y_1 = [_r w$, $z_1 = [_{r'} w'$. As $[_r, [_{r'} \in V_1 - V$ we have $w = w'$, $r = r'$ which implies $y_1 = z_1$ and $u = a$. If $y_2 = \Lambda$, then $z_2 = \Lambda$ and $v = b$; further, $y = y_1 c = z_1 c = z$. If $y_2 \neq \Lambda$, then $z_2 \neq \Lambda$ and we get from $y_2 v = z_2 b$ the conditions $y_2 = z_2$, $v = b$ similarly as in the case (α). It implies $y = y_1 c y_2 = z_1 c z_2 = z$.

We have proved that the negation of the disjunction of the first two possibilities implies the third possibility. Thus, (*ii*) is proved.

The proof of (*iii*) will be obtained by induction with respect to $| uyv |$. Clearly $| y | \geqq 2$ and thus $| uyv | \geqq 2$.

By $G(n)$ we denote the following assertion: If $u, v \in V_1^*$, $(x, y) \in R_1$, $s \in V_1$ are such elements that $s \overset{*}{\Rightarrow} uyv\ (G_1)$, $| uyv | = n$, then $s \overset{*}{\Rightarrow} uxv\ (G_1)$.

$G(2)$ holds true. Indeed, let us have such elements $u, v \in V_1^*$, $(x, y) \in R_1$, $s \in V_1$ that $s \overset{*}{\Rightarrow} uyv\ (G_1)$, $| uyv | = 2$. Then $s \rightarrow uyv\ (G_1)$ according to (*i*). It implies the existence of elements $w \in V^*$, $r \in R$ such that $(s, w) = r$, $uyv = [_r w]_r$. As we have $(x, y) \in R_1$, then there exist such elements $r' \in R$, $w' \in V^*$ that $r' = (x, w')$, $y = [_{r'} w']_{r'}$. It implies $[_r w]_r = uyv = u[_{r'} w']_{r'} v$. As $w \in V^*$ and $[_{r'},]_{r'} \in V_1 - V$ we have $u = \Lambda = v$ which implies $[_r w]_r = [_{r'} w']_{r'}$. It follows $r = r'$, $w = w'$. Thus, $(x, w) = (x, w') = r' = r = (s, w)$ which implies $x = s$. Thus, $s \overset{*}{\Rightarrow} s\ (G_1)$ and $s = uxv$. Therefore, $G(2)$ holds true.

Let $m > 2$ be a natural number. We prove that the validity of $G(2), G(3), \ldots, G(m - 1)$ implies the validity of $G(m)$.

Let $u, v \in V_1^*$, $(x, y) \in R_1$, $s \in V_1$ be such elements that $s \overset{*}{\Rightarrow} uyv\ (G_1)$, $| uyv | = m$. Then there exists an s-derivation $s_0, s_1, \ldots, s_p$ of uyv in G_1. As $| uyv | > 2$, we have $p \geqq 1$. Further we have $s_{p-1} \Rightarrow uyv\ (G_1)$. Thus, there exist some elements $a, b \in V_1^*$, $(t, z) \in R_1$ such that $s_{p-1} = atb$, $uyv = s_p = azb$. According to (*ii*) we have three possibilities:

(α) There exist such strings $b_1, b_2 \in V_1^*$ that $u = azb_1$, $b_1 b_2 = b$, $b_2 = yv$. In this case we have $s_{p-1} = atb = atb_1 b_2 = atb_1 yv$ and $| atb | < | azb | = | uyv | = m$. As $G(2), \ldots, G(m - 1)$ are valid, then the fact that $| atb_1 yv | < m$ implies $s \overset{*}{\Rightarrow} atb_1 xv\ (G_1)$. Clearly, $atb_1 xv \Rightarrow azb_1 xv\ (G_1)$ and $azb_1 xv = uxv$. Thus $s \overset{*}{\Rightarrow} uxv\ (G_1)$.

(β) There exist such strings $a_1, a_2 \in V_1^*$ that $uy = a_1$, $a_1 a_2 = a$, $a_2 zb = v$. Similarly as in (α) we prove $s \overset{*}{\Rightarrow} uxv\ (G_1)$.

