Linguistic Inquiry Volume VI Number 4 (Fall, 1975) 533–554.

D. Terence Langendoen

Finite-State Parsing of Phrase-Structure Languages and the Status of Readjustment Rules in Grammar

1. A Theorem about Finite-State Parsing of Phrase-Structure Languages

In Chomsky (1959a,b) and independently in Bar-Hillel, Perles, and Shamir (1961), it is proved that a context-free phrase-structure (CFPS) language L can be generated by a finite-state (FS) grammar if and only if there is a noncenter-embedding (NCE) CFPS grammar that generates L; we call this result the Chomsky–Bar-Hillel theorem.[1] This theorem asserts that, given a NCE–CFPS grammar G, there is a weakly equivalent FS grammar G', a grammar that can also be thought of as a finite acceptor (FA) for $L(G)$; that is, a device that accepts all and only all of the sentences of $L(G)$. Chomsky (1959a), moreover, provides an algorithm for constructing FAs for sentences generated by arbitrary Chomsky-normal-form (CNF)[2] CFPS grammars with up to any fixed finite degree n of center embedding (CE).

Now let $L_B(G)$ be the language consisting of the set of structural descriptions (P-markers) that a CFPS grammar G associates with the sentences it generates. It is easy to see that $L_B(G)$ is a CFPS language; a CFPS grammar G_B that generates $L_B(G)$ can be constructed from G by replacing each production $A \rightarrow \omega$ in G by the production $A \rightarrow [_A\omega]_A$.[3] Can we find a FA for $L_B(G)$? The answer is negative just in case $L_B(G)$ is an infinite language; it does not matter if G is NCE or even FS, or if G generates only a finite language. If G associates infinitely many P-markers with the sentences it generates, $L_G(B)$ is not a FS language and hence cannot be accepted by a FA. Given that there is no FA for $L_B(G)$, there can also be no finite transducer (FT) that takes sentences of $L(G)$ as input and gives as output their P-markers with respect to G (such a FT may be called a finite parser (FP) for G). In other words, if we try to extend the Chomsky–Bar-Hillel theorem to strong equivalence in the direct sense that we determine the subclass of CFPS grammars for which there are FPs, we find that that

[1] For terminology and notation, see the aforementioned papers or Chomsky (1963). I thank C. Kaniklidis and N. Chomsky for helpful discussions, and two reviewers whose painstaking critical efforts have resulted in numerous improvements.

[2] The productions of a Chomsky-normal-form CFPS grammar are all of the form $A \rightarrow B\ C$, or $A \rightarrow a$ where A, B, and C are nonterminal elements and a is a terminal element.

[3] We assume that $[_A$ and $]_A$ are added to the terminal vocabulary of G_B for each nonterminal element A in G. The labels on either the left or right parentheses may be omitted without ambiguity. Either left or right parentheses may be omitted entirely under certain conditions; see footnote 4 for details.

subclass contains just those CFPS grammars that associate only a finite number of P-markers with the set of sentences they generate. The theorem and an outline of its proof are now given.

Theorem 1. If G is a CFPS grammar that generates a language $L(G)$ and associates with the sentences of $L(G)$ a set $L_B(G)$ of P-markers, then there is a FS grammar G' that generates $L_B(G)$ if and only if $L_B(G)$ is finite.

Proof (outline). Clearly, if $L_B(G)$ is finite, there is a FS grammar G' that generates it, since there is a FS grammar for any finite language.

Suppose $L_B(G)$ is infinite. Then, by the method of the "*uvwxy*" theorem (Hopcroft and Ullman 1969, 57–59), for every nonnegative integer i, there are strings u, v, w, x, y (all except w possibly null) and terminal elements $[_A$ and $]_A$ such that $u([_Av)^i[_Aw]_A(x]_A)^iy$ is a sentence of $L_B(G)$, and such that u, v, w, x, y do not contain unmatched occurrences of $[_A$ and $]_A$. Suppose that the grammar G' that generates $L_B(G)$ is FS. Then, by the Nerode-Myhill theorem (Rabin and Scott 1959, Theorem 2), there are distinct m and n such that $([_Av)^m$ and $([_Av)^n$ are equivalent, and hence such that $u([_Av)^m[_Aw]_A(x]_A)^my$ is in $L_B(G)$ if and only if $u([_Av)^n[_Aw]_A(x])^my$ is in $L_B(G)$. Since the former must be in $L_B(G)$, then so must the latter. But the latter cannot be in $L_B(G)$, since it contains an unequal number of occurrences of $[_A$ and $]_A$. Therefore no grammar G' that generates $L_B(G)$ can be a FS grammar. This completes the outline of the proof of the theorem.[4]

[4] As Chomsky (1963, 367) observes, however, it may not be necessary to construct a full P-marker in order to represent unambiguously the structural descriptions of sentences generated by a CFPS grammar. Under certain conditions, either the left or right parentheses may be omitted, yielding structures from which full P-markers can be unambiguously and effectively determined. Let us call a P-marker with left (respectively right) parentheses suppressed a right- (respectively, left-) semi-P-marker (RSP-marker; respectively, LSP-marker). Let $L_R(G)$ be the set of RSP-markers that a CFPS grammar G associates with the sentences it generates, and let $L_L(G)$ be the set of LSP-markers that G associates with the sentences it generates.

I. Suppose G is FS; that is, suppose it is either left linear (LL) or right linear (RL). If G is LL, then $L_R(G)$ is a FS language that unambiguously represents $L_B(G)$, and if G is RL, then $L_L(G)$ is a FS language that unambiguously represents $L_B(G)$. This is so, because if G is LL, then all the left parentheses in sentences of $L_B(G)$ occur at the beginning of the corresponding sentences of $L(G)$, and any structural ambiguity in sentences of $L(G)$ will be represented in the positions of or labels on the right parentheses in $L_B(G)$. Therefore, $L_R(G)$ unambiguously represents $L_B(G)$. To form a FS grammar G' that generates $L_R(G)$, replace each production $A \rightarrow (B)\ x$ in G by the corresponding production $A \rightarrow (B)\ x]_A$. By the same argument, if G is RL, $L_L(G)$ is a FS language that unambiguously represents $L_B(G)$.

