

N71-34520

TERMINAL CONTEXT IN  
CONTEXT-SENSITIVE GRAMMARS<sup>†</sup>

by

Ronald V. Book

10-71

Center for Research in Computing Technology

COPY



Harvard University  
Cambridge, Massachusetts 02138

TERMINAL CONTEXT IN  
CONTEXT-SENSITIVE GRAMMARS<sup>†</sup>

by

Ronald V. Book

10-71

Center for Research in Computing Technology

Harvard University

Cambridge, Massachusetts 02138

<sup>†</sup>This research was supported in part by the National Aeronautics and Space Administration under Grant No. NGR-22-007-176 and by the National Science Foundation under Grant No. GJ-803. Part of this research was performed while the author was visiting the Department of System Science, University of California, Los Angeles.

# Terminal Context in Context-Sensitive Grammars

by

Ronald V. Book

## Abstract

If every non-context-free rule of a context-sensitive (with erasing) grammar has as left context a string of terminal symbols and the left context is at least as long as the right context, then the language generated is context-free. If every non-context-free rule of a context-sensitive (with erasing) grammar has as left and right context strings of terminal symbols, then the language generated is context-free.

Introduction.

It is well-known that the family of context-sensitive grammars generates languages which are not context-free and that it is undecidable whether a context-sensitive grammar generates a context-free language. However the mechanism by which the use of context allows a non-context-free language to be generated is not well understood (in fact, the question itself is vague: what does context do for you?). In this paper it is shown that when certain nontrivial constraints are placed on the form of the rules of a context-sensitive (with erasing) grammar only context-free languages will be generated. These constraints involve the use of terminal strings as part of context. The first restriction is that for every non-context-free rule, the left context is a string of terminal symbols which is at least as long as the (arbitrary) right context. (It is shown that the length restriction cannot be weakened.) The second restriction is that both left and right context be strings of terminal symbols.

If one is constructing a context-sensitive grammar to generate some non-context-free language, then one often proceeds as if context can be used to "store and transmit" information. Thus one builds rules so that "messages" or "pulses" are transmitted along a string in the course of the derivation. Sometimes this effect is achieved by building a grammar which imitates the action of a Turing machine; hence, the action of the

read-write head must be imitated as it travels back and forth across the tape.

The "ability to send messages" has not been formalized in such a way as to explain "what context does for you," although some properties of the structure of derivations have been studied [1,5,6]. However this notion does provide an intuitive "handle" for studying some questions and for gaining perspective on some results on context-sensitive grammars and languages.<sup>1</sup> The results established in this paper may be interpreted as constraining the "message-sending" capacity by means of strings of terminal symbols which act as "barriers" when used as context.

There are two somewhat related results in the literature. Hibbard [7] has shown that if  $G = (V, \Sigma, R, X)$  is a grammar and  $<$  as a partial order on  $V$  with the property that for every rule  $Z_1 \dots Z_p \rightarrow Y_1 \dots Y_q$  in  $R$ , there exists  $Y \in \{Y_1, \dots, Y_q\}$  such that for every  $Z \in \{Z_1, \dots, Z_p\}$ ,  $Z < Y$ , then  $L(G)$  is context-free. Ginsburg and Greibach [4] have shown that if  $G = (V, \Sigma, R, X)$  is a grammar such that every rule in  $R$  is of the form  $\rho \rightarrow \theta$  where  $\rho \in (V - \Sigma)^*$  and  $\theta \in V^* \Sigma V^*$ , then  $L(G)$  is context-free. Both of these results may be interpreted as constraining the "message-sending" capacity by erecting "barriers." Neither result appears to imply or be implied by the results established here.

---

<sup>1</sup>There are two other approaches to these problems. In [2,8,9,10] it is shown that for an arbitrary grammar certain types of derivations yield only context-free languages. Also, one can consider "regulating" the application of rewriting rules, such as in matrix grammars, programmed grammars, etc. See [11] for a summary of results in this area.

Section 1.

For the most part the notation used in this paper is that of [3] and the reader is referred to [3] for facts about context-free languages, regular sets, and gsm mappings. However there are certain conventions which need to be emphasized here.

A grammar is quadruple  $G = (V, \Sigma, R, X)$  where  $V$  is a finite set of symbols,  $\Sigma \subset V$  is the set of terminal symbols,  $X \in V - \Sigma$ , and  $R$  is a finite set of rewriting rules (productions) of the form  $\alpha_1 y_1 \dots \alpha_n y_n \alpha_{n+1} \rightarrow \alpha_1 w_1 \dots \alpha_n w_n \alpha_{n+1}$  with each  $\alpha_i \in \Sigma^*$ ,  $y_i \in (V - \Sigma)^*(V - \Sigma)$ ,  $w_i \in V^*$ , and for some  $i$ ,  $w_i \neq y_i$ .<sup>2</sup> If  $\rho \rightarrow \theta \in R$ , then for any  $\alpha, \beta \in V^*$ , write  $\alpha\rho\beta \Rightarrow \alpha\theta\beta$  and say that the rule  $\rho \rightarrow \theta$  is applicable to the string  $\alpha\rho\beta$  and that  $\rho \rightarrow \theta$  transforms  $\alpha\rho\beta$ . A derivation in  $G$  is a sequence  $\Gamma_0, \Gamma_1, \dots, \Gamma_n \in V^*$  such that for  $i = 1, \dots, n$ ,  $\Gamma_{i-1} \Rightarrow \Gamma_i$ . The transitive reflexive closure of  $\Rightarrow$  is  $\xRightarrow{*}$ . The language generated by  $G$  is  $L(G) = \{w \in \Sigma^* \mid X \xRightarrow{*} w\}$ .

If  $G = (V, \Sigma, R, X)$  is a grammar and  $\Gamma_0 \Rightarrow \Gamma_1 \Rightarrow \dots \Rightarrow \Gamma_n$  is a derivation in  $G$ , then a production sequence [5] for this derivation is a sequence of  $n$  ordered pairs,  $\{(B_i, P_i, C_i), (B_i, Q_i, C_i)\}_{i=1}^n$  where for each  $i = 1, \dots, n$ ,  $B_i P_i C_i = \Gamma_{i-1}$ ,  $B_i Q_i C_i = \Gamma_i$ , and  $P_i \rightarrow Q_i \in R$ .

A grammar  $G = (V, \Sigma, R, X)$  is Type 0 (or context-sensitive with erasing) if each rule in  $R$  is of the form  $\alpha Z \beta \rightarrow \alpha \gamma \beta$ ,  $\alpha, \beta, \gamma \in V^*$ ,  $Z \in V - \Sigma$ .

---

<sup>2</sup>For any set  $T$  of symbols,  $T^*$  is the free semigroup with identity generated by  $T$ .

It is well-known that a set  $L$  is recursively enumerable if and only if there is a Type 0 grammar  $G$  such that  $L(G) = L$ .

