Czechoslovak Mathematical Journal

Jozef Gruska

On structural unambiguity of formal languages

Czechoslovak Mathematical Journal, Vol. 15 (1965), No. 2, 283–294

Persistent URL: http://dml.cz/dmlcz/100670

Чекословацкий математический журнал, т. 15 (90) 1965, Прага

ON STRUCTURAL UNAMBIGUITY OF FORMAL LANGUAGES

JOZEF GRUSKA, Bratislava

(Received February 4, 1964)

1. INTRODUCTION AND SUMMARY

The ambiguity problem of Chomsky's context-free grammars has been attacked primarily from the negative point of view and it has been proved by many authors (see [1], [3], [4], e.g.) that it is unsolvable. Of course this does not mean that it is impossible to devise methods which may be useful to decide, at least for some languages, whether or not they are structurally ambiguous (s.a.). Some methods of this kind have been recently investigated by Fabian [2] for slightly more general languages. Moreover, the relationships between structural unambiguity (s.u.) and semantics were studied in [2] and it was shown that from a point of view it is sufficient to require, from a language, that it be weakly structurally unambiguous (w.s.u.). (Strong) structural unambiguity is, however, better suited for a study.

This paper brings essentially two results, the first concerning a property of non-cyclicity of languages, the second the relations between weak and strong structural unambiguity.

Every cyclic (i.e. such that there is a text t derivable from the same text t) language is structurally ambiguous and, under certain reasonable conditions, weakly structurally ambiguous (w.s.a.), too. There are provided quite efficient means for verifying whether or not a language is cyclic. Hence the question of structural ambiguity is interesting only for non-cylic languages. But the assumption of non-cyclicity may simplify many methods and proofs used in [2]. In particular, it is shown that for a certain class of languages, which contains Chomsky's context-free grammars, some conditions required for a reducing transformation (a basic concept in [2]) are always satisfied.

In [2] it has been shown that for some languages both structural unambiguity and weak s.u. coincide. The present paper shows that to a given language $\mathscr{L}$, which satisfies certain conditions, a language $\mathscr{L}_0$ may be constructed such that $\mathscr{L}$ is w.s.u. if and only if $\mathscr{L}_0$ is s.u. (If those conditions are not satisfied then $\mathscr{L}$ is w.s.a.). Hence it is sufficient to study structural unambiguity of formal languages.

The present paper uses notations and definitions of [2]. The reader should be familiar with sections 1 to 8, [2] and with Definition 9.1, [2].

2. CYCLIC LANGUAGES AND STRUCTURAL UNAMBIGUITY

2.1. Notes. We shall use sometimes the following notation:

$$\mathscr{L}: \quad A_1 \Rightarrow \alpha_1^1, \ldots, A_1 \Rightarrow \alpha_{k_1}^1, \quad A_2 \Rightarrow \alpha_1^2, \ldots, A_n \Rightarrow \alpha_{k_n}^n$$

for defining a particular language $\mathscr{L}$ such that $\mathbf{d}\mathscr{L} = \{A_1, \ldots, A_n\}$, $\mathscr{L}A_i = \{\alpha_1^i, \ldots, \alpha_{k_i}^i\}$. If $g = [A, t] \in \mathbf{g}\mathscr{L}$, we shall denote the set of all structures of g in $\mathscr{L}$ by $S_{\mathscr{L}}[A, t]$ or $S_{\mathscr{L}}g$. If there is no danger of misunderstanding the symbol specifying the language will be deleted.

The following lemma can be easily proved using Lemmas 6.3, 6.1 and 5.5, [2].

2.2. Lemma. *If $t_1 \to t_2$, then there exists a t_1-decomposition τ of t_2 such that $[t_1 i] \to \tau i$ at least for one $i \in \mathbf{d}t_1$.*

2.3. Definition. A language $\mathscr{L}$ is called cyclic (primitive cyclic) if there exists a $t \in \boldsymbol{\sigma}\mathscr{L}$, $(A \in \mathbf{d}\mathscr{L})$ such that $t \to t$ $([A] \to [A])$.

2.4. Lemma. *Let $t \to t$ for some $t \in \boldsymbol{\sigma}\mathscr{L}$. Then either $\lambda t = 1$ or there is a $t_0 \in \boldsymbol{\sigma}\mathscr{L}$ such that $t_0 \to t_0$ and $\lambda t_0 < \lambda t$.*

Proof. Let $t \to t$ and $\lambda t > 1$. By Lemma 2.2 there is a t-decomposition τ of t and an $i_1 \in \mathbf{d}t$ such that $[t i_1] \to \tau i_1$. If $\tau i = \Lambda$ for no $i \in \mathbf{d}\tau$, then $\tau i = [ti]$ for each i and the assertion of Lemma holds with $t_0 = [t i_1]$, which has length 1. It remains the case that there is a $j \in \mathbf{d}t$ such that $\tau j = \Lambda$. Then the assertion of Lemma holds with $t_0 = t^{(1,j-1)} \times t^{(j+1,\lambda t)}$. Indeed, since $[tj] \to \tau j = \Lambda$, we have

$$t_0 \rightrightarrows \prod(\tau^{(1,j,-1)} \times \tau^{(j+1,\lambda\tau)}) = \prod \tau = t \to t_0$$

and $\lambda t_0 = \lambda t - 1$.

2.5. Theorem. *A language $\mathscr{L}$ is cyclic if and only if it is primitive cyclic.*

Proof. The Theorem follows immediately from the preceding Lemma and from Definition 2.3.

2.6. Theorem. *Every cyclic language is structurally ambiguous.*

Proof. By Theorem 2.5 and Theorem 7.6, [2].

2.7. Remark. The converse of this Lemma is not true.
Example: $\mathscr{L}$: $A \Rightarrow B, A \Rightarrow D, B \Rightarrow D$.

2.8. Remark. A cyclic language can be weakly structurally unambiguous.
Example: $\mathscr{L}$: $A \Rightarrow B, B \Rightarrow A$.

2.9. Theorem. *Let $\mathscr{L}$ be a cyclic language and let there exist an $A \in \mathbf{d}\mathscr{L}$ such that $[A] \to [A]$, $\mathbf{t}_t(\mathscr{L}, A) \neq \Lambda$. Then $\mathscr{L}$ is not w.s.u.*