II. Suppose G is an unambiguous CFPS grammar. Then, both $L_L(G)$ and $L_R(G)$ unambiguously represent $L_B(G)$. Otherwise, there is some sentence s in $L_L(G)$ or in $L_R(G)$ that ambiguously represents sentences s' and s'' in $L_B(G)$. But then, there is a single sentence s''' in $L(G)$ that has two P-markers with respect to G, namely s' and s'', contrary to assumption.

III. Suppose G is a NCE–CFPS grammar. Then, either it generates a finite language, or it permits subderivations of the type $A \overset{*}{\rightarrow} x\ A$ or $B \overset{*}{\rightarrow} B\ y$, or both (but not such that there are subderivations of the type $C \overset{*}{\rightarrow} w\ C\ z$, where both w and z are nonnull). If subderivations of the first type are permitted, we say that G is (or permits) right embedding (RE); similarly, if subderivations of the second type are permitted, we say that G is (or permits) left embedding (LE).

Suppose G is unambiguous. From II, it follows that if G is RE, then $L_L(G)$ unambiguously represents $L_B(G)$, and that if G is LE, then $L_R(G)$ unambiguously represents $L_B(G)$. Moreover, if G is RE and not LE, then there is a RE–CFPS grammar G_L that generates $L_L(G)$, and if G is LE and not RE, then there is a LE–CFPS grammar G_R that generates $L_R(G)$. From the Chomsky–Bar-Hillel theorem, then, it follows that if G is an unambiguous strictly RE–CFPS grammar, then there is a FS grammar G' that generates $L_L(G)$, which unambig-

2. On Strong Equivalence between Finite Transducers and Phrase-Structure Grammars

As we previously noted, the first of Chomsky's proofs of the Chomsky–Bar-Hillel theorem is based on an algorithm for constructing a FA that accepts $L(G)$, given a NCE–CNF–CFPS grammar G (with certain additional limitations on its form). Moreover, there is an effective, one-one mapping Φ from the P-markers of each sentence generated by G onto the sequence of states that this FA goes through in accepting that sentence. Suppose, then, we equip this FA with an output tape on which it prints the sequences of states that it goes through when accepting each sentence x in $L(G)$. Such a device is a FT T that accepts x and assigns it elements $y_1, \ldots, y_n$ that the mapping Φ effectively and uniquely associates with the P-markers $z_1, \ldots, z_n$ that x has with respect to G. Briefly, we say that T generates $(x, \Phi(z_1, \ldots, z_n))$. We may now say, following Chomsky (1963, 396), that T and G are strongly equivalent if and only if T generates $(x, \Phi(z_1, \ldots, z_n))$ just in case G generates x with the P-markers $z_1, \ldots, z_n$ and no others. Let us call the FT, obtainable by Chomsky's algorithm from a NCE–CNF–CFPS grammar G, $\Psi(G)$. We have, then, the following theorem (Chomsky 1963, Theorem 34).

Theorem 2. There is an effective procedure Ψ such that, given a NCE–CNF–CFPS grammar G, $\Psi(G)$ is a FT that is strongly equivalent to G.

However, we see that the effective, one-one mapping Ψ cannot in general be a mapping that itself can be carried out by a FT, for if it could be so carried out, one

---

uously represents $L_B(G)$; and if G is an unambiguous strictly LE–CFPS grammar, then there is a FS grammar G' that generates $L_R(G)$, which unambiguously represents $L_B(G)$.

This result cannot be strengthened to unambiguous NCE–CFPS grammars in general, since such grammars may be both RE and LE, and if so, then any grammar that generates either $L_L(G)$ or $L_R(G)$ will not be FS.

IV. Even if G is ambiguous, it may be the case that either $L_L(G)$ or $L_R(G)$ unambiguously represents $L_B(G)$. Suppose, for example, that there is a variant of English that can be generated by a CFPS grammar G_1, and that the only parsing ambiguity occurs in sentences of the type illustrated in (i).

(i) a. Ann said that Bill entered quietly.
b. $[_S$Ann said $[_S$that Bill entered quietly$]_S]_S$
c. $[_S$Ann said $[_S$that Bill entered$]_S$ quietly$]_S$

Clearly, $L_R(G_1)$ unambiguously represents $L_B(G_1)$. On the other hand, suppose that there is another variant of English that can be generated by a CFPS grammar G_2, and that the only parsing ambiguity occurs in sentences of the type (ii). (The symbol $\bar{N}$ may be read "noun phrase"—throughout this article, we use a simplified version of Chomsky's bar notation (Chomsky 1970) for representing phrase categories.)

(ii) a. The happy young children's teacher arrived.
b. $[_{\bar{N}}[_{\bar{N}}$the happy young children$]_{\bar{N}}$'s teacher$]_{\bar{N}}$ arrived
c. $[_{\bar{N}}$the happy $[_{\bar{N}}$young children$]_{\bar{N}}$'s teacher$]_{\bar{N}}$ arrived

Clearly, $L_L(G_2)$ unambiguously represents $L_B(G_2)$. Ambiguity of the type represented in (i) arises from the joint possibility of both RE and CE on the same recursive category, while ambiguity of the type illustrated in (ii) arises from the joint possibility of both LE and CE on the same recursive category. Thus, it would appear that if a CFPS grammar G does not permit both RE and LE, then either LSP-markers or RSP-markers are sufficient to represent unambiguously the structural descriptions of sentences with respect to G. It is the joint possibility of both LE and RE in CFPS English-like languages that contain sentences of both types (i) and (ii) that obligates the use of full P-markers to represent unambiguously the structural descriptions of the sentences of those languages with respect to the grammars that generate them.

could construct a FP for G from $\Psi(G)$ and Φ, since FTs are closed under composition. From Theorem 1, we know that such a FP exists only if the set of P-markers associated with the sentences of $L(G)$ is finite. Thus, even though a general-purpose FS device (such as the human mind may be assumed to be) could internalize the procedure Ψ, it would not always be able to carry out Ψ in the course of associating P-markers with the sentences it accepts.