(γ) We have $u = a$, $y = z$, $v = b$. Then $(t, z) = (t, [_r w]_r)$ for a suitable $r = (t, w) \in R$ and $(x, y) = (x, [_{r'} w']_{r'})$ for a suitable $r' = (x, w') \in R$. From $[_r w]_r = z = y = [_{r'} w']_{r'}$ we have $r = r'$, $w = w'$ which implies $t = x$. Thus, $s_{p-1} = atb = axb = uxv$ and $s \overset{*}{\Rightarrow} s_{p-1}(G_1)$. Therefore, $s \overset{*}{\Rightarrow}$ $\overset{*}{\Rightarrow} uxv\ (G_1)$.

In all three cases we have proved $s \overset{*}{\Rightarrow} uxv\ (G_1)$. Thus the validity of

$G(2), \ldots, G(m-1)$ implies the validity of $G(m)$. Therefore $G(m)$ holds true for $m = 2, 3, \ldots$. It follows: If $u, v \in V_1^*$, $(x, y) \in R_1$, $s \in V_1$ are such elements that $s \overset{*}{\Rightarrow} uyv\ (G_1)$, then $s \overset{*}{\Rightarrow} uxv\ (G_1)$. Thus (*iii*) has been proved.

The proof of (*iv*) will be obtained by induction with respect to the index i of s_i.

Let $s \in V_1$, $x \in V_1^*$ be such elements that $s_0, s_1, \ldots, s_p$ is an s-derivation of x in G_1 with the property $p \geqq 1$. By $H(n)$ we denote the following assertion: The first [last] symbols of s_1 and s_n are equal.

$H(1)$ holds true trivially.

Let m be a natural number, $1 < m \leqq p$. We prove that the validity of $H(m-1)$ implies the validity of $H(m)$. We have $s_{m-1} \Rightarrow s_m(G_1)$. Then there exist some strings $u, v \in V_1^*$ and an element $(t, z) \in R_1$ such that $s_{m-1} = utv$, $uzv = s_m$. As $t \in V$ and the first [last] symbol of s_{m-1} is, according to $H(m-1)$, equal to the first [last] symbol of s_1 which is an element of $V_1 - V$, we have $u \neq \Lambda \neq v$. Clearly, the first [last] symbol of $s_m = uzv$ is equal to the first [last] symbol of s_{m-1} which is equal to the first [last] symbol of s_1.

We have proved that for each m, $1 < m \leqq p$ the validity of $H(m-1)$ implies the validity of $H(m)$. Thus $H(m)$ holds true for $m = 1, 2, \ldots, p$. Clearly, the conjunction of $H(1), H(2), \ldots, H(p)$ is (*iv*).

Proof of (*v*). There exist some $r, r' \in R$, $w, w' \in V^*$ such that $s_1 = [_r w]_r$, $t_1 = [_{r'} w']_{r'}$. If $u \neq \Lambda$, then the first symbol of u is $[_r$ which follows from the fact that $s_0, s_1, \ldots, s_p$ is an s-derivation of uxv in G_1 according to (*iv*). In the same manner we obtain the first symbol of u being $[_{r'}$. It implies $r = r'$ from which it follows $w = w'$ and $s_1 = t_1$. If $v \neq \Lambda$, then the proof can be obtained similarly.

The proof of (*vi*) will be obtained by induction with respect to $|xy|$. We have $1 = |s| \leqq |x| \leqq |xy|$ according to (*i*).

We denote by $I(n)$ the following assertion: If $s \in V_1$, $x, y \in V_1^*$ are such elements that $s \overset{*}{\Rightarrow} x\ (G_1)$, $s \overset{*}{\Rightarrow} xy\ (G_1)$, $xy = |n|$ then $y = \Lambda$.

$I(1)$ holds true. Indeed, let us have $s \in V_1$, $x, y \in V_1^*$ such that $s \overset{*}{\Rightarrow} x\ (G_1)$, $s \overset{*}{\Rightarrow} xy\ (G_1)$, $|xy| = 1$. Then $s = xy$ and $1 = |s| \leqq |x|$ according to (*i*). It implies $s = x$ and $y = \Lambda$.

Let $m > 1$ be a natural number. We prove that the validity of $I(1)$, $I(2), \ldots, I(m-1)$ implies the validity of $I(m)$.