Proof. Let n be the smallest integer such that there exist $A \in \mathbf{d}\mathscr{L}$, $t \in \boldsymbol{\sigma}_t\mathscr{L}$ such that $[A] \to [A]$, $[A] \to t$ and $\delta[A, t] = n$. We shall prove that $[A, t]$ is s.a. Since $[A] \to [A]$ and $[A] \notin \mathscr{L}A$, there is an α_1 such that $[A] \Rightarrow \alpha_1 \to [A] \to t$. If $n = 1$, then $[A] \Rightarrow t$ and $[A, t]$ has two different structures $[A, [t]]$ and $[\alpha_1, \tau_1]$ where τ_1 is an α_1-decomposition of t. Let $n > 1$. By Theorem 6.5, [2], there exists such an $[\alpha, \tau] \in S[A, t]$ that

$$(1) \qquad \delta[A, t] = 1 + \sum_{i=1}^{\lambda\alpha} \delta_0[\alpha i, \tau i]$$

where $\delta_0[\alpha i, \tau i] = 0$ if $[\alpha i] = \tau i$ and $\delta_0[\alpha i, \tau i] = \delta[\alpha i, \tau i]$ if $[\alpha i] \neq \tau i$. Let τ_2 be an α_1-decomposition of A. Then there is an i_0 such that $\tau_2 i_0 = [A]$. First let $\tau i_0 \neq t$ or $\alpha \neq \alpha_1$. Denote τ_3 the decomposition defined as follows: $\mathbf{d}\tau_3 = \mathbf{d}\tau_2$, $\tau_3 i = A = \tau_2 i$ if $\tau_2 i \neq [A]$ and $\tau_3 i = t$ if $\tau_2 i = [A]$. Then $[A, t]$ has two different structures $[\alpha, \tau]$ and $[\alpha_1, \tau_3]$. Secondly let $\tau i_0 = t$ and $\alpha = \alpha_1$. Then $[\alpha i_0] \overset{\rightarrow}{=} t$ and $[\alpha i_0] = [\alpha_1 i_0] \overset{\rightarrow}{=} [A]$. If $\alpha i_0 = A$, then $[\alpha i_0] \to t$ and, by (1), $\delta[A, t] = \delta[\alpha i_0, t] < \delta[A, t]$. This is a contradiction, hence $\alpha i_0 \neq A$ and, because $[\alpha i_0] \overset{\rightarrow}{=} [A]$, we get $[\alpha i_0] \to [A]$. Since $[\alpha i_0] \to A$ for $i \neq i_0$, we have $\alpha \to [\alpha i_0]$ and $[\alpha i_0] \to [A] \Rightarrow \alpha \overset{\rightarrow}{=} [\alpha i_0]$, $\delta[\alpha i_0, t] < \delta[A, t] = n$ which contradicts the choice of n. Thus, the case $\alpha = \alpha_1$ and $\tau i_0 = t$ is impossible.

2.10. Lemma. *Let $\mathscr{L}$ be a non-cyclic language and let*

$$(1) \qquad \mathbf{d}\mathscr{L} \text{ and } \{\alpha, \alpha \in \mathscr{L}A, \alpha \notin \boldsymbol{\sigma}_t\mathscr{L}\} \text{ be finite sets.}$$

Then the set $Q(t) = \{u; u \to t\}$ is finite for each text t.

Proof. Let t be such that $Q(t)$ is infinite. Because $\mathbf{d}\mathscr{L}$ is finite there is an $A \in \mathbf{d}\mathscr{L}$ such that the set $Q(A, t) = \{u, [A] \overset{\rightarrow}{=} u \overset{\rightarrow}{=} t\}$ is infinite. The set

$$N_t = \{[\alpha, i, \tau i], A \in \mathbf{d}\mathscr{L}, [A] \to t, [\alpha, \tau] \in S[A, t], i \in \mathbf{d}\alpha\}$$

is finite. Suppose that for each $[\alpha, i, \tau i] \in N_t$ the set $Q(\alpha i, \tau i)$ is finite. Then the set $P = \bigcup\{Q(\alpha i, \tau i); [\alpha i, i, \tau i] \in N_t\}$ is finite, too. If $u \in Q(A, t)$, $u \neq [A]$ and $u \notin \mathscr{L}A$ then there is an $[\alpha, \xi] \in S[A, u]$ and a u-decomposition ζ of t. By Lemma 6.1, [2], $\tau = \xi \otimes \zeta$ is an α-decomposition of t, and for each $i \in \mathbf{d}\alpha$ we have $[\alpha i] \overset{\rightarrow}{=} \xi i \overset{\rightarrow}{=} \tau i$. Since $[\alpha, i, \tau i] \in N_t$ we get $\xi i \in P$. Hence either $u = [A]$ or $u \in \mathscr{L}A$, $u \in \boldsymbol{\sigma}_t\mathscr{L}$ or $u = \prod_{i=1}^{\lambda\alpha} \xi i$ with $\xi i \in P$. Since $\{\alpha, \alpha \in \mathscr{L}A - \boldsymbol{\sigma}_t\mathscr{L}\}$ and P are finite sets, $Q(A, t)$ is finite which contradicts our assumption. Hence there is an $[\alpha, i, \tau i] \in N_t$ such that the set $Q(A_1, t_1)$ with $A_1 = \alpha i$, $t_1 = \tau i$ is infinite. If $t_1 = t$, then, since $[\alpha i] \to \tau i = t$ and $[A] \Rightarrow \alpha \to t$, we have $[A] \Rightarrow \alpha \overset{\rightarrow}{=} [\alpha i]$. Hence either $\lambda t_1 < \lambda t$ or $[A] \to [A_1]$. Repeating the argument we get a sequence $[A_1, t_1], [A_2, t_2], \ldots$ of grammatical elements such that for each i either $\lambda t_{i+1} < \lambda t_i$ or $[A_i] \to [A_{i+1}]$, $t_i = t_{i+1}$. Of course the first case may occur only finitely many times since $\lambda t_i \geqq 0$.

Hence we get, for some i, $[A_i] \to [A_{i+1}] \to [A_{i+2}] \ldots$ which is impossible since $\mathscr{L}$ is non-cyclic and $\mathbf{d}\mathscr{L}$ is finite. This contradiction shows that no $Q(t)$ is infinite.

2.11. Corollary. *Let $\mathscr{L}$ be a non-cyclic language for which (2.10.1) holds. Then for no grammatical element $[A, t]$ there is an infinite sequence σ such that $[A] \to \to \sigma(i+1) \to \sigma i \to t$ for each $i = 1, 2, \ldots$*

Proof. If for some grammatical element $[A, t]$ such an infinite sequence exists, then either $\sigma i = \sigma j$ for some $i \neq j$ and $\mathscr{L}$ is cyclic or $Q(t)$ is an infinite set.

2.12. Theorem. *Let $\mathscr{L}$ be a non-cyclic language for which (2.10.1) holds. Let $\langle V, R\rangle$ be a pair of transformations V and R defined on $\mathbf{g}\mathscr{L}$ such that for each $g = [A, t] \in \mathbf{g}\mathscr{L}$, each structure $[\alpha, \tau]$ of g and for $\varrho g = [A, Vg]$ we have*

(1) *Rg is a Vg-decomposition of t.*

(2) *If $[A] \Rightarrow t$, then $Vg \in \{A, t\}$.*

(3) *There is an α-decomposition ξ of Vg such that $\tau = \xi \otimes Rg$.*

(4) *If $\varrho g = g$, $i \in \mathbf{d}\alpha$, $\alpha i \neq \tau i$, then $\varrho[\alpha i, \tau i] = [\alpha i, \tau i]$.*

Then ϱ is a reducing transformation.

Proof. According to the definition of a reducing transformation (Def. 9.1, [2]), it is sufficient to prove that the following two conditions are also satisfied:

(5) If $t \neq Vg = [A]$, then $\alpha = [A]$.

(6) For no infinite sequence $g_1, g_2, \ldots$ we have $g_i \in \mathbf{g}\mathscr{L}$, $g_{i+1} = \varrho g_i$, $g_{i+1} \neq g_i$.