Now, the procedure Ψ of Theorem 2 can be thought of either as a model of how a CFPS grammar is represented in the mind (that is, as a model of competence),[5] or as a model of an aspect of linguistic performance that incorporates a CFPS grammar as a component. If Ψ is thought of as a model of competence, then from Theorem 2 we conclude that it is fully adequate for purposes of representing linguistic competence that a linguist would normally represent in the form of a CFPS grammar. However, there are at least three reasons for rejecting the view that Ψ is a model of competence, and correspondingly for accepting the view that it is a model of performance. First, if Ψ is a model of competence, then the theory of grammar must be formalized in terms of augmented FS grammars (the augmentation being required to deal with CE), rather than in terms of the conceptually more elegant theory of CFPS grammars. Second, given a linguistically adequate theory of competence with a CFPS base constructed in terms of Ψ, with transformations defined on the structures generated by that base, transformations would have to be defined as operations on sets of state sequences that Ψ goes through in generating a base string. While this could be done in principle, the characterization of the relations "factor of", "analyzable as", and "identical to", which are required by that theory of transformations, would be exceedingly complex and unnatural. Third, if Ψ is a model of competence, it is not at all clear how one would construct a reasonable model of performance that incorporated it; it would appear that the theory of performance would be totally independent of the theory of competence.

On the other hand, if Ψ is a model of performance, the base component (in the model of competence) could be given directly in the form of a CFPS grammar, and the basic relations of the transformational component could be directly defined as relations on strings (assuming, following the formalization of transformational grammar given in Peters and Ritchie (1973), that the base component directly generates strings with labeled brackets in them in the manner of G_B above). Further, the theory of performance is given directly by Ψ, which incorporates the grammar in the sense that the grammar is used in the construction of the states and instructions of Ψ (a matter which we take up in detail below in section 3).

However, if Ψ is thought of as a model of linguistic performance, then one is interested not in Chomsky's indirect notion of strong equivalence, which requires the

[5] We know that this hypothesis about competence is false, but only because the generative capacity (both weak and strong) of CFPS grammars is too small to be of linguistic interest. The class of CFPS grammars can still usefully serve, however, as a first approximation to the class of linguistically significant grammars.

carrying out of a mapping that in general exceeds the capacity of any FS device, but rather in the direct notion of strong equivalence introduced in section 1 above. We say that direct strong equivalence holds between a FT T and a CFPS grammar G if and only if T generates $(x; z_1, \ldots, z_n)$ just in case G generates x with the structural descriptions $z_1, \ldots, z_n$ and no others. From Theorem 1, it follows that there is no FT that is directly strongly equivalent to a CFPS grammar that associates an infinite number of P-markers with the set of sentences it generates.

From the conception of Ψ as a model of performance, moreover, we see how extremely limited the ability of FS devices is for parsing sentences generated by CFPS grammars. Since infinite numbers of phrase-markers can arise from CFPS grammars only by recursion (i.e. only by subderivations of the type $A \overset{*}{\rightarrow} \phi A \psi$), we conclude that it is recursiveness in general (not just CE) that gives rise to the inability of FPs to parse sentences generated by CFPS grammars. This observation, that it is recursiveness in general that limits the ability of a FP to associate P-markers with the sentences it accepts, leads us to conclude that if we wish to minimally augment the FP with a push-down store (PDS) so as to increase its direct strong generative capacity (in the manner in which Chomsky proposed to minimally, or in his words, optimally, augment the FA that accepts sentences generated by a NCE–CFPS grammar, to enable it to deal with up to some fixed finite degree n of CE), we would have to (ultimately) keep track of each recursion on the PDS. An algorithm for constructing a minimally augmented finite parser (MAFP) for any normal-form (NF)[6] CFPS grammar is presented in the next section.

3. Construction of a MAFP for a NF–CFPS Grammar

Let G be a NF–CFPS grammar that has the single axiom $\#S\#$, the nonterminal vocabulary U, and the terminal vocabulary V, and that generates $L(G)$ and associates with the sentences of $L(G)$ the set of P-markers $L_B(G)$. We must also either impose the restriction that a nonterminal symbol cannot appear more than once on the right-hand side of production of G, or allow that possibility but construct M in terms of productions in a new grammar G', in which symbols that recur on the right-hand side in productions of G are replaced by distinct symbols. One obvious way to do this is to form G' by replacing each production $A \rightarrow X_1\ B \cdots B X_{n+1}$ in G by the production $A \rightarrow X_1\ B^1 \cdots B^n\ X_{n+1}$, and for each rule $B \rightarrow \omega$ in G, to add to G' the productions $B^i \rightarrow \omega$ for each i, $1 \leqq i \leqq n$, and to use G' in the construction of the MAFP for G.[7]

[6] The productions of a NF–CFPS grammar are all of the form $A \rightarrow X$, or $A \rightarrow a$, where A is a nonterminal element, X is a nonnull string of nonterminal elements, and a is a terminal element.

[7] We are using the standard definition of CFPS grammar, which excludes the possibility of there being infinitely many production rules. However, the definition can readily be extended so as to include that possibility, so long as those productions can be indirectly represented by a finite number of finite rule schemata, and so long as from the schemata, the infinite set of rules they abbreviate can be enumerated by a procedure that

For ease in formulating the algorithm for constructing a MAFP for G, however, we accept the restriction on G that no nonterminal symbol may appear more than once on the right-hand side of each production of G, and note simply that this restriction is imposed for expository reasons only and that in principle it may be dispensed with.