Let us have such elements $s \in V_1$, $x, y \in V_1^*$ that $s \overset{*}{\Rightarrow} x\ (G_1)$, $s \overset{*}{\Rightarrow} xy\ (G_1)$, $|xy| = m$. We take an s-derivation $s_0, s_1, \ldots, s_p$ of x in G_1 and an s-derivation $t_0, t_1, \ldots, t_q$ of xy in G_1. We have $|xy| = m > 1$ which implies $q \geqq 1$ according to (*i*).

If $p = 0$ then $t_0 = s = s_0 = x$ and $|x| = 1$. It follows $t_0 \to t_1(G_1)$ and the first symbol x of $t_q = xy$ is equal to the first symbol of t_1 according to (*iv*). But $t_0 \to t_1(G_1)$ implies that, for the first symbol x of t_1, we have

$x \in V_1 - V$. Beyond, $t_0 \to t_1(G_1)$ and $x = t_0$ imply $x \in V$ which is a contradiction.

Thus, we have $p \geqq 1$. According to (v) $s_1 = t_1$ holds true. Clearly, $|s_1| \geqq 2$. We put $n = |s_1|$. Let $a_1, a_2, \ldots, a_n \in V_1$ be such elements that $s_1 = a_1a_2\ldots a_n = t_1$. According to 1.8 there exist such strings $x_i \in V_1^*$, $y_i \in V_1^*$ $(i = 1, 2, \ldots, n)$ that $a_i \overset{*}{\Rightarrow} x_i(G_1)$, $a_i \overset{*}{\Rightarrow} y_i(G_1)$ for $i = 1, 2, \ldots, n$ and $x_1x_2 \ldots x_n = x$, $y_1y_2 \ldots y_n = xy$.

Let us suppose that $x_i \neq y_i$ for at least one index i. We denote by i_0 the least index for which $x_{i_0} \neq y_{i_0}$. Then $x_1x_2 \ldots x_{i_0-1}\, x_{i_0} \ldots x_ny = xy = y_1y_2 \ldots y_n = x_1x_2 \ldots x_{i_0-1}y_{i_0} \ldots y_n$ which implies $x_{i_0} \ldots x_ny = y_{i_0} \ldots y_n$. It implies the existence of such a string $u \in V_1^*$, $u \neq \Lambda$, that either $x_{i_0} = y_{i_0}u$ or $y_{i_0} = x_{i_0}u$. We have $|y_{i_0}u| = |x_{i_0}| < |x| \leqq |xy| = m$ in the first case and $|x_{i_0}u| = |y_{i_0}| < |xy| = m$ in the second. Thus, the following conditions are satisfied in the first case: $a_{i_0} \in V_1$, $y_{i_0} \in V_1^*$, $u \in V_1^*$, $a_{i_0} \overset{*}{\Rightarrow} y_{i_0}(G_1)$, $a_{i_0} \overset{*}{\Rightarrow} y_{i_0}u(G_1)$, $|y_{i_0}u| < m$. As $I(1), I(2), \ldots, I(m-1)$ are valid, we have $u = \Lambda$, which is a contradiction. Similarly, in the second case, we obtain $u = \Lambda$, which is a contradiction, too. Thus, we have $x_i = y_i$ for $i = 1, 2, \ldots, n$ and $x = x_1x_2 \ldots x_n = y_1y_2 \ldots y_n = xy$ which implies $y = \Lambda$.

Thus, the validity of $I(1), I(2), \ldots, I(m-1)$ implies that of $I(m)$. Therefore, $I(m)$ is valid for $m = 1, 2, \ldots$, which is (vi).

The assertion (vi') can be demonstrated in a similar way.

The proof of (vii) will be obtained by induction with respect to $|uxv|$. Clearly, $|uxv| \geqq 1$.

We denote by $J(n)$ the following assertion: If $u, v, y \in V_1^*$, $x \in V_1$, $s \in V_1$ are such elements that $s \overset{*}{\Rightarrow} uxv\,(G_1)$, $s \overset{*}{\Rightarrow} uyv\,(G_1)$, $|x| < |y|$, $|uxv| = n$, then $x \overset{*}{\Rightarrow} y\,(G_1)$.