First let $t \neq Vg = [A]$. Then, by (3), for each $[\alpha, \tau] \in Sg$, $[A] \overset{\rightarrow}{=} \alpha \overset{\rightarrow}{=} Vg$. If $Vg = = [A]$, then $[A] \overset{\rightarrow}{=} \alpha \overset{\rightarrow}{=} [A]$, which is possible, since $\mathscr{L}$ is a non-cyclic language, only if $\alpha = [A]$. Secondly, suppose there is an infinite sequence $g_1, g_2, \ldots$ such that $g_i \in \mathbf{g}\mathscr{L}$, $\varrho g_i = g_{i+1}$ and $g_{i+1} \neq g_i$. Let $g_i = [A, t_i]$. Then $[A] \to t_i \to t_{i-1} \to t$ for each $i > 1$, but this contradicts Corollary 2.11.

3. RELATIONS BETWEEN LANGUAGES

3.1. Definition. Let $\mathscr{L}$, $\mathscr{L}_1$ be languages. We say that $\mathscr{L}_1$ is a part of $\mathscr{L}$, written $\mathscr{L}_1 \prec \mathscr{L}$, if

(1) $\qquad \mathbf{d}\mathscr{L}_1 \subset \mathbf{d}\mathscr{L}$ and $\mathscr{L}_1 A \subset \mathscr{L} A$ for each $A \in \mathbf{d}\mathscr{L}_1$.

3.2. Lemma. *Let $\mathscr{L}$ be a language and $\mathscr{L}_1$ be a transformation such that the condition (1) is satisfied. Then $\mathscr{L}_1$ is a language and $\mathscr{L}_1 \prec \mathscr{L}$.*

Proof. It is easy to verify that conditions (5.2.1) and (5.2.2), [2] hold with $\mathbf{A} = \mathbf{a}\mathscr{L}$.

3.3. Corollary. *Let $\mathscr{L}$ be a language and $\mathscr{L}_1 \subset \mathscr{L}$. Then $\mathscr{L}_1$ is a language and $\mathscr{L}_1 \prec \mathscr{L}$.*

3.4. Lemma. *Let $\mathscr{L}$, $\mathscr{L}_1$ be languages and $\mathscr{L}_1 \prec \mathscr{L}$. Then:*

(1) *the relations $\mathscr{L}_1$: $\Rightarrow$, $\mathscr{L}_1$: $\twoheadrightarrow$, $\mathscr{L}_1$: $\rightarrow$ are stronger than the relations $\mathscr{L}$: $\Rightarrow$, $\mathscr{L}$: $\twoheadrightarrow$. $\mathscr{L}$: $\rightarrow$, respectively.*

(2) *if $\mathscr{L}_1$ is cyclic, so is $\mathscr{L}$.*

(3) *$\mathbf{g}\mathscr{L}_1 \subset \mathbf{g}\mathscr{L}$ and if $g \in \mathbf{g}\mathscr{L}_1$ then $[\alpha, \tau] \in S_{\mathscr{L}_1}g$ implies $[\alpha, \tau] \in S_{\mathscr{L}}g$,*

(4) *if $\mathscr{L}$ is s.u., so is $\mathscr{L}_1$.*

Proof. The assertion (1) is obvious, (2) and (3) follow from (1) and (4) follows from (3).

3.5. Lemma. Let $\mathscr{L}$, $\mathscr{L}_1$ be languages and $\mathscr{L}$: $[A] \Rightarrow t$ implies $\mathscr{L}_1$: $[A] \rightarrow t$. Then $\mathbf{g}\mathscr{L} \subset \mathbf{g}\mathscr{L}_1$.

Proof. Obvious.

4. CYCLIC LANGUAGES

4.1. Denote

$$O_{\mathscr{L}} = \{A; A \in \mathrm{d}\mathscr{L}, [A] \rightarrow \Lambda\}$$

4.2. Theorem. *Let $\mathscr{L}$ be a language and $\mathscr{L}_0$ be a transformation defined in the following way:*

$$\mathbf{d}\mathscr{L}_0 = \{A; A \in \mathbf{d}\mathscr{L}, \chi A \neq \Lambda\}, \quad \mathscr{L}_0 A = \chi A$$

where

$$\chi A = \{\alpha i, \alpha \in \mathscr{L}A, i \in \mathbf{d}\alpha \text{ and } \alpha j \in O_{\mathscr{L}} \text{ for } j \neq i\}.$$

If $\mathbf{d}\mathscr{L}_0 = \Lambda$, then $\mathscr{L}$ is non-cyclic, if $\mathbf{d}\mathscr{L}_0 \neq \Lambda$ and $[B] \in \mathscr{L}_0 B$ for some $B \in \mathbf{d}\mathscr{L}_0$, then $\mathscr{L}$ is cyclic, if $\mathbf{d}\mathscr{L}_0 \neq \Lambda$ and $B \in \mathscr{L}_0 B$ for no $B \in \mathbf{d}\mathscr{L}_0$, then $\mathscr{L}_0$ is a language and $\mathscr{L}_0$ is cyclic if and only if so is $\mathscr{L}$.

Proof. First let $\mathbf{d}\mathscr{L}_0 = \Lambda$ and let $\mathscr{L}$ be cyclic. Then there exists an $A \in \mathbf{d}\mathscr{L}$ such that $\mathscr{L}$: $[A] \rightarrow [A]$. Since $[A] \notin \mathscr{L}A$, there is an $[\alpha, \tau] \in S_{\mathscr{L}}[A, A]$ such that $\mathscr{L}$: $[A] \Rightarrow \alpha \rightarrow [A]$. Moreover, there is a $i \in \mathbf{d}\alpha$ such that $\mathscr{L}$: $[\alpha i] \rightleftarrows [A]$ and $\alpha j = \Lambda$ for $j \neq i$. We have $\chi A \neq \Lambda$ which contradicts the assumption $\mathbf{d}\mathscr{L}_0 = \Lambda$. Hence $\mathscr{L}$ is non-cyclic.

Secondly let $\mathbf{d}\mathscr{L}_0 \neq \Lambda$ and $B \in \mathscr{L}_0 B$ for some $B \in \mathbf{d}\mathscr{L}_0$. Then it is easy to prove $\mathscr{L}$: $[B] \rightarrow [B]$ and hence $\mathscr{L}$ is cyclic.

Finally, suppose that $\mathbf{d}\mathscr{L}_0 \neq \Lambda$ and $B \in \mathscr{L}_0 B$ for no $B \in \mathbf{d}\mathscr{L}_0$. Clearly, $\mathscr{L}_0$ is a language.

Suppose that $\mathscr{L}$ is cyclic. For proving that $\mathscr{L}_0$ is cyclic it is sufficient to prove the following assertion:

(1) $$\mathscr{L}\colon [A] \rightarrow [B] \text{ implies } \mathscr{L}_0\colon [A] \rightarrow [B].$$

Denote M the set of all grammatical elements $g = [A, t]$ in $\mathscr{L}$ such that if $\lambda t = 1$, then $g \in \mathbf{g}\mathscr{L}_0$. If $\mathscr{L}$: $[A] \Rightarrow [B]$, then $\mathscr{L}_0$: $[A] \Rightarrow [B]$ and hence $[A, [B]] \in M$. Let $g = [A, [B]] \in \mathbf{g}\mathscr{L}$ have a weakly M-regular structure $[\alpha, \tau]$. Then there exists exactly one $i_0 \in \mathbf{d}\alpha$ such that $\tau i_0 \neq \Lambda$. We have $\tau i_0 = [B]$ and $\alpha i \in O_{\mathscr{L}}$ for all $i \neq i_0$, hence $\alpha i_0 \in \mathscr{L}_0 A$. Moreover, since $[\alpha, \tau]$ is a M-regular structure of g, $\mathscr{L}$: $[\alpha i_0] \rightarrow [B]$ implies $\mathscr{L}_0$:$[\alpha i_0] \rightarrow [B]$. Thus, M_0: $[A] \rightarrow [B]$ and $g \in M$. An application of Theorem 6.7, [2] yields $M = \mathbf{g}\mathscr{L}$. Hence (1) holds and $\mathscr{L}_0$ is cyclic.