A grammar  $G = (V, \Sigma, R, X)$  is context-free if each rule in  $R$  is of the form  $Z \rightarrow \gamma$ ,  $Z \in V - \Sigma$ ,  $\gamma \in V^*$ . (In any grammar a rule of this form is called a context-free rule.) A language  $L$  is context-free if and only if there is a context-free grammar  $G$  such that  $L(G) = L$ .

The length of a string  $w$  is denoted by  $|w|$ .

Section 2.

In this section we establish Theorem 1 below and show that the hypothesis cannot be weakened.

Theorem 1. Let  $G = (V, \Sigma, R, X)$  be a Type 0 grammar. If each non-context-free rule in  $R$  is of the form  $\alpha Z\beta \rightarrow \alpha\gamma\beta$  where  $\alpha \in \Sigma^*$ ,  $Z \in V - \Sigma$ ,  $\beta, \gamma \in V^*$ , and  $|\alpha| \geq |\beta|$ , then  $L(G)$  is context-free.

A non-context-free rule in  $R$  has left context which is a terminal string and right context which is no longer than the left (no other restriction is placed on the right context). These two restrictions imply that a "message" cannot be "transmitted" to the left over a string of symbols longer than  $1 + m(m+1)/2$  where  $m = \max\{|\alpha| \mid \alpha Z\beta \rightarrow \alpha\gamma\beta \in R\}$ . Thus any  $w \in L(G)$  can be generated by a derivation such that at each step the transformed symbol is no farther than  $m(m+1)/2$  from the leftmost nonterminal symbol in the string being transformed--hence,  $L(G)$  is context-free. In order to prove this, we first state some definitions and review some facts about grammars and languages.

Let  $G = (V, \Sigma, R, X)$  be a grammar. For  $\Gamma, \Psi \in V^*$ ,  $t \geq 1$ , if  $\Gamma \Rightarrow \Psi$  where  $\Gamma = \alpha Z \dots Y \beta$ ,  $Z$  is the leftmost nonterminal symbol in  $\Gamma$  (i.e.,  $\alpha \in \Sigma^*$  and  $Z \in V - \Sigma$ ),  $Y$  is the transformed symbol in  $\Gamma \Rightarrow \Psi$ , and  $|Z \dots Y| \leq t$ , then  $\Gamma \Rightarrow \Psi$  is  $t$  bounded. If  $\Gamma \Rightarrow \Psi$  is  $t$  bounded, then it is  $r$  bounded for every  $r \geq t$ . A derivation is a  $t$ -bounded derivation if each step is  $t$  bounded. For any  $t \geq 1$   $LEFT(t, G) =$



$\{w \in \Sigma^* \mid \text{there is a } t\text{-bounded derivation } X \Rightarrow \dots \Rightarrow w \text{ in } G\}$ . Clearly, for every  $t \geq 1$ ,  $\text{LEFT}(t,G) \subseteq L(G)$ . In [2,8] it is shown that for any grammar  $G$  and any  $t \geq 1$ ,  $\text{LEFT}(t,G)$  is a context-free language.

For any  $t \geq 1$ , let  $M(t) = t(t+1)/2$ . For a Type 0 grammar  $G = (V, \Sigma, R, X)$ , let  $m = \max\{|\alpha| \mid \alpha Z \beta \rightarrow \alpha \gamma \beta \in R\}$  and let  $M_G = 1 + M(m)$ . If  $G$  satisfies the hypothesis of Theorem 1, then for any rule  $\alpha Z \beta \rightarrow \alpha \gamma \beta \in R$ ,  $|\beta| \leq M(|\beta|) < M_G$  since  $|\alpha| \geq |\beta|$ . To prove Theorem 1 we shall show that  $\text{LEFT}(M_G, G) = L(G)$  so that  $L(G)$  is context-free because  $\text{LEFT}(M_G, G)$  is context-free.

The formal argument proving Theorem 1 is based on the following observation:

- (i) Consider  $\mu \alpha Z x_1 \dots x_n$  where  $\mu, \alpha \in \Sigma^*$ . Suppose one wishes to apply the rule  $\alpha Z \beta \rightarrow \alpha \gamma \beta$  in order to transform  $Z$ . Consider  $x_1 \dots x_n$ . If  $\beta$  is a prefix of  $x_1 \dots x_n$ , then this rule can be applied. If not, then other rules must be applied to  $x_1 \dots x_n$  in order to obtain a string with  $\beta$  as prefix.

- (ii) Suppose that  $\beta = x_1 \dots x_q$  where  $q < n$  and that some rule must be applied to  $x_1 \dots x_n$  in order to transform  $x_{q+1}$  so that eventually  $\beta$  is obtained as a prefix. Further, suppose that the rule to be applied is a context-sensitive rule, so that its left context is a suffix of  $x_1 \dots x_q$  (since  $Z \in V - \Sigma$ ,  $Z$  cannot occur as part of terminal left context). Thus the rule to be applied must have left context which is shorter than  $q + 1 \leq |\beta|$ . But  $|\alpha| \geq |\beta|$  so that the left context of this rule is shorter than  $\alpha$ .
- (iii) By induction, it is seen that if it is possible to apply rules to  $x_1 \dots x_n$  in order to obtain a string with  $\beta$  as prefix, then it is possible to do this by transforming symbols that are no farther than  $1+2+\dots+|\alpha|$  from  $Z$ , so that each step in the resulting derivation is  $M(|\alpha|)+1$  bounded. Since  $M_G \geq 1 + M(|\alpha|)$ , the resulting derivation is  $M_G$  bounded.

The result of these observations is argued formally through two lemmas (the proof of Lemma 1 containing the main argument). The theorem then follows easily.

Lemma 1. Let  $G = (V, \Sigma, R, X)$  be a Type 0 grammar which satisfies the hypothesis of Theorem 1. For any  $\beta \in V^*$  and any  $n \geq 1$ , if  $\Gamma_0 \Rightarrow \dots \Rightarrow \Gamma_n$  is a derivation of length  $n$  in  $G$  such that  $\beta$  is not a prefix of  $\Gamma_0$  but is a prefix of  $\Gamma_n$ , then there exist  $\Pi_0, \dots, \Pi_n \in V^*$  such that  $\Pi_0 = \Gamma_0$ ,  $\Pi_n = \Gamma_n$ ,  $\Pi_0 \Rightarrow \dots \Rightarrow \Pi_n$  is a derivation of length  $n$  in  $G$ , and the step  $\Pi_0 \Rightarrow \Pi_1$  is  $M(|\beta|)$  bounded.

Proof. Since  $\beta$  is not a prefix of  $\Gamma_0$ ,  $\beta \neq \epsilon$ . The proof proceeds by induction on  $|\beta|$ .