We construct the MAFP M for G as follows. First we give a procedure for enumerating the set of states Σ of M. Let the set K consist of all rooted subsequences of labels on brackets of sentences of $L_B(G)$, such that no label occurs more than once. That is, $K = \{A_1 \cdots A_m \mid m \geqq 1,\ A_1 = S$, and for all i, j such that $1 \leqq i < j \leqq m$, $A_i \neq A_j$, and there are strings X, Y such that $A_i \rightarrow X\ A_{i+1}\ Y$ is a production of $G\}$. Σ consists of the initial state S, the final state F, and all members of K subscripted by L or R (for "left" and "right" respectively). Thus, if $B_1 \cdots B_n$ is in K, then $(B_1 \cdots B_n)_L$ and $(B_1 \cdots B_n)_R$ are both in Σ. M has a reading head that scans from left to right an input tape, on which are written strings of the form $\#\sigma\#$, where σ is a string of elements of V together with left and right brackets labeled with elements of U (i.e. a string over the terminal alphabet of $L_B(G)$). The device M is also equipped with a PDS on which it may print or erase a string of elements of U, followed by the designated boundary symbol *, and an auxiliary reading head that scans the most recently printed string on the PDS. We say that M is in its initial configuration if it is in the state S, reading the first # on the input tape, and the PDS is blank. We say that M is in its final configuration if it is in the state F, reading the second # on the input tape, and the PDS is blank. Instructions in M are all in the form $Q \rightarrow I;\ x\ R$, where Q and R are members of Σ, x is a (possibly null) string of symbols on the input tape that the device must

---

is itself no more powerful than that of the theory of CFPS grammar (with a finite number of rules). The reason that the enumeration procedure must not be more powerful than the theory of CFPS grammar is to ensure that the weak generative capacity of CFPS grammars with possibly infinitely many productions is no greater than that of CFPS grammars with only a finite number of productions. For example, without this restriction, the well-known non-CFPS language $L = \{a^n b^n c^n : n > 0\}$ could be generated by the infinite CFPS grammar $S \rightarrow a^n b^n c^n$, $n > 0$ (the procedure that enumerates the rules abbreviated by this schema requires the power of a context-sensitive phrase-structure grammar).

The extension of the theory of CFPS grammars to include systems with infinitely many productions is linguistically motivated by consideration of the constituent structure of coordinate constructions in natural languages. However, the schemata that are so motivated are all of the general form $A \rightarrow \chi\psi^n\omega$, $n > 0$, for which the procedures that enumerate the productions, given these schemata, can all be carried out by FS devices (Chomsky and Schützenberger 1963, 133; Chomsky 1965, 224). Given such schemata, however, the procedure for constructing a grammar G' that does not contain repetitions of the same nonterminal symbol on the right-hand side of productions would not succeed, since it would require that G' have infinitely many nonterminal symbols. Fortunately, it turns out that for purposes of constructing the MAFP for a grammar G of that type, it is only necessary to index (or distinguish) the first occurrence of each repeating symbol. Thus, if G has the infinite set of rules abbreviated by the schema $A \rightarrow B\ C\ (B\ C)^n$, $n > 0$, then G' has the schema $A \rightarrow B^1\ C^1\ (B\ C)^n$, $n > 0$.

Note, finally that if the theory of CFPS grammar is extended to include grammars with infinitely many production rules (which, however, can be enumerated from finite schemata by a FS grammar), Theorem 1 is false, since clearly there is a FP for the grammar G consisting of the infinitely many productions abbreviated by the schema $S \rightarrow a^n$, $n > 0$, even though $L_B(G)$ is an infinite language. Indeed, Theorem 1 can be modified to state that for each CFPS grammar G that generates an infinite language solely by virtue of infinite sets of rules that can be enumerated by a finite automaton from finite schemata, there is a FT T that is directly strongly equivalent to G.

read to carry out that instruction, and I is a (possibly null) subinstruction to write and/or erase some string y^* (y a string of elements of U) on the PDS. The subinstruction $W(y^*)$ means to write y^* on the PDS and to push down whatever else appears on the PDS; the subinstruction $E(y^*)$ means to erase the string y^* on the PDS and to push up whatever else appears on the PDS. An erase subinstruction cannot be carried out if the auxiliary reading head is not scanning y^*. If both a write and an erase subinstruction appear in I, they are to be performed in the order indicated.

M accepts a string $\#\sigma\#$ on its input tape if and only if it can progress from its initial configuration to its final configuration exactly once while reading σ, using instructions constructed in accordance with the procedure now given.

I. If $B_n \rightarrow a$ is a production of G, then for all $B_1 \cdots B_n$ in K, the following is an instruction of M.

$$(B_1 \cdots B_n)_L \quad \rightarrow \qquad [_{B_n}a]_{B_n}(B_1 \cdots B_n)_R$$

II. If $B_n \rightarrow C_1 \cdots C_p$ is a production of G, then for all $B_1 \cdots B_n$ in K, the following are instructions of M.

A1. If $C_1 \neq B_k (1 \leqq k \leqq n)$:

$$(B_1 \cdots B_n)_L \quad \rightarrow \qquad [_{B_n} \quad (B_1 \cdots B_n C_1)_L$$

B1. If $C_i, C_{i+1} \neq B_k (1 \leqq i \leqq p-1; 1 \leqq k \leqq n)$:

$$(B_1 \cdots B_n C_i)_R \rightarrow \qquad (B_1 \cdots B_n C_{i+1})_L$$

C1. If $C_p \neq B_k (1 \leqq k \leqq n)$:

$$(B_1 \cdots B_n C_p)_R \rightarrow \qquad]_{B_n} \quad (B_1 \cdots B_n)_R$$

A2. If $C_1 = B_k (1 \leqq k \leqq n)$:

$$(B_1 \cdots B_n)_L \quad \rightarrow W(B_k \cdots B_n{}^*); \qquad [_{B_n} \quad (B_1 \cdots B_k)_L$$

B2a. If $C_i \neq B_j, C_{i+1} = B_k (1 \leqq i \leqq p-1; 1 \leqq j, k \leqq n)$:

$$(B_1 \cdots B_n C_i)_R \rightarrow W(B_k \cdots B_n{}^*); \qquad (B_1 \cdots B_k)_L$$

B2b. If $C_i = B_j, C_{i+1} \neq B_k (1 \leqq i \leqq p-1; 1 \leqq j, k \leqq n)$:

$$(B_1 \cdots B_j)_R \quad \rightarrow \quad E(B_j \cdots B_n{}^*) \qquad (B_1 \cdots B_n C_{i+1})_L$$

B2c. If $C_i = B_j, C_{i+1} = B_k (1 \leqq i \leqq p-1; 1 \leqq j, k \leqq n)$:

$$(B_1 \cdots B_j)_R \rightarrow E(B_j \cdots B_n{}^*), W(B_k \cdots B_n{}^*) \qquad (B_1 \cdots B_k)_L$$

C2. If $C_p = B_k (1 \leqq k \leqq n)$:

$$(B_1 \cdots B_k)_R \quad \rightarrow E(B_k \cdots B_n{}^*) \qquad]_{B_n} \quad (B_1 \cdots B_n)_R$$