Now let $\mathscr{L}_0$ be cyclic. In order to prove that $\mathscr{L}$ is cyclic it is sufficient, by Lemma 3.5, to prove that $\mathscr{L}_0$: $[A] \Rightarrow t$ implies $\mathscr{L}$: $[A] \rightarrow t$. Let $\mathscr{L}_0$: $[A] \Rightarrow t$. Then $t = [B]$, $B \in \mathbf{a}\mathscr{L}$ and there are $\alpha \in \mathscr{L}A$, $i_0 \in \mathbf{d}\alpha$ such that $\alpha i_0 = B$ and $\alpha i \in O_{\mathscr{L}}$ for all $i \in \mathbf{d}\alpha$, $i \neq i_0$. By Lemma 5.5, [2], $\mathscr{L}$: $[A] \rightarrow [B] = t$. This completes the proof.

4.3. Remark. If $\mathbf{d}\mathscr{L}$ is a finite set then so is $\mathbf{t}\mathscr{L}_0$ and it is easy to verify whether or not $\mathscr{L}_0$ is cyclic.

The set $O_{\mathscr{L}}$ plays an important role in the theorems of this paper. The following theorem shows a way of constructing the set $O_{\mathscr{L}}$.

4.4. Theorem. *Let* $N_0 = \{A;\ [A] \Rightarrow \Lambda\}$ *and* $N_i = \{A;\ \alpha \in \mathscr{L}A,\ \alpha j \in N_{i-1}$ *for all* $j \in \mathbf{d}\alpha\} \cup N_{i-1}$ *for all integers i. Then*

$$O_{\mathscr{L}} = \bigcup_{i=0}^{\infty} N_i\,.$$

Proof. By induction it is easy to prove $\bigcup_{i=0}^{\infty} N_i \subset O_{\mathscr{L}}$. Now denote M the set of all $g = [A, t] \in \mathbf{g}\mathscr{L}$ such that if $t = \Lambda$, then $A \in \bigcup_{i=0}^{\infty} N_i$. If $[A] \Rightarrow \Lambda$, then $A \in N_0$ and hence $g \in M$. Let $g = [A, \Lambda]$ have a weakly M-regular structure $[\alpha, \tau]$. Then $[\alpha j] \rightarrow \Lambda$ for each $j \in \mathbf{d}\alpha$ and, with respect to M-regularity of $[\alpha, \tau]$, $\alpha j \in \bigcup_{i=0}^{\infty} N_i$. Since $N_i \subset N_{i+1}$, there is an i_0 such that $\alpha j \in N_{i_0}$ for all $j \in \mathbf{d}\alpha$. Hence $A \in N_{i_0+1}$ and $g \in M$. An application of Theorem 6.7, [2] shows that $M = \mathbf{g}\mathscr{L}$ and $O_{\mathscr{L}} \subset \bigcup_{i=0}^{\infty} N_i$.

5. STRUCTURAL UNAMBIGUITY AND WEAK STRUCTURAL UNAMBIGUITY

5.1. Definition. Let $\mathscr{L}$ be a language and $t \in \boldsymbol{\sigma}\mathscr{L}$. Denote $Q_{\mathscr{L}} = \{A;\ A \in \mathbf{d}\mathscr{L},\ \mathbf{t}_t(\mathscr{L}, A) = \{\Lambda\}\}$ and $O_t = \{i;\ i \in \mathbf{d}t,\ ti \in O_{\mathscr{L}}\}$.

5.2. Definition. Let $t \in \boldsymbol{\sigma}\mathscr{L}$ and $N \subset \mathbf{d}t$. Denote $s(t, N)$ the product $\prod \tau$ of the decomposition τ defined by

$$(1) \qquad \lambda\tau = \lambda t,\ \tau i = ti \quad \text{if} \quad i \notin N \quad \text{and} \quad \tau i = \Lambda \quad \text{if} \quad i \in N\,.$$

Moreover, put

(2) $$\psi t = s(t, O_t).$$

The sequence $s(t, N)$ is obtained from t by deleting all symbols ti such that $i \in N$. Similarly, the sequence ψt is obtained from t by deleting all asymbols ti such that $ti \in O_{\mathscr{L}}$.

5.3. Lemma. *If $t \in \sigma\mathscr{L}$ and $N \subset O_t$, then*

(1) $$t \rightrightarrows s(t, N),$$

(2) $$t \rightrightarrows \psi t,$$

(3) $$\psi(s(t, N)) = \psi t.$$

Proof. Straightforward from Definition 5.2.

5.4. Definition. A language $\mathscr{L}$ is said to be regular if $\mathbf{t}_t(\mathscr{L}, A) \neq \Lambda$ for each $A \in \mathbf{d}\mathscr{L}$.

5.5. Definition. We say that a language $\mathscr{L}$ is Λ-structurally unambiguous if the following conditions are satisfied:

(1) If $A \in \mathbf{d}\mathscr{L}$, $\alpha_1 \in \mathscr{L}A$, $\alpha_2 \in \mathscr{L}A$, $\alpha_1 \neq \alpha_2$, then $\psi\alpha_1 \neq \psi\alpha_2$.

(2) If $A \in \mathbf{d}\mathscr{L}$, $\alpha \in \mathscr{L}A$, $i_1 \in \mathbf{d}\alpha$, $i_2 \in \mathbf{d}\alpha$, $i_1 < i_2$, $\alpha i_1 = \alpha i_2$, $\mathbf{t}_t(\mathscr{L}, \alpha i_1) \neq \{\Lambda\}$, then $\alpha j \notin O_{\mathscr{L}}$ at least for one $j \in \langle i_1, i_2 \rangle$.

5.6. Lemma. *Let $\mathscr{L}$ be a regular language. If $\mathscr{L}$ is weakly structurally unambigous, then $\mathscr{L}$ is Λ-structurally unambigous.*

Proof. Conversely, suppose that $\mathscr{L}$ is not Λ-structurally unambiguous. First let condition (5.5.1) be not satisfied. Then there are A, α_1, α_2 such that $A \in \mathbf{d}\mathscr{L}$, $\alpha_1, \alpha_2 \in \mathscr{L}A$, $\psi\alpha_1 = \psi\alpha_2$ and $\alpha_1 \neq \alpha_2$. Moreover, since $\mathscr{L}$ is regular, there exists a $t_1 \in \mathbf{t}_t\mathscr{L}$ such that $\psi\alpha_1 = \psi\alpha_2$. Then, using Lemma 5.3, we have $[A] \Rightarrow \alpha_i \rightrightarrows \psi\alpha_i \rightrightarrows t_1$, $i = 1, 2$. Hence $[A, t_1] \in \mathbf{g}_t\mathscr{L}$ and $[A, t_1]$ is s.a., which contradicts the assumption of Lemma.