(i) For any  $n \geq 1$ , let  $\Gamma_0 \Rightarrow \dots \Rightarrow \Gamma_n$  be a derivation of length  $n$  in  $G$ , let  $\{ \langle (B_i, P_i, C_i), (B_i, Q_i, C_i) \rangle \}_{i=1}^n$  be a production sequence for this derivation, and let  $\Gamma_0 = Y_1 \dots Y_t$ ,  $t \geq 1$ , each  $Y_i \in V$ . If  $|\beta| = 1$ , then  $\beta \in V$ . Since  $\beta$  is not a prefix of  $\Gamma_0 = Y_1 \dots Y_t$ ,  $\beta \neq Y_1$ . But  $\beta$  is a prefix of  $\Gamma_n$  so that there is some step in  $\Gamma_0 \Rightarrow \dots \Rightarrow \Gamma_n$  which transforms  $Y_1$ . Hence,  $Y_1 \in V - \Sigma$ . Thus  $Y_1$  does not occur as part of the left context of any rule in  $R$  (since such context is in  $\Sigma^*$ ). Let  $\Gamma_{k-1} \Rightarrow \Gamma_k$  be the first step which transforms  $Y_1$ . Since  $Y_1$  is the leftmost symbol in  $\Gamma_0$  and since  $\Gamma_{k-1} \Rightarrow \Gamma_k$  is the first step which transforms  $Y_1$ , for each  $i = 1, \dots, k$ ,  $Y_1$  is the leftmost symbol of  $\Gamma_{i-1} = B_i P_i C_i$ . Thus  $Y_1$  cannot be used as part of right context for any of the rules  $P_i \rightarrow Q_i$ ,  $1 \leq i \leq k-1$ . Hence for each  $i = 1, \dots, k-1$ ,  $B_i = Y_1 D_i$  for some  $D_i \in V^*$ ,  $B_k = \epsilon$ , and  $P_k \rightarrow Q_k$  is a context-free rule

with  $P_k = Y_1$ . Thus the derivation  $\Pi_0 \Rightarrow \dots \Rightarrow \Pi_n$  can be constructed by first applying  $P_k \rightarrow Q_k$  to  $Y_1$  in  $\Gamma_0$ , then imitating the derivations  $\Gamma_0 \Rightarrow \dots \Rightarrow \Gamma_{k-1}$  and  $\Gamma_k \Rightarrow \dots \Rightarrow \Gamma_n$ . Since  $Y_1$  is transformed by  $\Pi_0 \Rightarrow \Pi_1$ , this step is  $M(|\beta|) = 1$  bounded. The construction of the derivation  $\Pi_0 \Rightarrow \dots \Rightarrow \Pi_n$  is carried out below.

Let  $D_k = e$ . Thus  $\Gamma_{k-1} = B_{k-1}Q_{k-1}C_{k-1} = Y_1D_{k-1}Q_{k-1}C_{k-1}$  and  $\Gamma_{k-1} = B_kP_kC_k = Y_1C_k$  so that  $C_k = D_{k-1}Q_{k-1}C_{k-1}$ . Construct  $\Pi_0, \dots, \Pi_n$  as follows. Let  $B'_1 = e$ ,  $P'_1 = Y_1 = P_k$ ,  $Q'_1 = Q_k$ , and  $C'_1 = Y_2 \dots Y_t = D_1P_1C_1$ ; let  $\Pi_0 = B'_1P'_1C'_1 = P_kD_1P_1C_1$  and  $\Pi_1 = B'_1Q'_1C'_1 = Q_kD_1P_1C_1$  so that  $\Pi_0 = Y_1 \dots Y_t = \Gamma_0$  and  $\Pi_0 \Rightarrow \Pi_1$ . For  $j = 2, \dots, k$ , let  $B'_j = Q_kD_{j-1}$ ,  $P'_j = P_{j-1}$ ,  $Q'_j = Q_{j-1}$ ,  $C'_j = C_{j-1}$ , and  $\Pi_j = B'_jQ'_jC'_j$ . Then  $B'_2P'_2C'_2 = Q_kD_1P_1C_1 = \Pi_1$ , and for  $j = 2, \dots, k$ ,  $\Pi_j = B'_jQ'_jC'_j = Q_kD_{j-1}Q_{j-1}C_{j-1} = Q_kD_jP_jC_j = B'_{j+1}P'_{j+1}C'_{j+1}$  (since  $\Gamma_{j-1} = B_{j-1}Q_{j-1}C_{j-1} = Y_1D_{j-1}Q_{j-1}C_{j-1} = B_jP_jC_j$ ). Thus  $\Pi_k = Q_kD_{k-1}Q_{k-1}C_{k-1} = Q_kC_k = \Gamma_k$ . For  $i = k+1, \dots, n$ , let  $\Pi_i = \Gamma_i$ . Thus  $\Pi_0 \Rightarrow \dots \Rightarrow \Pi_k = \Gamma_k \Rightarrow \dots \Rightarrow \Pi_n = \Gamma_n$  is a derivation of length  $n$  in  $G$ ,  $\Pi_0 = \Gamma_0$ ,  $\Pi_n = \Gamma_n$ , and the step  $\Pi_0 \Rightarrow \Pi_1$  is  $M(|\beta|) = 1$  bounded.

- (ii) Assume the result for all  $\beta \in V^*$  such that  $|\beta| < r$  for some  $r > 1$ , all  $n \geq 1$ , and all derivations of length  $n$  in  $G$ .

(iii) Consider  $\beta \in V^*$  such that  $|\beta| = r$  and a derivation  $\Gamma_0 \Rightarrow \dots \Rightarrow \Gamma_n$  of length  $n$  in  $G$  such that  $\beta$  is not a prefix of  $\Gamma_0$  and  $\beta$  is a prefix of  $\Gamma_n$ . Let  $\{ \langle (B_i, P_i, C_i), (B_i, Q_i, C_i) \rangle \}_{i=1}^n$  be a production sequence for this derivation. Let  $\Gamma_0 = Y_1 \dots Y_t$ ,  $t \geq 1$ , each  $Y_i \in V$ . If some step of  $\Gamma_0 \Rightarrow \dots \Rightarrow \Gamma_n$  transforms  $Y_1$ , then the argument is just as in (i). Suppose no step of  $\Gamma_0 \Rightarrow \dots \Rightarrow \Gamma_n$  transforms any part of  $Y_1 \dots Y_{q-1}$  for some  $q \leq t$  but some step transforms  $Y_q$ . Since  $\beta$  is not a prefix of  $\Gamma_0 = Y_1 \dots Y_t$ , this implies  $q \leq |\beta| = r$ . Since  $Y_q$  is transformed at some step,  $Y_q \in V - \Sigma$ . Let  $\Gamma_{k-1} \Rightarrow \Gamma_k$  be the first step such that  $Y_q$  is transformed. Since no part of  $Y_1 \dots Y_{q-1}$  is transformed in  $\Gamma_0 \Rightarrow \dots \Rightarrow \Gamma_n$  and  $Y \in V - \Sigma$ , no part of  $Y_1 \dots Y_q$  can serve as part of left context in any step of  $\Gamma_0 \Rightarrow \dots \Rightarrow \Gamma_{k-1}$ .