III. The following are also instructions of M.

A. $S \qquad \rightarrow \qquad \# \qquad (S)_L$

B. $(S)_R \qquad \rightarrow \qquad \# \qquad F$

The states Σ of M are designed to indicate what symbol on the input tape M is scanning when it is in that state. Thus, if M is in fact in the course of a computation that accepts a sentence of $L_B(G)$ written on its input tape, and it is in the state $(B_1 \cdots B_n)_L$,

M is scanning the symbol $[_{B_n}$ on the input tape; similarly, if it is in the state $(B_1 \cdots B_n)_R$, M is scanning the symbol $]_{B_n}$. Moreover, the sequence of elements of U in the name of each state represents the labels on the unmatched left brackets that M has already read, in the order in which they were encountered, provided that no unmatched left bracket has been encountered more than once. When a given unmatched labeled left bracket is reencountered on the input tape, M enters the state that is appropriate for the first encounter of that bracket, in accordance with step IIA2, IIB2a, or IIB2c of the construction, and the labels of the unmatched left brackets that were encountered between the previous occurrence of the recurring bracket and the one being scanned are recorded on the PDS. When the labeled right bracket corresponding to the most recently scanned occurrence of the recursive labeled left bracket is encountered on the input tape, the PDS is erased back to the second occurrence of the boundary symbol * (if there is only one occurrence of * on the PDS, it becomes blank), and the erased symbols (except for *) are read back into the state name of the next configuration of M in accordance with step IIB2b, IIB2c, or IIC2 of the construction.[8] In this way M is able to keep a complete record of the unmatched left brackets it has previously read on the input tape, so as to be able to read labeled right brackets in the inverse order, as must be the case with sentences in $L_B(G)$.

The proof of the theorem that M accepts a string $\#\sigma\#$ just in case σ is a sentence of $L_B(G)$ would be too lengthy to give here. Instead, an illustration of the construction of a MAFP for a linguistically interesting grammar is given.

4. Illustration of the Construction

We illustrate the construction of a MAFP by considering the NF–CFPS grammar G, whose productions are given in (1). The symbols $\bar{C}$, $\bar{V}$, and P may be read "complement phrase", "verb phrase", and "possessive", respectively.

(1) a. $\mathrm{S} \rightarrow \bar{\mathrm{C}}\ \bar{\mathrm{V}}$
b. $\mathrm{S} \rightarrow \bar{\mathrm{N}}\ \bar{\mathrm{V}}$
c. $\bar{\mathrm{V}} \rightarrow \mathrm{V}\ \bar{\mathrm{C}}$
d. $\bar{\mathrm{V}} \rightarrow \mathrm{V}\ \bar{\mathrm{N}}$
e. $\bar{\mathrm{C}} \rightarrow \mathrm{C}\ \mathrm{S}$
f. $\bar{\mathrm{N}} \rightarrow \mathrm{D}\ \mathrm{N}$
g. $\mathrm{D} \rightarrow \bar{\mathrm{N}}\ \mathrm{P}$
h. $\mathrm{C} \rightarrow$ that
j. $\mathrm{D} \rightarrow$ the
k. $\mathrm{P} \rightarrow$'s
l. $\mathrm{N} \rightarrow$ {adult, boy, child, friend, girl, man, woman}
m. $\mathrm{V} \rightarrow$ {amazes, believes, bothers, knows}

[8] Step IIB2c of the construction is designed to deal with the situation in which M scans successively the right bracket of some recursive symbol followed by the left bracket of another recursive symbol. In this case, one can think of the string $B_j \ldots B_n$ as being read into the state name temporarily, so as to permit the string $B_k \ldots B_n$ to be transferred from the state name to the PDS.

G is designed to illustrate certain recursions that are typical of English: subject complementation, object complementation, and possessive modification. The MAFP M that parses sentences as generated by G has the instructions given in (2).[9]