Secondly let the condition (5.5.2) be not satisfied. Then there are $A \in \mathbf{d}\mathscr{L}$, $\alpha \in \mathscr{L}A$ and $i_1, i_2 \in \mathbf{d}\alpha$ such that $i_1 < i_2$, $\alpha i_1 = \alpha i_2$, $\mathbf{t}_t(\mathscr{L}, \alpha i_1) \neq \{\Lambda\}$ and $\alpha i \in O_{\mathscr{L}}$ for all $i \in \langle i_1, i_2 \rangle$. Let $\Lambda \neq t_0 \in \mathbf{t}_t(\mathscr{L}, \alpha i_1)$ and $t_i \in \mathbf{t}_t(\mathscr{L}, \alpha i)$ for $i \in (\mathbf{d}t - \langle i_1, i_2 \rangle)$. Put $t = t_1 \times t_2 \times \ldots \times t_{i_1 - 1} \times t_0 \times t_{i_2 + 1} \times \ldots \times t\lambda t$, $\alpha_1 = \alpha_2 = \alpha$, $\mathbf{d}\tau_1 = \mathbf{d}\tau_2 = \mathbf{d}\alpha$, $\tau_1 i = t_2 = \tau_2 i$ for $i \in (\mathbf{d}t - \langle i_1, i_2 \rangle)$, $\tau_1 i_1 = t_0$, $\tau_2 i_2 = t_0$, $\tau_1 i = \Lambda$ for all other $i \in \mathbf{d}\tau_1$ and $\tau_2 i = \Lambda$ for all other $i \in \mathbf{d}\tau_2$. Then $\tau_1 \neq \tau_2$ and $[\alpha_1, \tau_1]$, $[\alpha_2, \tau_2]$ are two different structures of $[A, t] \in \mathbf{g}_t\mathscr{L}$, which contradicts the assumptions of Lemma.

5.7. Lemma. *If $\mathscr{L}$ is a Λ-structurally unambiguous language, then each grammatical element $[A, A]$ is s.u.*

Proof. Let a grammatical element $[A, \Lambda]$ have two different structures $[\alpha_i, \tau_i]$, $i = 1, 2$. Obviously, $\alpha_1 \neq \alpha_2$. If $\alpha_1 \neq [A] \neq \alpha_2$, then $\psi\alpha_1 = \psi\alpha_2 = \Lambda$. If $\alpha_1 = [A] \neq \alpha_2$, then $\psi t = \psi\alpha_2 = \Lambda$. Similarly for the case $\alpha_1 \neq [A] = \alpha_2$. Hence, in all the cases condition (5.5.1) is not satisfied which is the desired contradiction.

5.8. Definition. Let $\mathscr{L}$ be a language. Denote $\mathscr{L}_0$ the transformation defined as follows:

(1) $$\mathbf{d}\mathscr{L}_0 = \mathbf{d}\mathscr{L} - Q_{\mathscr{L}}$$

and for each $A \in \mathbf{d}\mathscr{L}_0$

(2) $\mathscr{L}_0 A = \{s(\alpha, N);\ \alpha \in \mathscr{L}A,\ N \subset O_\alpha,\ s(\alpha, N)\, i \in Q_{\mathscr{L}}$ for no $i \in \mathbf{d}s(\alpha, N)$ and $s(\alpha, N) \neq \Lambda\}$.

5.9. Lemma. *Let $\mathscr{L}$ be a language and $\mathscr{L}_0$ be defined as in Definition* 5.8. *Then*

(1) $\mathscr{L}_0$ is a language,

(2) $\mathbf{a}_t\mathscr{L} = \mathbf{a}_t\mathscr{L}_0$,

(3) $\mathscr{L}_0$: $[A] \to t$ *implies* $\mathscr{L}$: $[A] \to t$.

Proof. To prove (1) it is sufficient to show that $\mathscr{L}_0$: $[A] \Rightarrow [A]$ for no $A \in \mathbf{d}\mathscr{L}_0$. Conversely, suppose that $\mathscr{L}_0$: $[A] \Rightarrow [A]$ for some $A \in \mathbf{d}\mathscr{L}_0$. Then there are $\alpha \in \mathscr{L}A$, $N \subset \mathbf{d}\alpha$, such that $[A] = s(\alpha, N)$. By Lemma 5.3, $\mathscr{L}$: $\alpha \overset{\rightarrow}{=} [A]$. We get $\mathscr{L}$: $[A] \to [A]$ which is impossible since $\mathscr{L}$ is non-cyclic.

Now we prove (2). First let $B \in \mathbf{a}_t\mathscr{L}$. Then there are $A \in \mathbf{d}\mathscr{L}$, $\alpha \in \mathscr{L}A$, $i \in \mathbf{d}\alpha$ such that $\alpha i = B$. We have $A \notin Q_{\mathscr{L}}$ and hence $A \in \mathbf{d}\mathscr{L}_0$. By Definition 5.2 $\psi\alpha = \prod\tau$, where $\lambda\tau = \lambda\alpha$ and $\tau i = [\alpha i]$ if and only if $i \notin O_\alpha$. Since $\alpha i \in \mathbf{a}_t\mathscr{L}$, we have $\alpha i \notin Q_{\mathscr{L}}$ and there is a $v \in \mathbf{d}\psi\alpha$ such that $(\psi\alpha)\, v = \alpha i = B$ which implies $B \in \mathbf{a}_t\mathscr{L}_0$.

Secondly let $B \in \mathbf{a}_t\mathscr{L}_0$. Then either $B \in \mathbf{a}_t\mathscr{L}$ or $B \in \mathbf{d}\mathscr{L} - \mathbf{d}\mathscr{L}_0$. If $B \in \mathbf{d}\mathscr{L} - \mathbf{d}\mathscr{L}_0$, then $B \in Q_{\mathscr{L}}$. Since $B \in \mathbf{a}_t\mathscr{L}_0$, there are $A \in \mathbf{d}\mathscr{L}_0$, $\alpha \in \mathscr{L}_0 A$ and $i \in \mathbf{d}\alpha$ such that $\alpha i = B \in Q_{\mathscr{L}}$. But this is impossible since, by (5.8.2), $\alpha i \in Q_{\mathscr{L}}$ for no $\alpha \in \mathscr{L}_0 A$. Thus, $B \in \mathbf{a}_t\mathscr{L}$ and $\mathbf{a}_t\mathscr{L} = \mathbf{a}_t\mathscr{L}_0$.

In order to prove (3) it is sufficient, by Lemma 3.5, to show that $\mathscr{L}_0$: $[A] \Rightarrow t$ implies $\mathscr{L}$: $[A] \to t$. Suppose $\mathscr{L}_0$: $[A] \Rightarrow t$. By Definition 5.2, there are $\alpha \in \mathscr{L}A$, $N \subset O_\alpha$ such that $\mathscr{L}$: $[A] \Rightarrow \alpha$, $t = s(\alpha, N)$ and, by Lemma 5.3, we have $\mathscr{L}$: $[A] \to \alpha \overset{\rightarrow}{=} t$. This completes the proof.