If the rule applied in  $\Gamma_{k-1} \Rightarrow \Gamma_k$  is a context-free rule, then the argument is just as in (i). If the rule used in  $\Gamma_{k-1} \Rightarrow \Gamma_k$  is a non-context-free rule, say  $\alpha_1 Y_q \alpha_2 \rightarrow \alpha_1 \gamma \alpha_2$ , then by choice of  $q$ ,  $\alpha_1$  must be a suffix of  $Y_1 \dots Y_{q-1}$ , and so  $|\alpha_1| \leq q-1 < q \leq |\beta| = r$ . Thus  $\Gamma_{k-1} = B_k P_k C_k$  where  $P_k = \alpha_1 Y_q \alpha_2$  and  $B_k \alpha_1 = Y_1 \dots Y_{q-1}$ . If  $B_k \alpha_1 Y_q \alpha_2$  is a prefix of  $\Gamma_0$ , then the argument is just as in (i), applying  $\alpha_1 Y_q \alpha_2 \rightarrow \alpha_1 \gamma \alpha_2$  first and noticing that this step is  $q \leq |\beta| < M(|\beta|)$  bounded.

If  $B_{k-1} \alpha_1 Y_q \alpha_2$  is not a prefix of  $\Gamma_0$ , then by choice of  $q$  the application of rules in  $\Gamma_0 \Rightarrow \dots \Rightarrow \Gamma_{k-1}$  transform no symbols in  $\Gamma_0 = Y_1 \dots Y_t$  to the left of  $Y_{q+1}$  and these steps use no part of  $B_{k-1} \alpha_1 Y_q = Y_1 \dots Y_q$  as context. Thus for each  $i = 1, \dots, k-1$ ,  $B_{k-1} \alpha_1 Y_q = Y_1 \dots Y_q$  is a prefix of  $B_i$ , say  $B_i = B_{k-1} \alpha_1 Y_q D_i$  for some  $D_i \in V^*$ . If for some  $i = 1, \dots, k-1$ ,  $\Delta_{i-1} = D_i P_i C_i$ , and  $\Delta_{k-1} = D_{k-1} Q_{k-1} C_{k-1}$ , then  $\Delta_0 \Rightarrow \dots \Rightarrow \Delta_{k-1}$  is a derivation of length  $k-1$  in  $G$  (with production sequence  $\{ \langle (D_i, P_i, C_i), (D_i, Q_i, C_i) \rangle \}_{i=1}^{k-1}$ ) such that  $\alpha_2$  is not a prefix of  $\Delta_0$  but  $\alpha_2$  is a prefix of  $\Delta_{k-1}$ . Since  $\alpha_1 Y_q \alpha_2 \rightarrow \alpha_1 Y \alpha_2$  is a rule in  $R$ ,  $|\alpha_1| \geq |\alpha_2|$ . Also,  $|\alpha_1| < r$ . Hence,  $|\alpha_2| < r$  and  $k-1 < n$  so that the induction hypothesis applies to the derivation  $\Delta_0 \Rightarrow \dots \Rightarrow \Delta_{k-1}$  and the string  $\alpha_2$ , that is, there is a derivation  $\nabla_0 \Rightarrow \dots \Rightarrow \nabla_{k-1}$  of length  $k-1$  in  $G$  such that  $\nabla_0 = \Delta_0$ ,  $\nabla_{k-1} = \Delta_{k-1}$ , and the step  $\nabla_0 \Rightarrow \nabla_1$  is  $M(|\alpha_2|)$  bounded.

Construct the derivation  $\Pi_0 \Rightarrow \dots \Rightarrow \Pi_n$  as follows. For each  $i = 1, \dots, k-1$ , let  $\Pi_i = B_{k-1} \alpha_1 Y_q \nabla_i$ , so that  $\Pi_0 \Rightarrow \dots \Rightarrow \Pi_{k-1}$ . Now  $\Pi_0 = B_{k-1} \alpha_1 Y_q \nabla_0 = B_{k-1} \alpha_1 Y_q \Delta_0 = B_{k-1} \alpha_1 Y_q D_1 P_1 C_1 = B_1 P_1 C_1 = \Gamma_0$  and  $\Pi_{k-1} = B_{k-1} \alpha_1 Y_q \nabla_{k-1} = B_{k-1} \alpha_1 Y_q \Delta_{k-1} = B_{k-1} \alpha_1 Y_q D_{k-1} Q_{k-1} C_{k-1} = B_{k-1} Q_{k-1} C_{k-1} = \Gamma_{k-1}$ . For  $i = k, \dots, n$ , let  $\Pi_i = \Gamma_i$ . Thus,  $\Gamma_0 = \Pi_0 \Rightarrow \dots \Rightarrow \Pi_{k-1} = \Gamma_{k-1}$  and  $\Gamma_k = \Pi_k \Rightarrow \dots \Rightarrow \Pi_n = \Gamma_n$ , so that  $\Pi_0 \Rightarrow \dots \Rightarrow \Pi_n$ . Further, the step  $\Pi_0 \Rightarrow \Pi_1$  is

$|B_{k_1} \alpha_1 Y_q| + M(|\alpha_2|)$  bounded since  $V_0 \Rightarrow V_1$  is  $M(|\alpha_2|)$  bounded. But  $|B_{k_1} \alpha_1 Y_q| = q \leq |\beta|$  and  $|\alpha_2| \leq |\alpha_1| < |\beta|$  so that  $|B_{k_1} \alpha_1 Y_q| + M(|\alpha_2|) \leq M(|\beta|)$ . Thus  $\Pi_0 \Rightarrow \Pi_1$  is  $M(|\beta|)$  bounded.  $\square$

Lemma 2. Let  $G = (V, \Sigma, R, X)$  be a Type 0 grammar which satisfies the hypothesis of Theorem 1. For any  $w \in L(G)$ , if  $X \Rightarrow \Gamma_1 \Rightarrow \dots \Rightarrow \Gamma_n = w$  is a derivation in  $G$  such that there is a least  $t$  where the step  $\Gamma_t \Rightarrow \Gamma_{t+1}$  is not  $M_G$  bounded, then there exist  $\Pi_1, \dots, \Pi_n \in V^*$  such that  $X \Rightarrow \Pi_1 \Rightarrow \dots \Rightarrow \Pi_n$  is a derivation in  $G$ ,  $\Pi_n = w$ , and the derivation  $X \Rightarrow \Pi_1 \Rightarrow \dots \Rightarrow \Pi_{t+1}$  is  $M_G$  bounded.