$J(1)$ holds true. Indeed, if $u, v, y \in V_1^*$, $x \in V_1$, $s \in V_1$ are such elements that $s \overset{*}{\Rightarrow} uxv\,(G_1)$, $s \overset{*}{\Rightarrow} uyv\,(G_1)$, $|x| < |y|$, $|uxv| = 1$, then $s = uxv$ according to (i). It follows $s = x$, $u = \Lambda = v$ and $x = s \overset{*}{\Rightarrow} uyv\,(G_1)$, $uyv = y$. Thus, $x \overset{*}{\Rightarrow} y\,(G_1)$.

Let $m > 1$ be a natural number. We prove that the validity of $J(1)$; $J(2), \ldots, J(m-1)$ implies the validity of $J(m)$.

Let $u, v, y \in V_1^*$, $x \in V_1$, $s \in V_1$ be such elements that $s \overset{*}{\Rightarrow} uxv\,(G_1)$, $s \overset{*}{\Rightarrow} uyv\,(G_1)$, $|x| < |y|$, $|uxv| = m$. If $u = \Lambda = v$, then $m = |uxv| = |x| = 1$, which is a contradiction. Thus, $u \neq \Lambda$ or $v \neq \Lambda$. Let $s_0, s_1, \ldots, s_p$ be an s-derivation of uxv in G_1, $t_0, t_1, \ldots, t_q$ an s-derivation of uyv in G_1. As $|uxv| = m > 1$, $|uyv| > |uxv| > 1$, then we have $p \geqq 1$, $q \geqq 1$ according to (i). According to (v) we have $s_1 = t_1$. Clearly, $|s_1| \geqq 2$. We put $n = |s_1|$, $s_1 = a_1a_2 \ldots a_n = t_1$ where $a_i \in V_1$ for $i = 1, 2, \ldots, n$. There exist such strings $x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n \in V_1^*$ that $a_i \overset{*}{\Rightarrow} x_i(G_1)$, $a_i \overset{*}{\Rightarrow} y_i(G_1)$ for $i = 1, 2, \ldots, n$, $x_1x_2 \ldots x_n = uxv$, $y_1y_2 \ldots y_n = uyv$ according to 1.8. By i_0 we denote such an index that there exist some $b,\ c \in V_1^*$ that $bxc = x_{i_0}$, $u = x_1x_2 \ldots x_{i_0-1}b$, $v = cx_{i_0+1} \ldots x_n$.

Let us suppose that there exists such an index i, $1 \leqq i < i_0$, that $x_i \neq y_i$. We denote by i_1 the least index with this property. It follows $x_1x_2 \ldots x_{i_1-1}y_{i_1} \ldots y_n = y_1y_2 \ldots y_n = uyv = x_1x_2 \ldots x_{i_1-1}x_{i_1} \ldots x_{i_0-1}byv$ which implies $y_{i_1} \ldots y_n = x_{i_1} \ldots x_{i_0-1}byv$. Thus, there exists such a string $w \in V_1^*$, $w \neq \Lambda$, that $y_{i_1} = x_{i_1}w$ or $x_{i_1} = y_{i_1}w$. Therefore, we have $a_{i_1} \in V_1$, $a_{i_1} \overset{*}{\Rightarrow} x_{i_1}(G_1)$, $a_{i_1} \overset{*}{\Rightarrow} x_{i_1}w\ (G_1)$ in the first case and $a_{i_1} \in V_1$, $a_{i_1} \overset{*}{\Rightarrow} y_{i_1}(G_1)$, $a_{i_1} \overset{*}{\Rightarrow} y_{i_1}w\ (G_1)$ in the second. According to (vi) we have $w = \Lambda$, which is a contradiction. Thus, $x_i = y_i$ for $i = 1, 2, \ldots, i_0 - 1$. Similarly, by means of (vi') we obtain $x_i = y_i$ for $i = i_0 + 1, \ldots, n$. Thus, $x_1x_2 \ldots x_{i_0-1}y_{i_0}x_{i_0+1} \ldots x_n = y_1y_2 \ldots y_n = uyv = x_1x_2 \ldots x_{i_0-1}bycx_{i_0+1} \ldots x_n$ which implies $y_{i_0} = byc$.

We have proved the existence of such strings $b,\ c \in V_1^*$ and of such an element $a_{i_0} \in V_1$ that $a_{i_0} \overset{*}{\Rightarrow} bxc\ (G_1)$, $a_{i_0} \overset{*}{\Rightarrow} byc\ (G_1)$, $|x| < |y|$, $|bxc| = |x_{i_0}| < |uxv| = m$. As $J(1), J(2), \ldots, J(m-1)$ are valid, we have $x \overset{*}{\Rightarrow} y\ (G_1)$.