(2) a.	i.	$(S)_L$	$\rightarrow$		$[_S$	$(S\bar{C})_L$	IIA1.
	ii.	$(S\bar{C})_R$	$\rightarrow$			$(S\bar{V})_L$	IIB1.
	iii.	$(S\bar{V})_R$	$\rightarrow$		$]_S$	$(S)_R$	IIC1.
b.	i.	$(S)_L$	$\rightarrow$		$[_S$	$(S\bar{N})_L$	IIA1.
	ii.	$(S\bar{N})_R$	$\rightarrow$			$(S\bar{V})_L$	IIB1.
c.	i.	$(S\bar{V})_L$	$\rightarrow$		$[_{\bar{V}}$	$(S\bar{V}V)_L$	IIA1.
	ii.	$(S\bar{V}V)_R$	$\rightarrow$			$(S\bar{V}\bar{C})_L$	IIB1.
	iii.	$(S\bar{V}\bar{C})_R$	$\rightarrow$		$]_{\bar{V}}$	$(S\bar{V})_R$	IIC1.
d.	i.	$(S\bar{V}V)_R$	$\rightarrow$			$(S\bar{V}\bar{N})_L$	IIB1.
	ii.	$(S\bar{V}\bar{N})_R$	$\rightarrow$		$]_{\bar{V}}$	$(S\bar{V})_R$	IIC1.
e.	i.	$(S\bar{C})_L$	$\rightarrow$		$[_{\bar{C}}$	$(S\bar{C}C)_L$	IIA1.
	ii.	$(S\bar{C}C)_R$	$\rightarrow$	$W(S\bar{C}^*)$		$(S)_L$	IIB2a.
	iii.	$(S)_R$	$\rightarrow$	$E(S\bar{C}^*)$	$]_{\bar{C}}$	$(S\bar{C})_R$	IIC2.
	iv.	$(S\bar{V}\bar{C})_L$	$\rightarrow$		$[_{\bar{C}}$	$(S\bar{V}\bar{C}C)_L$	IIA1.
	v.	$(S\bar{V}\bar{C}C)_R$	$\rightarrow$	$W(S\bar{V}\bar{C}^*)$		$(S)_L$	IIB2a.
	vi.	$(S)_R$	$\rightarrow$	$E(S\bar{V}\bar{C}^*)$;	$]_{\bar{C}}$	$(S\bar{V}\bar{C})_R$	IIC2.
f.	i.	$(S\bar{N})_L$	$\rightarrow$		$[_{\bar{N}}$	$(S\bar{N}D)_L$	IIA1.
	ii.	$(S\bar{V}\bar{N})_L$	$\rightarrow$		$[_{\bar{N}}$	$(S\bar{V}\bar{N}D)_L$	IIA1.
	iii.	$(S\bar{N}D)_R$	$\rightarrow$			$(S\bar{N}N)_L$	IIB1.
	iv.	$(S\bar{V}\bar{N}D)_R$	$\rightarrow$			$(S\bar{V}\bar{N}N)_L$	IIB1.
	v.	$(S\bar{N}N)_R$	$\rightarrow$		$]_{\bar{N}}$	$(S\bar{N})_R$	IIC1.
	vi.	$(S\bar{V}\bar{N}N)_R$	$\rightarrow$		$]_{\bar{N}}$	$(S\bar{V}\bar{N})_R$	IIC1.
g.	i.	$(S\bar{N}D)_L$	$\rightarrow$	$W(\bar{N}D^*)$;	$[_D$	$(S\bar{N})_L$	IIA2.
	ii.	$(S\bar{V}\bar{N}D)_L$	$\rightarrow$	$W(\bar{N}D^*)$;	$[_D$	$(S\bar{V}\bar{N})_L$	IIA2.
	iii.	$(S\bar{N})_R$	$\rightarrow$	$E(\bar{N}D^*)$		$(S\bar{N}DP)_L$	IIB2b.
	iv.	$(S\bar{V}\bar{N})_R$	$\rightarrow$	$E(\bar{N}D^*)$		$(S\bar{V}\bar{N}DP)_L$	IIB2b.
	v.	$(S\bar{N}DP)_R$	$\rightarrow$		$]_D$	$(S\bar{N}D)_R$	IIC1.
	vi.	$(S\bar{V}\bar{N}DP)_R$	$\rightarrow$		$]_D$	$(S\bar{V}\bar{N}D)_R$	IIC1.
h.	i.	$(S\bar{C}C)_L$	$\rightarrow$	$[_C\text{that}]_C$		$(S\bar{C}C)_R$	I.
	ii.	$(S\bar{V}\bar{C}C)_L$	$\rightarrow$	$[_C\text{that}]_C$		$(S\bar{V}\bar{C}C)_R$	I.
j.	i.	$(S\bar{N}D)_L$	$\rightarrow$	$[_D\text{the}]_D$		$(S\bar{N}D)_R$	I.
	ii.	$(S\bar{V}\bar{N}D)_L$	$\rightarrow$	$[_D\text{the}]_D$		$(S\bar{V}\bar{N}D)_R$	I.
k.	i.	$(S\bar{N}DP)_L$	$\rightarrow$	$[_P\text{'s}]_P$		$(S\bar{N}DP)_R$	I.
	ii.	$(S\bar{V}\bar{N}DP)_L$	$\rightarrow$	$[_P\text{'s}]_P$		$(S\bar{V}\bar{N}DP)_R$	I.

[9] The right-hand column in (2) gives the step of the construction that permits the establishment of each instruction of M. The numbering of the instructions reflects the numbering of the productions of the grammar G given in (1).

TABLE I. TRANSITIONS OF M WHILE ACCEPTING SENTENCE (*4a*)