Proof. Let  $\Gamma_t = a_1 \dots a_p Z_1 \dots Z_q$  where each  $a_i \in \Sigma$ , each  $Z_j \in V$ , and  $Z_1 \in V - \Sigma$ . Since the step  $\Gamma_t \Rightarrow \Gamma_{t+1}$  is not  $M_G$  bounded,  $Z_2 \dots Z_q \neq e$ . Since  $X \xRightarrow{*} \Gamma_t \xRightarrow{*} \Gamma_n$  and  $\Gamma_n = w \in \Sigma^*$ , there is some first step of  $\Gamma_t \Rightarrow \dots \Rightarrow \Gamma_n$  which transforms  $Z_1$ . Let  $\Gamma_j \Rightarrow \Gamma_{j+1}$  be that step, so that  $t < j$ . Since  $Z_1 \in V - \Sigma$ , no part of  $a_1 \dots a_p Z_1$  serves as part of the left context of any rule applied in  $\Gamma_t \Rightarrow \dots \Rightarrow \Gamma_j$ . Since  $Z_1$  is the leftmost nonterminal symbol in each of  $\Gamma_t, \Gamma_{t+1}, \dots, \Gamma_j$ , no part of  $a_1 \dots a_p Z_1$  serves as part of the right context of any rule applied in  $\Gamma_t \Rightarrow \dots \Rightarrow \Gamma_j$ . Thus, as in the proof of Lemma 1, the derivation  $X \Rightarrow \Pi_1 \Rightarrow \dots \Rightarrow \Pi_n$  can be constructed by rearranging the order of application of rules in the derivation  $\Gamma_t \Rightarrow \dots \Rightarrow \Gamma_j$ .

In particular, if the rule applied in  $\Gamma_j \Rightarrow \Gamma_{j+1}$  is a context-free rule  $Z_1 \rightarrow \gamma$  or is a context-sensitive rule  $\alpha Z_1 \beta \rightarrow \alpha \gamma \beta$  and  $\Gamma_t = a_1 \dots a_p Z_1 \dots Z_q = a_1 \dots a_{p-|\alpha|} \alpha Z_1 \beta Z_{|\beta|+2} \dots Z_q$ , then this rule can be applied to  $\Pi_t$  to yield  $\Pi_{t+1} = a_1 \dots a_p \gamma Z_2 \dots Z_q$ , where  $\Pi_i = \Gamma_i$  for  $i = 1, \dots, t, j+1, \dots, n$ , and  $\Pi_{t+2}, \dots, \Pi_{j+1}$  are obtained from  $\Gamma_t \Rightarrow \dots \Rightarrow \Gamma_j$  just as in the proof of Lemma 1. In this case the step  $\Pi_t \Rightarrow \Pi_{t+1}$  is 1-bounded.

If the rule applied in  $\Gamma_j \Rightarrow \Gamma_{j+1}$  is a context-sensitive rule  $\alpha Z_1 \beta \rightarrow \alpha \gamma \beta$  and  $\alpha Z_1 \beta$  is not a prefix of  $\Gamma_t$ , then  $\beta$  is not a prefix of  $Z_2 \dots Z_q$  but is a prefix of  $\delta$  where  $\Gamma_j = \alpha Z_1 \delta$ . Since the rules applied in  $\Gamma_t \Rightarrow \dots \Rightarrow \Gamma_j$  use no part of  $a_1 \dots a_p Z_1$  as either left or right context, this sequence of rules can be applied to  $Z_2 \dots Z_q$  to obtain  $\delta$  in a derivation of length  $j-t$ . By Lemma 1 this derivation can be converted to another derivation  $Z_2 \dots Z_q \Rightarrow \dots \Rightarrow \delta$  of length  $j-t$  such that the first step is  $M(\beta)$  bounded. From the latter derivation (just as in the proof of Lemma 1),  $\Pi_t \Rightarrow \dots \Rightarrow \Pi_j$  is obtained such that  $\Pi_t = \Gamma_t$ ,  $\Pi_j = \Gamma_j$ , and the step  $\Pi_t \Rightarrow \Pi_{t+1}$  is  $M(\beta) + 1 \leq M_G$  bounded. Letting  $\Pi_i = \Gamma_i$  for  $i = 1, \dots, t-1, j+1, \dots, n$ , one obtains derivation  $X \Rightarrow \Pi_1 \Rightarrow \dots \Rightarrow \Pi_n = \Gamma_n = w$  where the derivation  $X \Rightarrow \Pi_1 \Rightarrow \dots \Rightarrow \Pi_{t+1}$  is  $M_G$  bounded.  $\square$



Proof of Theorem 1.

For any  $w \in L(G)$ , consider any derivation  $X \Rightarrow \Gamma_1 \Rightarrow \dots \Rightarrow \Gamma_n = w$  in  $G$ . Either this derivation is  $M_G$  bounded, or applying Lemma 2 at most  $n-2$  times yields a derivation  $X \Rightarrow \Pi_1 \Rightarrow \dots \Rightarrow \Pi_n = w$  in  $G$  which is  $M_G$  bounded. Hence  $w \in \text{LEFT}(M_G, G)$ . Thus  $L(G) \subseteq \text{LEFT}(M_G, G) \subseteq L(G)$ . Since  $\text{LEFT}(M_G, G)$  is context-free,  $L(G)$  is context-free.  $\square$

Let  $G = (V, \Sigma, R, X)$  be a Type 0 grammar which satisfies the hypothesis of Theorem 1, so that if  $\alpha Z \beta \rightarrow \alpha \gamma \beta \in R$ , then  $|\alpha| \geq |\beta|$ . If this restriction on length is weakened, then  $L(G)$  is not necessarily context-free. To see that this is true, note that there exist Type 0 grammars  $G = (V, \Sigma, R, X)$  such that each non-context-free rule is of the form  $ZY \rightarrow Z'Y$  where  $Z \in V - \Sigma$ , and  $Z', Y \in V$ , and such that  $L(G)$  is not context-free.<sup>3</sup> With no loss of generality, assume that in such a grammar  $G$ , each context-free rule is of the form  $Z \rightarrow \gamma$  where  $|\gamma| \leq 2$ . Let  $c$  be a new symbol not in  $V$ ,  $\Sigma_1 = \Sigma \cup \{c\}$ , and  $V_1 = V \cup \{c\}$ . Let

---

<sup>3</sup>This fact has been observed by L.H. Haines and others.

$$\begin{aligned}
R_1 = & \{Z \rightarrow e \mid Z \rightarrow e \in R\} \cup \{Z \rightarrow Y \mid Z, Y \in V, Z \rightarrow Y \in R\} \cup \\
& \{Z \rightarrow cY_1cY_2c, Z \rightarrow cY_1cY_2, Z \rightarrow cY_1Y_2, Z \rightarrow cY_1Y_2c, \\
& Z \rightarrow Y_1cY_2c, Z \rightarrow Y_1Y_2c, Z \rightarrow Y_1cY_2 \mid Z, Y_1, Y_2 \in V, Z \rightarrow Y_1Y_2 \in R\} \cup \\
& \{cZcY \rightarrow cZ'cY \mid Z, Y, Z' \in V, ZY \rightarrow Z'Y \in R\}.
\end{aligned}$$

The grammar  $G_1 = (V_1, \Sigma_1, R_1, X)$  is such that each non-context-free rule is of the form  $\alpha Z\beta \rightarrow \alpha\gamma\beta$  where  $\alpha \in \Sigma_1$  and  $1 = |\alpha| < |\beta| = 2$ . If  $h: \Sigma_1^* \rightarrow \Sigma^*$  is the homomorphism determined by defining  $h(c) = e$  and for every  $a \in \Sigma$ ,  $h(a) = a$ , then clearly  $h(L(G_1)) = L(G)$ . Since  $L(G)$  is not context-free,  $L(G_1)$  is not context-free. Hence, the length restriction cannot be weakened.