We have proved thad the validity of $J(1), J(2), \ldots, J(m-1)$ implies the validity of $J(m)$. Thus, $J(m)$ is valid for $m = 1, 2, \ldots$ which is (vii).

5. CHARACTERIZATION OF CONTEXT-FREE LANGUAGES

5.1. Lemma. *Let $G = \langle V, \{\sigma\}, R\rangle$ be a special context-free grammar having the property* (P) *of* 1.17. *Let Z_1, Z_2 be sets with the following properties: there exists a bijection f_1 of R onto Z_1, a bijection f_2 of R onto Z_2 and the sets V, Z_1, Z_2 are mutually disjoint. We put $f_1(r) = [_r, f_2(x) =]_r$ for every $r \in R$, $V_1 = V \cup Z_1 \cup Z_2$, $R_1 = \{(x, [_ry]_r);\ (x, y) = r \in R\}$, $G_1 = \langle V_1, \{\sigma\}, R_1\rangle$.*

Then the following assertions hold true:

(i) *If $(x, y) \in R_1$, then $x \equiv y\ (V_1, \mathscr{L}(G_1))$.*

(ii) *G is the trace of G_1 in V^*.*

(iii) *We have $R_1 \subseteq E(V_1, \mathscr{L}(G_1))$.*

(iv) *The sets V_1, $A(V_1, \mathscr{L}(G_1))$ are finite.*

(v) *If $(x, y) \in E(V_1, \mathscr{L}(G_1))$, then $x \overset{*}{\Rightarrow} y\ (G_1)$.*

(vi) *We have $F(V_1, \mathscr{L}(G_1)) \subseteq R_1$ and the set $F(V_1, \mathscr{L}(G_1))$ is finite.*

(vii) *If $(x, y) \in F(V_1, \mathscr{L}(G_1))$, then $x \in V$.*

Proof of (i). Let us have $(x, y) \in R_1$.

If $uxv \in \mathscr{L}(G_1)$ for some strings $u, v \in V_1^*$, then $\sigma \overset{*}{\Rightarrow} uxv\ (G_1)$. As we have $uxv \Rightarrow uyv\ (G_1)$, it follows $\sigma \overset{*}{\Rightarrow} uyv\ (G_1)$ and $uyv \in \mathscr{L}(G_1)$. Thus, $x > y\ (V_1, \mathscr{L}(G_1))$.

If $uyv \in \mathscr{L}(G_1)$ for some strings $u, v \in V_1^*$, then $\sigma \overset{*}{\Rightarrow} uyv\ (G_1)$. It follows from 4.1 ($iii$) that $\sigma \overset{*}{\Rightarrow} uxv\ (G_1)$, i.e. $uxv \in \mathscr{L}(G_1)$. Thus, $y > x\ (V_1, \mathscr{L}(G_1))$.

We have proved that $(x, y) \in R_1$ implies $x \equiv y\ (V_1, \mathscr{L}(G_1))$.

(ii) is clear.

Proof of *(iii)*. According to *(ii)* $G = \langle V, \{\sigma\}, R\rangle$ is the trace of $G_1 = \langle V_1, \{\sigma\}, R_1\rangle$ in V^* and $(x, y) \in R_1$ implies $x \in V$. According to 2.5 the language $(V, \mathscr{L}(G))$ is the trace of $(V_1, \mathscr{L}(G_1))$ in V^*.

We prove that G_1 has the property (P) of 1.17.