Step	*Instruction*	*In State*	*Scanning*	*To State*	*Reading*	*PDS Content*
1.	n. i.	S	#	$(S)_L$	#	——
2.	a. i.	$(S)_L$	$[_S$	$(S\bar{C})_L$	$[_S$	——
3.	e. i.	$(S\bar{C})_L$	$[_{\bar{C}}$	$(S\bar{C}C)_L$	$[_{\bar{C}}$	——
4.	h. i.	$(S\bar{C}C)_L$	$[_C$	$(S\bar{C}C)_R$	$[_C\text{that}]_C$	——
5.	e. ii.	$(S\bar{C}C)_R$	$[_S$	$(S)_L$	——	$S\bar{C}*$
6.	b. i.	$(S)_L$	$[_S$	$(S\bar{N})_L$	$[_S$	$S\bar{C}*$
7.	f. i.	$(S\bar{N})_L$	$[_{\bar{N}}$	$(S\bar{N}D)_L$	$[_{\bar{N}}$	$S\bar{C}*$
8.	j. i.	$(S\bar{N}D)_L$	$[_D$	$(S\bar{N}D)_R$	$[_D\text{the}]_D$	$S\bar{C}*$
9.	f. iii.	$(S\bar{N}D)_R$	$[_N$	$(S\bar{N}N)_L$	——	$S\bar{C}*$
10.	l. i.	$(S\bar{N}N)_L$	$[_N$	$(S\bar{N}N)_R$	$[_N\text{boy}]_N$	$S\bar{C}*$
11.	f. v.	$(S\bar{N}N)_R$	$]_{\bar{N}}$	$(S\bar{N})_R$	$]_{\bar{N}}$	$S\bar{C}*$
12.	b. ii.	$(S\bar{N})_R$	$[_{\bar{V}}$	$(S\bar{V})_L$	——	$S\bar{C}*$
13.	c. i.	$(S\bar{V})_L$	$[_{\bar{V}}$	$(S\bar{V}V)_L$	$[_{\bar{V}}$	$S\bar{C}*$
14.	m.	$(S\bar{V}V)_L$	$[_V$	$(S\bar{V}V)_R$	$[_V\text{amazes}]_V$	$S\bar{C}*$
15.	d. i.	$(S\bar{V}V)_R$	$[_{\bar{N}}$	$(S\bar{V}\bar{N})_L$	——	$S\bar{C}*$
16.	f. ii.	$(S\bar{V}\bar{N})_L$	$[_{\bar{N}}$	$(S\bar{V}\bar{N}D)_L$	$[_{\bar{N}}$	$S\bar{C}*$
17.	j. ii.	$(S\bar{V}\bar{N}D)_L$	$[_D$	$(S\bar{V}\bar{N}D)_R$	$[_D\text{the}]_D$	$S\bar{C}*$
18.	f. iv.	$(S\bar{V}\bar{N}D)_R$	$[_N$	$(S\bar{V}\bar{N}N)_L$	——	$S\bar{C}*$
19.	l. ii.	$(S\bar{V}\bar{N}N)_L$	$[_N$	$(S\bar{V}\bar{N}N)_R$	$[_N\text{girl}]_N$	$S\bar{C}*$
20.	f. vi.	$(S\bar{V}\bar{N}N)_R$	$]_{\bar{N}}$	$(S\bar{V}\bar{N})_R$	$]_{\bar{N}}$	$S\bar{C}*$
21.	d. ii.	$(S\bar{V}\bar{N})_R$	$]_{\bar{V}}$	$(S\bar{V})_R$	$]_{\bar{V}}$	$S\bar{C}*$
22.	a. iii.	$(S\bar{V})_R$	$]_S$	$(S)_R$	$]_S$	$S\bar{C}*$
23.	e. iii.	$(S)_R$	$]_{\bar{C}}$	$(S\bar{C})_R$	$]_{\bar{C}}$	——
24.	a. ii.	$(S\bar{C})_R$	$[_{\bar{V}}$	$(S\bar{V})_L$	——	——
25.	c. i.	$(S\bar{V})_L$	$[_{\bar{V}}$	$(S\bar{V}V)_L$	$[_{\bar{V}}$	——
26.	m.	$(S\bar{V}V)_L$	$[_V$	$(S\bar{V}V)_R$	$[_V\text{bothers}]_V$	——
27.	d. i.	$(S\bar{V}V)_R$	$[_{\bar{N}}$	$(S\bar{V}\bar{N})_L$	——	——
28.	f. ii.	$(S\bar{V}\bar{N})_L$	$[_{\bar{N}}$	$(S\bar{V}\bar{N}D)_L$	$[_{\bar{N}}$	——
29.	j. ii.	$(S\bar{V}\bar{N}D)_L$	$[_D$	$(S\bar{V}\bar{N}D)_R$	$[_D\text{the}]_D$	——
30.	f. iv.	$(S\bar{V}\bar{N}D)_R$	$[_N$	$(S\bar{V}\bar{N}N)_L$	——	——
31.	l. ii.	$(S\bar{V}\bar{N}N)_L$	$[_N$	$(S\bar{V}\bar{N}N)_R$	$[_N\text{man}]_N$	——
32.	f. vi.	$(S\bar{V}\bar{N}N)_R$	$]_{\bar{N}}$	$(S\bar{V}\bar{N})_R$	$]_{\bar{N}}$	——
33.	d. ii.	$(S\bar{V}\bar{N})_R$	$]_{\bar{V}}$	$(S\bar{V})_R$	$]_{\bar{V}}$	——
34.	a. iii.	$(S\bar{V})_R$	$]_S$	$(S)_R$	$]_S$	——
35.	n. ii.	$(S)_R$	#	F	#	——

l.	i.	$(S\bar{N}N)_L$	$\rightarrow [_N\{\text{adult}, \ldots\}]_N$	$(S\bar{N}N)_R$	I.
	ii.	$(S\bar{V}\bar{N}N)_L$	$\rightarrow [_N\{\text{adult}, \ldots\}]_N$	$(S\bar{V}\bar{N}N)_R$	I.
m.		$(S\bar{V}V)_L$	$\rightarrow [_V\{\text{amazes}, \ldots\}]_V$	$(S\bar{V}V)_R$	I.
n.	i.	S	$\rightarrow \#$	$(S)_L$	IIIA.
	ii.	$(S)_R$	$\rightarrow \#$	F	IIIB.

M is capable of accepting any sentence of $L_B(G)$; in other words, of parsing any sentence of $L(G)$ in accordance with *G*. For example, *M* accepts the structures of (4), which are the P-markers of the sentences of (3) with respect to *G*.

(3) a. That the boy amazes the girl bothers the man.
b. The woman believes that the adult knows the child's friend.

(4) a. $\# \ [_S[_{\bar{C}}[_C\text{that}]_C\ [_S[_{\bar{N}}[_D\text{the}]_D\ [_N\text{boy}]_N]_{\bar{N}}\ [_{\bar{V}}[_V\text{amazes}]_V\ [_{\bar{N}}[_D\text{the}]_D\ [_N\text{girl}]_N]_{\bar{N}}]_{\bar{V}}]_S]_{\bar{C}}\ [_{\bar{V}}[_V\text{bothers}]_V\ [_{\bar{N}}[_D\text{the}]_D\ [_N\text{man}]_N]_{\bar{N}}]_{\bar{V}}]_S \ \#$
b. $\# \ [_S[_{\bar{N}}[_D\text{the}]_D\ [_N\text{woman}]_N]_{\bar{N}}\ [_{\bar{V}}[_V\text{believes}]_V\ [_{\bar{C}}[_C\text{that}]_C\ [_S[_{\bar{N}}[_D\text{the}]_D\ [_N\text{adult}]_N]_{\bar{N}}\ [_{\bar{V}}[_V\text{knows}]_V\ [_{\bar{N}}[_D[_{\bar{N}}[_D\text{the}]_D\ [_N\text{child}]_N]_{\bar{N}}\ [_P\text{'s}]_P]_D\ [_N\text{friend}]_N]_{\bar{N}}]_{\bar{V}}]_S]_{\bar{C}}]_{\bar{V}}]_S \ \#$

The sequence of transitions that *M* goes through in accepting (4a), the P-marker of (3a) with respect to *G*, is given in Table 1.[10]

5. Some Properties of MAFPs

Since MAFPs are minimally augmented, their use of the extra power of a PDS is limited to just those situations in which that power is necessary for effective parsing. Consequently, they do not need to use their PDS to parse all noun phrases within sentences, unlike augmented transition networks (see Woods 1969; 1970; Wanner and Maratsos 1971), which use the PDS for parsing major subconstituents regardless of recursion. The state names of MAFPs also directly reflect the grammatical relation (if any) of the constituent undergoing parsing (see Chomsky 1965, 71), so that the device *M* of section 4 requires no special routine to determine whether it is parsing, for example, a subject noun phrase or an object noun phrase. *M* is parsing a subject noun phrase with respect to *G* if and only if it is in states beginning $(S\bar{N}\ldots$, and a direct object noun phrase if and only if it is in states beginning $(S\bar{V}\bar{N}\ldots$.