It should be noted that the theory of phrase-structure grammars or rewriting systems is symmetric with respect to left and right. Hence, if the restriction on the form of the rules in Theorem 1 is altered to  $\alpha Z\beta \rightarrow \alpha\gamma\beta$  where  $\beta \in \Sigma^*$ ,  $Z \in V - \Sigma$ ,  $\alpha, \gamma \in V^*$ , and  $|\alpha| \leq |\beta|$ , then again  $L(G)$  is context-free.

Section 3.

This section is devoted to showing that if the rules of a grammar have only terminal strings as context, then the language generated is context-free. Formally, this is stated in the following theorem.

Theorem 2. If  $G = (V, \Sigma, R, X)$  is a Type 0 grammar such that every non-context-free rule is of the form  $\alpha Z\beta \rightarrow \alpha\gamma\beta$  where  $\alpha \in \Sigma^*$  and  $\beta \in \Sigma^*$ ,  $Z \in V - \Sigma$ , then  $L(G)$  is context-free.

Intuitively one sees that "messages" cannot be transmitted over sufficiently long terminal strings and here it is only terminal strings which are allowed as context. However, the formal proof of the theorem is based on a lemma which allows one to reduce the length of terminal context, so that by repeated use a context-free grammar is generated.

Notation. For any Type 0 grammar  $G = (V, \Sigma, R, X)$ , let  $L_G = \max\{|\alpha| \mid \alpha Z\beta \rightarrow \alpha\gamma\beta \in R\}$  and  $R_G = \max\{|\beta| \mid \alpha Z\beta \rightarrow \alpha\gamma\beta \in R\}$ .

Lemma 3. Let  $G = (V, \Sigma, R, X)$  be a Type 0 grammar satisfying the hypothesis of Theorem 2. If  $R_G \geq L_G$  and  $R_G \geq 1$ , then one can construct a Type 0 grammar  $G_1$ , a regular set  $T$ , and a gsm  $f$  such that

- (i)  $G_1$  satisfies the hypothesis of Theorem 2.
- (ii)  $R_{G_1} = R_G - 1$  and  $L_{G_1} = L_G$ , and
- (iii)  $f(L(G_1) \cap T) = L(G)$ .

---

<sup>4</sup>See [3] for definitions and facts about regular sets and gsm's.

Proof of Theorem 2.

The theory of phrase-structure grammars is symmetric with respect to left and right, so that Lemma 3 still holds if  $R_G$  and  $L_G$  (and  $R_{G_1}$  and  $L_{G_1}$ ) are interchanged throughout--refer to the result as Lemma 3'. Given  $G$  as in the hypothesis, applying Lemmas 3 and 3'  $m = R_G + L_G$  times yields a sequence  $G_1, \dots, G_m$  of grammars, a sequence  $T_1, \dots, T_m$  of regular sets, and a sequence  $f_1, \dots, f_m$  of gsm's such that for  $i = 1, \dots, m$ ,  $f_i(L(G_i) \cap T_i) = L(G_{i-1})$ , where  $G_0 = G$  and  $G_m$  is context-free [since  $L_{G_m} = R_{G_m} = 0$  implies  $G_m$  is context-free]. Since the family of context-free languages is closed under intersection with regular sets and under gsm mappings, this implies that  $L(G) = L(G_0)$  is context-free.  $\square$

The proof of Lemma 3 rests on the following observation. If  $X \Rightarrow \Gamma_1 \Rightarrow \dots \Rightarrow \Gamma_n$  is any derivation in  $G$  and some  $\Gamma_i$  has a terminal substring of length  $R_G$  of greater, then every step of  $\Gamma_i \Rightarrow \dots \Rightarrow \Gamma_n$  either transforms a symbol to the right of  $\beta$  independent of what is to the left of  $\beta$  or transforms a symbol to the left of  $\beta$  independent of what is to the right of  $\beta$ .

Proof of Lemma 3.

Without loss of generality, assume that every non-context-free rule in  $R$  either has left context but no right context, or has right context but no left context. Partition the set  $R$  as follows: let

$$S_1 = \{Z \rightarrow \gamma \in R \mid Z \in V - \Sigma\},$$

$$S_2 = \{\alpha Z \rightarrow \alpha \gamma \in R \mid \alpha \in \Sigma^*, Z \in V - \Sigma, \text{ and } |\alpha| < R_G\},$$

$$S_3 = \{\alpha Z \rightarrow \alpha \gamma \in R \mid \alpha \in \Sigma^*, Z \in V - \Sigma, \text{ and } |\alpha| = R_G\},$$

$$S_4 = \{Z\beta \rightarrow \gamma\beta \in R \mid Z \in V - \Sigma, \beta \in \Sigma^*, \text{ and } |\beta| < R_G\}, \text{ and}$$

$$S_5 = \{Z\beta \rightarrow \gamma\beta \in R \mid Z \in V - \Sigma, \beta \in \Sigma^*, \text{ and } |\beta| = R_G\}.$$

For each  $\beta \in \Sigma^*$  such that  $|\beta| = R_G$  and there exist  $Z \in V - \Sigma$ ,  $\gamma \in V^*$  with  $Z\beta \rightarrow \gamma\beta \in S_5$ , let  $[\beta]$  be a new symbol, and let  $\Sigma_1$  be the set of such new symbols. For each such  $\beta$ , let  $\beta_{-1}$  be the prefix of  $\beta$  of length  $|\beta| - 1$ , i.e., if  $\beta = \delta a$  where  $a \in \Sigma$ , then  $\beta_{-1} = \delta$ ; since  $|\beta| = R_G$ ,  $|\beta_{-1}| = R_G - 1$ .

Construct the grammar  $G_1 = (V_1, \Sigma \cup \Sigma_1, R_1, X)$  as follows. Let  $V_1 = V \cup \Sigma_1$ . Let  $U_5 = \{Z\beta_{-1} \rightarrow \gamma\beta_{-1}[\beta]\beta_{-1} \mid Z \in V - \Sigma \text{ and } Z\beta \rightarrow \gamma\beta \in S_5\}$ . Let  $R_1 = (R - S_5) \cup U_5$ . From the construction it is clear that  $G_1$  is a Type 0 grammar such that  $L_{G_1} = L_G$  and  $R_{G_1} = R_G - 1$ .