If $(x, y) \in R_1$, then $x \in V$ and there exists a string $y' \in V^*$ such that $(x, y') \in R$. As G has the property (P), then there exist some strings $u, v \in V^*$ with the property $\sigma \overset{*}{\Rightarrow} uxv\ (G)$, i.e. $uxv \in \mathscr{L}(G)$. As $(V, \mathscr{L}(G))$ is the trace of $(V_1, \mathscr{L}(G_1))$ in V^*, there exists some string $w \in \mathscr{L}(G_1)$ such that $\mathbf{t}^V_*(w) = uxv$. According to 2.2 (B), there exist some strings $u', x', v' \in V_1^*$ such that $u'x'v' = w$, $\mathbf{t}^V_*(u') = u$, $\mathbf{t}^V_*(x') = x$, $\mathbf{t}^V_*(v') = v$. Clearly, $x' \neq \Lambda$. We put $x' = x_1x_2 \dots x_n$ for a suitable natural number n and for suitable elements $x_i \in V_1$ $(i = 1, 2, \dots, n)$. We have $\mathbf{t}^V(x_1)\,\mathbf{t}^V(x_2) \dots \dots \mathbf{t}^V(x_n) = \mathbf{t}^V_*(x') = x$ according to 2.1. It follows that there exists an index i_0 $(1 \leqq i_0 \leqq n)$ such that $\mathbf{t}^V(x_{i_0}) = x$. It follows from 2.1 that $x_{i_0} = x$. Thus, $w = u'x_1 \dots x_{i_0-1}x\,x_{i_0+1} \dots x_nv'$. Thus, for every $(x, y) \in R_1$ there exist some strings $w_1 = u'x_1 \dots x_{i_0-1}$, $w_2 = x_{i_0+1} \dots x_nv'$, w_1, $w_2 \in V_1^*$, such that $w_1xw_2 = w \in \mathscr{L}(G_1)$, i.e. $\sigma \overset{*}{\Rightarrow} w_1xw_2(G_1)$.

Thus G_1 has the property (P).

Let us have $(x, y) \in R_1$. As G_1 has the property (P), we have $xv(V_1, \mathscr{L}(G_1))$ which implies $yv(V_1, \mathscr{L}(G_1))$. According to *(i)* we have $x \equiv y\ (V_1, \mathscr{L}(G_1))$. Moreover, we have $1 = |x| < |y|$. Thus y is a configuration of $(V_1, \mathscr{L}(G_1))$ and x is its result. Thus, $(x, y) \in E(V_1, \mathscr{L}(G_1))$ and we have proved $R_1 \subseteq E(V_1, \mathscr{L}(G_1))$.

Proof of *(iv)*. We put $M = \{y;\ (x, y) \in R_1\}$. According to *(iii)* we have $M \subseteq C(V_1, \mathscr{L}(G_1))$. It follows $\mathscr{L}(G_1) - V_1^*MV_1^* \supseteq \mathscr{L}(G_1) - V_1^*C(V_1, \mathscr{L}(G_1))\,V_1^* = A(V_1, \mathscr{L}(G_1))$. But clearly $\mathscr{L}(G_1) - V_1^*MV_1^* = \{\sigma\}$. It follows that $A(V_1, \mathscr{L}(G_1))$ is finite.

The finiteness of V_1 is clear.

Proof of *(v)*. Let us have $(x, y) \in E(V_1, \mathscr{L}(G_1))$. Thus y is a configuration of $(V_1, \mathscr{L}(G_1))$, x its result. Then $yv(V_1, \mathscr{L}(G_1))$, $y \equiv x\ (V_1, \mathscr{L}(G_1))$ and $1 = |x| < |y|$. Thus there exist some strings $u, v \in V_1^*$ such that $uyv \in \mathscr{L}(G_1)$, i.e. $\sigma \overset{*}{\Rightarrow} uyv\ (G_1)$. From $y \equiv x\ (V_1, \mathscr{L}(G_1))$ it follows $uxv \in \mathscr{L}(G_1)$ and $\sigma \overset{*}{\Rightarrow} uxv\ (G_1)$; from $|x| < |y|$ we have $x \overset{*}{\Rightarrow} y\ (G_1)$ according to *(vii)* of 4.1.

Proof of *(vi)*. Let us have $(x, y) \in F(V_1, \mathscr{L}(G_1))$. Thus y is a configuration of $(V_1, \mathscr{L}(G_1))$ and x its result. According to *(v)* we have $x \overset{*}{\Rightarrow} y\ (G_1)$. Thus there exists an x-derivation $s_0, s_1, \dots, s_p$ of y in G_1 with $p \geqq 1$. We have $s_{p-1} \Rightarrow y\ (G_1)$. Thus there exist some strings u, $v \in V_1^*$ and $(t, z) \in R_1$ such that $s_{p-1} = utv$, $uzv = y$. According to *(iii)* we have $z \in C(V_1, \mathscr{L}(G_1))$. From $y \in P(V_1, \mathscr{L}(G_1))$ it follows $u = \Lambda = v$,

$y = z$ and $s_{p-1} = t$. According to 4.1 (*i*) we have $p - 1 = 0$ and $x = s_0 = t$. Thus, $(x, y) = (t, z) \in R_1$.