5.10. Lemma. *Let $\mathscr{L}$ be a non-cyclic regular and Λ-structurally unambiguous language and $\mathscr{L}_0$ be as in Definition* 5.8. *Let $g = [A, t] \in \mathbf{g}_t\mathscr{L}$, $t \neq \Lambda$. Then*

(1) *$g \in \mathbf{g}_t\mathscr{L}_0$ and*

(2) *if g is s.a. in $\mathscr{L}$, then so is it in $\mathscr{L}_0$, too.*

Proof. Let $g = [A, t] \in \mathbf{g}_t\mathscr{L}$ and $[\alpha, \tau] \in S_{\mathscr{L}}g$. Denote

$$N_{\alpha,\tau} = \{i;\ i \in \mathbf{d}\alpha,\ \tau i = \Lambda\}\,, \quad \bar{\alpha} = s(\alpha, N_{\alpha,\tau}) \quad \text{and} \quad \bar{\tau} = s(\tau, N_{\alpha,\tau})\,.$$

As the first step we shall prove the assertion (1). Denote M the set of all $g = [A, t]$ such that $g \in \mathbf{g}_t\mathscr{L}$, $t \neq \Lambda$ imply $g \in \mathbf{g}_t\mathscr{L}_0$. Let $g = [A, t]$ where $\mathscr{L}$: $[A] \Rightarrow t \in \mathbf{t}_t\mathscr{L}$, $t \neq \Lambda$. Then, by definition of $\mathscr{L}_0$, $A \in \mathbf{d}\mathscr{L}_0$, $\mathscr{L}_0$: $[A] \Rightarrow t$ and, by (5.9.2), $g \in \mathbf{g}_t\mathscr{L}_0$ and $g \in M$. We have proved:

(3) If $\mathscr{L}$: $[A] \Rightarrow t \in \mathbf{t}_t\mathscr{L}$, $t \neq \Lambda$, then $[A, [t]]$ is a structure of g in $\mathscr{L}_0$.

Next, let $t \neq \Lambda$ and $g = [A, t] \in \mathbf{g}_t\mathscr{L}$ have a weakly M-regular structure $[\alpha, \tau]$. Then $\bar{\alpha} \neq \Lambda$ and $\mathscr{L}$: $[\bar{\alpha}j] \overset{\rightarrow}{=} \bar{\tau}j$ for each $j \in \mathbf{d}\bar{\alpha}$ and $\Lambda \neq \bar{\tau}j \in \mathbf{t}_t\mathscr{L}$. By definition 5.8, we have $A \in \mathbf{d}\mathscr{L}_0$, $\bar{\alpha} \in \mathscr{L}_0 A$. Moreover, $\bar{\tau}$ is an $\bar{\alpha}$-decomposition of t in $\mathscr{L}$. If $\mathscr{L}$: $[\bar{\alpha}j] \rightarrow \tau j$, then, by weak M-regularity of g, $[\bar{\alpha}j, \bar{\tau}j] \in \mathbf{g}_t\mathscr{L}_0$, hence $\bar{\tau}$ is an $\bar{\alpha}$-decomposition of t in $\mathscr{L}_0$ and $\mathscr{L}_0$: $\bar{\alpha} \overset{\rightarrow}{=} t$. Because $\mathscr{L}_0$: $[A] \Rightarrow \bar{\alpha}$, we get $g = [A, t] \in \mathbf{g}\mathscr{L}_0$ and, by (5.9.2), $g \in \mathbf{g}_t\mathscr{L}_0$. This implies $g \in M$ and, moreover, if $\mathscr{L}_0$: $\bar{\alpha} \rightarrow t$, then $[\bar{\alpha}, \bar{\tau}] \in \in S_{\mathscr{L}}\, g$. By Theorem 6.7, [2], this implies $M = \mathbf{g}\mathscr{L}$ and (3) holds.

Since $M = \mathbf{g}\mathscr{L}$, every structure $[\alpha, \tau]$ of a terminal grammatical element $g = [A, t]$, where $t \neq \Lambda$ and $[A] \neq \alpha$, is weakly M-regular. Therefore, we have:

(4) If $[A, t] \in \mathbf{g}_t\mathscr{L}$, $t \neq \Lambda$, $[\alpha, \tau] \in S_{\mathscr{L}}[A, t]$, $\alpha \neq [A]$, then $[A, t] \in \mathbf{g}_t\mathscr{L}_0$, $\mathscr{L}_0$: $\bar{\alpha} \overset{\rightarrow}{=} t$ and, if $\mathscr{L}_0$: $\bar{\alpha} \rightarrow t$ then $[\bar{\alpha}, \bar{\tau}] \in S_{\mathscr{L}_0}g$.

Now we shall prove the assertion (2). Let $g = [A, t] \in \mathbf{g}\mathscr{L}$, $t \neq \Lambda$ and let g have two different structures $[\alpha i, \tau i]$, $i = 1, 2$, in $\mathscr{L}$.

Let $\alpha_1 = [A]$. Then $\mathscr{L}$: $[A] \Rightarrow \alpha_2$, $\mathscr{L}$: $[A] \Rightarrow t$, $\mathscr{L}$: $\alpha_2 \rightarrow t$ and, since $\mathscr{L}$ is non-cyclic, $\alpha_2 \neq t$. By (5.5.1), $\psi\alpha_2 \neq \psi t$, by (5.3.3) $\psi\bar{\alpha}_2 = \psi\alpha_2$ and hence $\bar{\alpha}_2 \neq t$. By (3) and (4), $[A, [t]]$ and $[\bar{\alpha}_2, \bar{\tau}_2]$ are two different structures of g in $\mathscr{L}_0$. Similarly, if $\alpha_2 = [A]$. It remains the case $\alpha_1 \neq [A] \neq \alpha_2$.

First suppose $\bar{\alpha}_1 \neq \bar{\alpha}_2$ (this will be the case if, in particular, $\alpha_1 \neq \alpha_2$, according to (5.3.3) and (5.5.1)). If $\bar{\alpha}_1 \neq t \neq \bar{\alpha}_2$ then, by (4), g is s.a. in $\mathscr{L}_0$. If $\bar{\alpha}_1 = t \neq \bar{\alpha}_2$, then $[A, [t]]$ and $[\bar{\alpha}_2, \bar{\tau}_2]$ are two different structures of g in $\mathscr{L}_0$. Similarly if $\bar{\alpha}_1 \neq t = \bar{\alpha}_2$.

Secondly we have the case $\alpha_1 = \alpha_2$, $\bar{\alpha}_1 = \bar{\alpha}_2$. Denote $\alpha = \alpha_1$, $\bar{\alpha} = \bar{\alpha}_1$. Since $\tau_1 \neq \tau_2$, we may determine the smallest integer $j \in \mathbf{d}\alpha$ such that $\tau_1 j \neq \tau_2 j$. Suppose $\tau_1 j \neq \Lambda = \tau_2 j$. If we put $j_0 = \mathbf{min}\,\{i;\ i > j,\ \tau_2 i \neq \Lambda\}$, then, since $\bar{\alpha}_1 = \bar{\alpha}_2$, we have $\alpha_2 j_0 = \alpha_2 j$, $\mathbf{t}_t(\mathscr{L}, \alpha_2 j) \neq \{\Lambda\}$, $\alpha_2 i \in O_{\mathscr{L}}$ for $i \in \langle j, j_0\rangle$ which contradicts (5.5.2). Similarly if $\tau_1 j = \Lambda = \tau_2 j$. Finally let $\tau_1 j \neq \Lambda = \tau_2 j$. Then there is an $i \in \mathbf{d}\bar{\alpha}$ such that $\bar{\tau}_1 i \neq \bar{\tau}_2 i$. Hence at least one $\bar{\tau}j$ is not primitive, $\mathscr{L}_0$: $\bar{\alpha} \rightarrow t$ and g has in $\mathscr{L}_0$ two different structures $[\bar{\alpha}, \bar{\tau}_1]$, $[\bar{\alpha}, \bar{\tau}_2]$. Thus in all cases g is s.a. in $\mathscr{L}_0$.