Let  $T = (\Sigma \cup (\bigcup_{[\beta] \in \Sigma_1} (\beta_{-1}[\beta])^* \beta_{-1}[\beta]\beta))^*$  so that  $T$  is a regular set. Let  $f$  be a gsm which yields the identity mapping on  $\Sigma$  until a symbol  $[\beta] \in \Sigma_1$  is scanned. Strings of the form  $[\beta]\beta_{-1}$  are erased and  $f$  returns to its initial mode of operation. Any other operation outputs "garbage". Hence for any  $[\beta] \in \Sigma_1$  and any  $n \geq 0$ ,

$$f(\beta_{-1}([\beta]\beta_{-1})^n[\beta]\beta) = \beta_{-1}f([\beta]\beta_{-1})^{n+1}a = \beta_{-1}a = \beta \text{ where } a \in \Sigma \text{ and } \beta = \beta_{-1}a.$$

Also,  $f(w) = w$  if and only if  $w \in \Sigma^*$ .

Before proving that  $f(L(G_1) \cap T) = L(G)$ , let us informally explain the construction of  $G_1$  and the role of  $T$  and  $f$ . Since  $R - S_5 = R_1 - U_5$  and  $U_5$  is a "copy" of  $S_5$ , it is enough to explain the use of rules in  $U_5$ . When a rule  $Z\beta_{-1} \rightarrow \gamma\beta_{-1}[\beta]\beta_{-1}$  is applied to a string  $EZ(\beta_{-1}[\beta])^t\beta_{-1}F$  to obtain  $E\gamma(\beta_{-1}[\beta])^{t+1}\beta_{-1}F$ , one "guesses" that  $\beta$  is a prefix of  $\beta_{-1}F$  so that an application of  $Z\beta \rightarrow \gamma\beta$  in  $G$  is being imitated. The new occurrence of the symbol  $[\beta]$  serves as a "marker" to indicate this guess and also as a "barrier" so that further steps take place either to the right of  $(\beta_{-1}[\beta])^{t+1}$  or to the left of  $([\beta]\beta_{-1})^{t+1}$ . The "new" copy of  $\beta_{-1}$  in  $\gamma\beta_{-1}[\beta]\beta_{-1}$  is available for use as right context in the future application of some rule. The "old"  $\beta_{-1}$  can still serve as part of left context, since if  $\beta = \beta_{-1}a$ , one still wishes to be able to apply a rule such as  $\delta a\gamma \rightarrow \delta a\psi$  if such a rule is in  $R$ . By hypothesis  $L_G \leq R_G$ , so that  $|\delta a| \leq |\beta| = R_G$ . Hence if one were able to apply this rule in a derivation in  $G$  and if the guess that  $\beta$  is present is correct,

then one can still apply this rule in an "imitating" derivation in  $G_1$  -- the symbol  $[\beta]$  does not cause a conflict since  $|[\beta]\beta_{-1}a| = 1 + |\beta_{-1}a| = 1 + |\beta| = 1 + R_G > L_G$ .

The regular set  $T$  serves as a "filter" to restrict attention to terminal strings which do have substrings in  $(\beta_{-1}[\beta])^{t+1}\beta$  -- i.e., to check that the guesses were correct. The symbol  $[\beta]$  also serves as a marker to tell the gsm to erase the substring  $[\beta]\beta_{-1}$ .

The equality  $f(L(G_1) \cap T) = L(G)$  is established by showing that a derivation in  $G$  resulting in a string in  $L(G)$  can be imitated by a derivation in  $G_1$  resulting in a string in  $L(G_1) \cap T$ , with the role of  $f$  being obvious. The inclusion  $f(L(G_1) \cap T) \subseteq L(G)$  is somewhat more complicated only because a derivation in  $G_1$  resulting in a string in  $L(G_1) \cap T$  may need to be "re-arranged" in order to be imitated in  $G$  -- the "guess" may have been made too soon. The proofs of these inclusions are only sketched since the detailed induction arguments do not yield any additional insight.

Claim 1.  $L(G) \subseteq f(L(G_1) \cap T)$ .

Sketch of the proof. It is sufficient to show that derivations in  $G$  can be imitated in  $G_1$  such that any resulting terminal string is also in  $T$ . Derivations in  $G$  which do not use rules in  $S_5$  can be considered to be derivations in  $G_1$  and  $L(G) \subseteq \Sigma^* \subset T$ . Thus one need be concerned only with those derivations in  $G$  which do use rules in  $S_5$ .

The set  $R_1$  of rules of  $G_1$  contains all the rules in  $R - S_5$  as well as the set  $U_5$ , which is a "copy" of  $S_5$ . A rule in  $U_5$  is of the form  $Z\beta_{-1} \rightarrow \alpha\beta_{-1}[\beta]\beta_{-1}$  where  $Z \in V - \Sigma$ ,  $\beta_{-1} \in \Sigma^*$ ,  $|\beta_{-1}| = R_G - 1$ , and  $Z\beta \rightarrow \gamma\beta$  is in  $S_5$ . Hence the symbol  $[\beta]$  is generated only as part of a string  $\beta_{-1}[\beta]\beta_{-1}$ . If one "imitates" a derivation of  $G$  in  $G_1$ , then to imitate an application of  $Z\beta \rightarrow \gamma\beta$ , the rule  $Z\beta_{-1} \rightarrow \gamma\beta_{-1}[\beta]\beta_{-1}$  in  $U_5$  is applied to a string  $\delta_1 Z(\beta_{-1}[\beta])^k \beta \delta_2$ ,  $k \geq 0$ , to yield  $\delta_1 \gamma(\beta_{-1}[\beta])^{k+1} \beta \delta_2$  so that  $[\beta]$  is generated as part of  $(\beta_{-1}[\beta])^{k+1}\beta$ . Thus it is easy to see that if  $X \Rightarrow \Gamma_1 \Rightarrow \dots \Rightarrow \Gamma_n$  is a derivation in  $G$  with  $\Gamma_n \in \Sigma^*$ , then one can construct a derivation  $X \Rightarrow \Pi_1 \Rightarrow \dots \Rightarrow \Pi_n$  in  $G_1$  with  $\Pi_n \in (\Sigma \cup \Sigma_1)^*$  such that  $\Pi_n \in T$  and  $f(\Pi_n) = \Gamma_n$ . Hence,  $L(G) \subseteq f(L(G_1) \cap T)$ .  $\square$

Claim 2.  $f(L(G_1) \cap T) \subseteq L(G)$ .

Sketch of the proof. It is sufficient to show that derivations in  $G_1$  which generate strings in  $L(G_1) \cap T$  can be imitated in  $G$ . Such derivations in  $G_1$  generate strings in  $L(G_1) \cap \Sigma^*$  if and only if they use no rules from  $U_5$ . But clearly such derivations are already derivations in  $G$ , so  $L(G_1) \cap \Sigma^* \subseteq L(G)$ ; also,  $f(L(G_1) \cap \Sigma^*) = L(G_1) \cap \Sigma^*$ . Hence one need consider only those derivations in  $G_1$  which use at least one application of a rule in  $U_5$ .