We have proved $F(V_1, \mathscr{L}(G_1)) \subseteq R_1$. Since the set R_1 is finite, then the set $F(V_1, \mathscr{L}(G_1))$ is finite as well.

(*vii*) is an immediate consequence of the inclusion $F(V_1, \mathscr{L}(G_1)) \subseteq R_1$ and of the fact that $(x, y) \in R_1$ implies $x \in V$.

5.2. Main Theorem. *Let (U, L) be a language. Then the following assertions are equivalent*:

(A) *(U, L) is a context-free language.*

(B) *There exists a language of strong depth 1 (V_1, L_1), a set V essential with respect to (V_1, L_1) and a full language (W, W^*) such that (U, L) is the intersection of (W, W^*) with the trace of (V_1, L_1) in V^*.*

Proof. Let (A) be fulfilled. According to 1.16 and 1.17 there exists such a context-free grammar $H = \langle V, U, \{\sigma\}, R\rangle$ in the usual sense having the property (P) that $L = \mathscr{L}(H)$. We put $G = \langle V, \{\sigma\}, R\rangle$; thus, G is a special context-free grammar having the property (P). We construct the grammar $G_1 = \langle V_1, \{\sigma\}, R_1\rangle$ in the same manner as in 5.1. Then the sets V_1, $A(V_1, \mathscr{L}(G_1))$, $F(V_1, \mathscr{L}(G_1))$ are finite according to 5.1 (*iv*) and (*vi*), and thus, $(V_1, \mathscr{L}(G_1))$ is a language of strong depth 1 according to 3.11. The set V is essential with respect to $(V_1, \mathscr{L}(G_1))$ and $(V, \mathscr{L}(G))$ is the trace of $(V_1, \mathscr{L}(G_1))$ in V^* according to 5.1 (*vii*) and (*ii*) and according to 2.5. We have now $\mathscr{L}(H) = \mathscr{L}(G) \cap U^*$, thus (U, L) is the intersection of $(V, \mathscr{L}(G))$ with the full language (U, U^*).

We have proved that (A) implies (B).

Let (B) be fulfilled. Then (V_1, L_1) is a special context-free language according to 3.13 and $K(V_1, L_1) = \langle V_1, A(V_1, L_1), F(V_1, L_1)\rangle$ is a special context-free grammar generating (V_1, L_1) according to 3.8 and 3.12. Since V is essential with respect to (V_1, L_1), then the trace (V, L_2) of (V_1, L_1) in V^* is generated by the trace G of $K(V_1, L_1)$ in V^* according to 2.5. From the fact that V is essential with respect to (V_1, L_1) it follows that G is a special context-free grammar. Thus (V, L_2) is a special context-free language. Thus, (U, L) which is the intersection of (V, L_2) and the full language (W, W^*), is a context-free language.

We have proved that (B) implies (A).

BIBLIOGRAPHY

[1] Novotný, M.: *Algebraic structures of mathematical linguistics*. Bull. Math. de la Soc. Sci Math. de la R. S. de Roumanie **12** (**60**) (1969), 87—101.

[2] Kulagina, O. S.: *Ob odnom sposobe opredelenija grammatičeskich ponjatij na baze teorii množestv*. Problemy kibernetiki **1** (1958), 203—214.

[3] Gladkij, A. V.: *Konfiguracionnye charakteristiki jazykov*. Problemy kibernetiki **10** (1963), 251—260.

[4] Novotný, M.: *On configurations in the sense of Kulagina.* Publ. Fac. Sci. Univ. J. E. Purkyně Brno No 507 (1969), 301—316.
[5] Ginsburg, S.: *The Mathematical Theory of Context-Free Languages.* Mc Graw-Hill, New York 1966.
[6] Ginsburg, S. and Harison, M. A.: *Bracketed context free languages.* System Development Corp. Rept. TM-738/023/00 Jan. 4, 1966.

Department of Mathematics
J. E. Purkyně University, Brno
Czechoslovakia