5.11. Lemma. *Let $\mathscr{L}$ be a non-cyclic regular and Λ-structurally unambiguous language and $\mathscr{L}_0$ be as in Definition 5.8. Let $g = [A, t] \in \mathbf{g}_t\mathscr{L}_0$. Then*

(1) $g \in \mathbf{g}_t\mathscr{L}$,

(2) *If g is s.a. in $\mathscr{L}_0$ then so it is in $\mathscr{L}$, too.*

Proof. The assertion (1) follows from (5.9.2) and (5.9.3). Now we shall prove that if a $g \in \mathbf{g}_t\mathscr{L}_0$ has two different structures $[\alpha_i, \tau_i]$, $i = 1, 2$ in $\mathscr{L}_0$, then g is s.a. in $\mathscr{L}$, too. Note that if $\mathscr{L}_0$: $[A] \Rightarrow \alpha \rightarrow t$, then, by (5.9.3), $\mathscr{L}$: $[A] \rightarrow \alpha \rightarrow t$. We distinguish several cases.

Case 1. $\alpha_1 = [A] \neq \alpha_2$. Then $\mathscr{L}_0$: $[A] \Rightarrow t$ and $[\alpha_1, \tau_1] = [A, [t]]$. If $t \in \mathscr{L}A$, then $[A, [t]] \in S_{\mathscr{L}}[A, t]$. Moreover, $\mathscr{L}$: $[A] \rightarrow \alpha_2 \rightarrow t$ and therefore there is a $[\beta, \tau] \in S_{\mathscr{L}}[A, t]$ such that $\beta \neq [A]$, i.e. $[A, t]$ is s.a. in $\mathscr{L}$. Now let $t \notin \mathscr{L}A$. Since $t \in \mathscr{L}_0A$, there are β_1, β_2, N_1 and N_2 such that $N_1 \subset O_{\beta_1}, N_2 \subset O_{\beta_2}$; $\beta_1, \beta_2 \in \mathscr{L}A$. $s(\beta_1, N_1) = t$ and $s(\beta_2, N_2) = \alpha_2$. We have $\mathscr{L}$: $[A] \Rightarrow \beta_1$. $\mathscr{L}$: $[A] \Rightarrow \beta_2$. By (5.3.1) $\mathscr{L}$: $\beta_1 \overset{\rightarrow}{=} t$ and $\mathscr{L}$: $\beta_2 \overset{\rightarrow}{=} \alpha_2$. Since $t \notin \mathscr{L}A$, we have $t \neq \beta_1$ and $\mathscr{L}$: $\beta_1 \rightarrow t$. Thus $\mathscr{L}$: $[A] \Rightarrow \beta_1 \rightarrow t$, $\mathscr{L}$: $[A] \Rightarrow \beta_2 \rightarrow t$. If $\beta_1 \neq \beta_2$, then g has two different structures in $\mathscr{L}$. Let $\beta_1 = \beta_2$. We have $s(\beta_1, N_1) = t \in \mathbf{t}_t\mathscr{L}_0$ (see (5.9.2)), and hence $t = s(\beta_1, O_{\beta_1})$. Because $\mathscr{L}_0$: $\alpha_2 \rightarrow t \in \mathbf{t}_t\mathscr{L}_0$ we have $\alpha_2 \neq t$ and, since $\alpha_2 = s(\beta_2, N_2) = s(\beta_1, N_2)$, we get $N_2 \neq O_{\beta_1}$. Hence $\lambda\alpha_2 > \lambda t$. But this is impossible since, by using the fact that $\alpha_2 \rightarrow t$ and $\Lambda \in \mathscr{L}_0A$ for no $A \in \mathbf{d}\mathscr{L}_0$, we have $\lambda\alpha_2 \leqq \lambda t$. Hence $\beta_1 \neq \beta_2$ and g is s.a. in $\mathscr{L}$.

Case 2. $\alpha_1 \neq [A] = \alpha_2$. The proof follows the same pattern as above.

Case 3. $\alpha_1 \neq [A] \neq \alpha_2$. There are $\beta_1, \beta_2 \in \mathscr{L}A$ and N_1, N_2 such that $N_1 \subset O_{\beta_1}$, $N_2 \subset O_{\beta_2}$, and $\alpha_1 = s(\beta_1, N_1)$, $\alpha_2 = s(\beta_2, N_2)$. We have $\mathscr{L}$: $[A] \Rightarrow \beta_1 \rightarrow t$, and $\mathscr{L}$: $[A] \Rightarrow \beta_2 \rightarrow t$. Let $[i_1, i_2, \ldots, i_{k_1}]$ and $[j_1, j_2, \ldots, j_{k_2}]$ be the increasing sequences of all $i \in (\mathbf{d}\beta_1 - N_1)$, $j \in (\mathbf{d}\beta_2 - N_2)$, respectively. Let τ_3, τ_4 be decompositions defined as follows: $\mathbf{d}\tau_3 = \mathbf{d}\beta_1$, $\tau_3 i = \Lambda$ if $i = i_\nu$ for no $\nu \in \langle 1, k_1 \rangle$ and $\tau_3 i_\nu = \tau_1 \nu$ in the opposite case. Similarly $\mathbf{d}\tau_4 = \mathbf{d}\beta_2$, $\tau_4 i = \Lambda$ if $i = j_\mu$ for no $\mu \in \langle 1, k_2 \rangle$ and $\tau_4 i_\mu = \tau_2 \mu$ in the opposite case. Since $[\alpha_i, \tau_i] \in S_{\mathscr{L}_0}[A, t]$ for $i = 1, 2$, we have $\mathscr{L}_0$: $\alpha_i j \overset{\rightarrow}{=} \tau_i j$ for $j \in \mathbf{d}\alpha$ and hence, according to (5.9.3), $\mathscr{L}$: $\alpha_i j \overset{\rightarrow}{=} \tau_i j$. Hence τ_{i+2} is a β_i-decomposition of t in $\mathscr{L}$.

Obviously g is s.a. in $\mathscr{L}$ if $\beta_1 \neq \beta_2$. Suppose $\beta_1 = \beta_2$. If $\alpha_1 = \alpha_2 = \alpha$ then $[i_1, i_2, \ldots, i_{k_1}] = [j_1, j_2, \ldots, j_{k_2}]$ and $\tau_1 \neq \tau_2$ implies $\tau_3 \neq \tau_4$. Next, if $\alpha_1 \neq \alpha_2$ then $[i_1, i_2, \ldots, i_{k_1}] \neq [j_1, j_2, \ldots, j_{k_2}]$ and, since $O_{\mathscr{L}_0} = \Lambda$, $\tau_3 \neq \tau_4$. Hence in all cases g is s.a. in $\mathscr{L}$.