Suppose  $X \Rightarrow \Gamma_1 \Rightarrow \dots \Rightarrow \Gamma_n$  is a derivation in  $G_1$  such that  $\Gamma_n \in L(G_1) \cap T$  and such that for some  $k \leq n$  the rule applied at the step  $\Gamma_{k-1} \Rightarrow \Gamma_k$  is a rule in  $U_5$ . Let  $\{ \langle (B_i, P_i, C_i), (B_i, Q_i, C_i) \rangle \}_{i=1}^n$  be a production sequence for  $X \Rightarrow \Gamma_1 \Rightarrow \dots \Rightarrow \Gamma_n$ , and let  $P_k \rightarrow Q_k$  be  $Z\beta_{-1} \rightarrow \gamma\beta_{-1}[\beta]\beta_{-1}$ , where  $Z \in V - \Sigma$  and  $\beta_{-1} \in \Sigma^*$ .

If  $P_k C_k = Z(\beta_{-1}[\beta])^t \beta D$  for some  $t \geq 0$  and  $D \in V_1^*$ , then in the "imitating" derivation  $X \Rightarrow \Pi_1 \Rightarrow \dots \Rightarrow \Pi_n$  in  $G$ , the rule  $Z\beta \rightarrow \gamma\beta$  can be applied at the step  $\Pi_{k-1} \Rightarrow \Pi_k$ . Since  $f((\beta_{-1}[\beta])^{t+1}\beta) = \beta$ , the substring  $(\beta_{-1}[\beta])^{t+1}\beta$  of  $\Gamma_k$  is mapped onto the appropriate substring  $\beta$  of  $\Pi_k$ .

Suppose for some  $t \geq 0$ ,  $Z(\beta_{-1}[\beta])^t \beta_{-1}$  is a prefix of  $P_k C_k$  but for every  $m \geq 0$ ,  $Z(\beta_{-1}[\beta])^m \beta$  is not a prefix of  $P_k C_k$ . Let  $\Gamma_{q-1} \Rightarrow \Gamma_q$  be the first step which generates an occurrence of  $[\beta]$ . Since no rule of  $R_1$  has any symbol of  $\Sigma_1$  on its lefthand side and since  $|\beta_{-1}| = R_{G_1}$ , one loses no generality by assuming that for some  $m$ ,  $q < m \leq n$  every step of  $\Gamma_q \Rightarrow \dots \Rightarrow \Gamma_m$  transforms a symbol to the left of  $[\beta]$  and every step of  $\Gamma_m \Rightarrow \dots \Rightarrow \Gamma_n$  transforms a symbol to the right of  $[\beta]$ . Since  $\Gamma_n \in T$ , the derivation  $\Gamma_m \Rightarrow \dots \Rightarrow \Gamma_n$  produces a substring  $[\beta]\beta$ . Hence in the imitating derivation  $X \Rightarrow \Pi_1 \Rightarrow \dots \Rightarrow \Pi_n$  in  $G$ , the portion  $\Pi_{q-1} \Rightarrow \dots \Rightarrow \Pi_{q-1+n-m}$  imitates  $\Gamma_m \Rightarrow \dots \Rightarrow \Gamma_n$  so that when imitating  $\Gamma_{q-1} \Rightarrow \Gamma_q$ , the string  $\beta$  is available to use as right context. This step is imitated by  $\Pi_{q-1+n-m} \Rightarrow \Pi_{q+n-m}$  and then  $\Gamma_q \Rightarrow \dots \Rightarrow \Gamma_m$  is imitated by  $\Pi_{q+n-m} \Rightarrow \dots \Rightarrow \Pi_n$ .  $\square$

A construction similar to that in the proof of Lemma 3 can be used in conjunction with the result of Ginsburg and Greibach cited in Section 1 to establish the following generalization of Theorem 2.

Corollary. If  $G = (V, \Sigma, R, X)$  is a grammar such that every non-context-free rule is of one of the forms

$$(i) \quad \alpha\rho \rightarrow \alpha\theta \quad \text{where } \alpha \in \Sigma^*, \quad |\alpha| \geq 1, \quad \text{and } \rho \in (V - \Sigma)^*,$$

$$(ii) \quad \rho\beta \rightarrow \theta\beta \quad \text{where } \beta \in \Sigma^*, \quad |\beta| \geq 1, \quad \text{and } \rho \in (V - \Sigma)^*,$$

then  $L(G)$  is context-free.

There is one further restriction which generalizes the hypothesis of both Theorems 1 and 2, that is, let  $G = (V, \Sigma, R, X)$  be a Type 0 grammar such that every non-context-free rule is of one of the forms

$$(i) \quad \alpha Z\beta \rightarrow \alpha\gamma\beta \quad \text{where } \alpha \in \Sigma^*, \quad Z \in V - \Sigma, \quad \beta, \gamma \in V^*, \quad \text{and } |\alpha| \geq |\beta|,$$

$$(ii) \quad \alpha Z\beta \rightarrow \alpha\gamma\beta \quad \text{where } \beta \in \Sigma^*, \quad Z \in V - \Sigma, \quad \alpha, \gamma \in V^*, \quad \text{and } |\alpha| \leq |\beta|.$$

Thus it is required that either left or right context be a terminal string and in either case, the terminal context has length at least as great as the other context. We conjecture that this restriction forces the language generated to be context-free.

References

1. R. Book, Time-bounded grammars and their languages, J. Computer and System Sciences 5 (1971), to appear.
2. R. Evey, The Theory and Application of Pushdown Store Machines, Ph.D. Dissertation Harvard Univ., 1963. Also appears as Math. Linguistics and Automatic Translation NSF-10 (1963), Computation Lab., Harvard Univ., Cambridge, Mass.
3. S. Ginsburg, The Mathematical Theory of Context-free Languages, McGraw-Hill, 1966.
4. S. Ginsburg and S.A. Greibach, Mappings which preserve context-sensitive languages, Information and Control 9 (1966), 563-582.
5. A. Gladkii, On the complexity of derivations in phrase structure grammars, Algebr i Logika Sem. 3 (1964), 29-44.
6. T. Griffiths, Some remarks on derivations in general rewriting systems, Information and Control 12 (1968), 27-54.
7. T. Hibbard, Scan Limited Automata and Context Limited Grammars, Ph.D. Dissertation, Univ. of California, Los Angeles, 1966.
8. G.H. Matthews, Discontinuity and asymmetry in phrase structure grammars, Information and Control 6 (1963), 137-146.
9. G.H. Matthews, A note on asymmetry in phrase structure grammars, Information and Control 7 (1964), 360-365.
10. G.H. Matthews, Two-way languages, Information and Control 10 (1967), 111-119.
11. A. Salomaa, Formal Languages, in preparation.