5.12. Theorem. *Let $\mathscr{L}$ be a non-cyclic regular and Λ-structurally unambiguous language. Let $\mathscr{L}_0$ be defined as in Definition* 5.8. *Then $\mathscr{L}$ is w.s.u. if and only if $\mathscr{L}_0$ is s.u.*

Proof. First let $\mathscr{L}$ be w.s.a. Then there is a $g = [A, t] \in \mathbf{g}_t\mathscr{L}$ which is s.a. By Lemma 5.7, $t \neq \Lambda$. According to Lemma 5.10 $g \in \mathbf{g}_t\mathscr{L}_0$ and g is s.a. in $\mathscr{L}_0$ and $\mathscr{L}_0$ is s.a.

Secondly let $\mathscr{L}_0$ be s.a. Then there is a $g \in \mathbf{g}\mathscr{L}_0$ which is s.a. in $\mathscr{L}_0$. By definition of $\mathscr{L}_0$ and by Lemma 5.9, $\mathscr{L}_0$ is regular and hence, by Theorem 7.7, [2], $\mathscr{L}_0$ is w.s.a. if and only if $\mathscr{L}_0$ is s.a. Hence there is a $g_1 \in \mathbf{g}_t\mathscr{L}_0$ which is s.a. Thus, by Lemma 5.11, $g_1 \in \mathbf{g}_t\mathscr{L}$ and g_1 is s.a. in $\mathscr{L}$ and $\mathscr{L}$ is s.a. This completes the proof.

5.13. Remark. If $\mathscr{L}$ is not Λ-structurally unambiguous then the assertion of Theorem can be false.

Example: $\mathscr{L}$: $A \Rightarrow BCCD$, $C \Rightarrow G$, $C \Rightarrow \Lambda$.

In the following we study the problem of determining whether or not a given language is regular.

5.14. Theorem. *Let $\mathscr{L}$ be a language. Denote $\mu\mathscr{L}$ a transformation defined as follows:*

$$\mathbf{d}\mu\mathscr{L} = \mathbf{d}\mathscr{L} - \{A;\ \alpha \in \mathscr{L}A,\ \alpha \in \boldsymbol{\sigma}_t\mathscr{L}\}$$

$$\mu\mathscr{L}A = \mathscr{L}A \quad \text{for each} \quad A \in \mathbf{d}\mu\mathscr{L}.$$

Then $\mu\mathscr{L}$ is a language and $\mathscr{L}$ is regular if and only if $\mu\mathscr{L}$ is.

Proof. Since $\mu\mathscr{L} \prec \mathscr{L}$, we have, by Lemma 3.2, that $\mu\mathscr{L}$ is a language. First let $\mathscr{L}$ be regular. Denote $\gamma A = \min\{\delta[A, t];\ [A, t] \in \mathbf{g}_t\mathscr{L}\}$ for each $A \in \mathbf{d}\mathscr{L}$. It is easy to prove, by induction by γA, that $\mu\mathscr{L}$ is regular. Secondly let $\mu\mathscr{L}$ be regular and let $A \in \mathbf{d}\mathscr{L}$. If $A \in (\mathbf{d}\mathscr{L} - \mathbf{d}\mu\mathscr{L})$, there is an $\alpha \in \mathscr{L}A \cap \boldsymbol{\sigma}_t\mathscr{L}$ and we have $\mathscr{L}$: $[A] \Rightarrow \alpha \in \boldsymbol{\sigma}_t\mathscr{L}$. Let $A \in \mathbf{d}\mu\mathscr{L}$. Then there is a t such that $\mu\mathscr{L}$: $[A] \rightarrow t \in \mathbf{t}_t\mu\mathscr{L}$. By Lemma 3.4, we have $\mathscr{L}$: $[A] \rightarrow t_i \in \mathbf{t}_t\mu\mathscr{L}$. If $ti \notin \mathbf{a}_t\mathscr{L}$, then $ti \in (\mathbf{d}\mathscr{L} - \mathbf{d}\mu\mathscr{L})$ and there exists an $\alpha \in \mathscr{L}ti \cap \boldsymbol{\sigma}_t\mathscr{L}$. Hence there exists a t_1 such that $\mathscr{L}$: $[A] \rightarrow t_1 \in \mathbf{t}_t\mathscr{L}$ and $\mathscr{L}$ is regular.

5.15. Theorem. *If $\mu\mathscr{L} = \mathscr{L} \neq \Lambda$, then $\mathscr{L}$ is not regular.*

Proof. If $\mathscr{L} \neq \Lambda$ and $\mathscr{L}$ is regular then $\mathbf{g}_t\mathscr{L} \neq \Lambda$, $\mathscr{L}A \cap \boldsymbol{\sigma}_t\mathscr{L} \neq \Lambda$ for some $A \in \mathbf{d}\mathscr{L}$ since otherwise $\mathrm{Cl}_t\Lambda = \Lambda$ and $\mathbf{g}_t\mathscr{L} = \Lambda$ according to Theorem 6.9, [2].

5.16. Corollary. *If $\mathscr{L}$ is a language and $\mathbf{d}\mathscr{L}$ is a finite set then $\mathscr{L}$ is regular if and only if there exists a k such that $\mu^k\mathscr{L} = \Lambda$.* [1])

Acknowledgement. The author wishes to acknowledge VÁCLAV FABIAN, who made many suggestions which improved the presentation.

References

[1] *D. G. Cantor:* On the ambiguity Problem of Backus Systems. J. Assoc. Comp. Mach. *9* (1962), 477—479.

[2] *V. Fabian:* Structural unambiguity of formal languages. Czech. Math. J. *14* (*89*) 1964, 394 to 430.

[3] *R. W. Floyd:* On ambiguity on Phrase Structure Languages. Comm. A.C.M. *5* (1962), 527 to 529.

[4] *S. A. Greibach:* The Undecidability of the Ambiguity Problem for Minimal Linear Grammars. Information and Control *6* (1963), 119—125.

[1]) Denote $\mu^1\mathscr{L} = \mu\mathscr{L}$ and $\mu^{k+1}\mathscr{L} = \mu\mu^k\mathscr{L}$ for $k \geqq 1$.

Резюме

О СТРУКТУРНОЙ ОДНОЗНАЧНОСТИ ФОРМАЛЬНЫХ ЯЗЫКОВ

ЙОСИФ ГРУСКА (Josef Gruska), Братислава

В работе изучается проблема структурной однозначности формальных языков, определенных в работе [2] В. Фабиана. Класс этих языков содержит класс Хомского грамматик типа 2. Настоящая работа приносит, в основном, два результата. Первый из них касается соотношения между структурной однозначностью и цикличностью языков, второй касается соотношения между структурной однозначностью и слабой структурной однозначностью.

В работе доказывается, что циклический язык (т.е. такой язык, в котором существует текст, выводимый из того же текста), не является структурно однозначным и, при некоторых дополнительных условиях, даже не является слабо структурно однозначным. Даются вполне эффективные средства для проверки цикличности языка. Это значит, что достаточно изучать проблему структурной однозначности для нециклических языков. Кроме того, условие нецикличности позволяет упростить некоторые методы и доказательства в работе [2]. В частности, доказывается для определенного класса нецилических языков, содержащего класс Хомского грамматик типа 2, что некоторые условия в определении редуцирующего преобразования (оснвное понятие в [2]), всегда выполнены. Кроме того доказывается, что если язык $\mathscr{L}$ удовлетворяет некоторым условиям, то можно построить такой язык $\mathscr{L}_0$, который является структурно однозначным тогда и только тогда, если язык $\mathscr{L}$ слабо структурно однозначен. (Если эти условия не выполнены, то язык $\mathscr{L}$ не является слабо структурно однозначным.)