The fact that access to a PDS is required for parsing not only CE structures, but also right-embedding (RE) and left-embedding (LE)[11] ones as well, follows from the fact that derivations of RE and LE structures with respect to any grammar that generates an infinite bracketing language are CE. In the case of RE structures, the labeled right brackets of the recursive symbol provide the necessary right-hand context for

[10] In fact, Table 1 gives a complete analysis of the internal configurations of *M* at each point in the process of accepting (4a).

[11] The terms "right embedding" and "left embedding", rather than the more customary designations "right branching" and "left branching", are used to emphasize the fact that one is dealing here with recursive structures, and also to highlight the parallelism with center embedding.

CE; for LE structures, the labeled left brackets provide the necessary left-hand context for CE. Thus, if $A \xrightarrow{*} \phi A$ is a subderivation with respect to a RE-CFPS G, where ϕ is nonnull, then $A \xrightarrow{*} \phi' A \psi$ is a subderivation with respect to G_B, where ϕ' is nonnull and where ψ consists of a nonnull string of right brackets; similarly, if $A \xrightarrow{*} A\psi$ is a subderivation with respect to G, where ψ is nonnull, then $A \xrightarrow{*} \phi A \psi'$ is a subderivation of G_B, where ψ' is nonnull and where ϕ consists of a nonnull string of left brackets. Since human beings do not have access to an unlimited PDS for parsing sentences, it follows that beyond some finite degree m of embedding, they must be unable to keep track of the number of recursions in LE and RE structures, and hence of the syntactic and semantic relations among their parts (the same, of course, is also true of CE structures).[12]

6. An Explanation for the Unacceptability of Embedding Constructions in Natural Languages

It has been known for some time that the human sentence recognition device is incapable of processing sentences of natural languages with greater than degree 3 or 4 of CE. From the foregoing discussion, it should also be the case that that device is unable to assign structural descriptions (say in the form of labeled bracketings for surface strings) to sentences with greater than some fixed finite amount of embedding, whether left, right, or center. Now, it may be observed that sentences of natural languages, like English, with degree of LE or RE greater than 3 or 4, are almost invariably produced with intonation breaks that do not correspond to the constituent structure assigned by the syntactic component of the grammar.[13] This readjustment of surface constituent structure effectively reduces multiple RE and LE structures to a kind of coordinate structure, in which the degree of embedding is reduced to degree 1. We take the fact that such readjustment is almost invariably performed in case the

[12] We can think of a perceptual or production model that incorporates a MAFP as imposing a fixed, finite limit m on the degree of embedding in the sentences of the bracketing language it accepts. Thus, sentences of $L_B(G)$ can be accepted by such a model if and only if those sentences have degree m or less of embedding. If the perceptual model also incorporates a minimally augmented FA for sentences of $L(G)$, in the manner suggested by Miller and Chomsky (1963), then it also imposes a (possibly different) fixed, finite limit n on the degree of CE in the sentences it accepts. Such a model predicts that, for a given CFPS grammar G, there may be sentences of $L(G)$ that are acceptable, whose corresponding structures in $L_B(G)$ are not acceptable; namely, all those sentences with less than degree n of CE, but with greater than degree m of embedding. In section 6 we point out that this prediction is borne out in natural languages, thus suggesting strongly that a model for human sentence recognition and production should incorporate devices for both recognition and production of sentences and their corresponding structural descriptions, perhaps in the form of a minimally augmented FT that pairs sentences with their structures with respect to the internalized grammar.

[13] As Chomsky (1965, 13–14) puts it:

> . . .there are no clear examples of unacceptability involving only left-branching or only right-branching, although these constructions are unnatural in other ways—thus, for example, in reading the right-branching construction [(5b)], the intonation breaks are ordinarily inserted in the wrong places (that is after 'cat' and 'rat', instead of where the main brackets appear). . .But it is unclear why left- and right-branching structures should become unnatural after a certain point, if they actually do.

The hedge at the end of this passage presumably reflects the fact that Chomsky was not aware at the time of any "unnatural" LE constructions. As we observe below, the unnaturalness in question does extend to LE constructions.

degree of embedding is greater than 3 or 4 to mean that the corresponding RE and LE structures are unacceptable, and the reason such structures are unacceptable to be that they cannot be recognized by the human sentence recognition device, since the latter incorporates a MAFP with a severe limit on the amount of PDS available to it for keeping track of embedding. In English, this readjustment of constituent structure has been specifically noted for RE structures of the type illustrated in (5).

(5) a. the book that was on the table that was near the door that was newly painted (Chomsky 1961, 127)
 b. This is the cat that caught the rat that stole the cheese. (Chomsky 1965, 13; Chomsky and Halle 1968, 372)

The structure assigned by the syntactic component of English grammar to the sentence (5b) is that given in (6a) (with irrelevant details omitted); however, such sentences are usually phrased as if they had the coordinate-like structure indicated in (6b).[14]

(6) a. $[_S$this is $[_{\bar{N}}$the cat $[_S$that caught $[_{\bar{N}}$the rat $[_S$that stole the cheese$]_S]_{\bar{N}}]_S]_{\bar{N}}]_S$
 b. $[_S[_S$this is $[_{\bar{N}}$the cat $]_{\bar{N}}]_S[_S$that caught $[_{\bar{N}}$the rat$]_{\bar{N}}]_S$ $[_S$that stole the cheese$]_S]_S$

Since the relation between structures like (6a) and (6b) is systematic, we may assume that there is a linguistic rule that relates them. Following Chomsky and Halle (1968, 371–372), we may call such a rule a readjustment rule (RR), about which class they have this to say (371):

> It seems clear that the grammar must contain readjustment rules that reduce surface structures, but it is very difficult to separate the study of these processes from the study of the theory of performance in any principled way.

7. The Status of Readjustment Rules in Grammar

One way that the study of the processes represented by RRs might be separated in a principled way from the study of the theory of performance is by an examination of the formal properties of RRs. If they should have formal properties in common with the formal properties of grammar, and not have properties in common with the properties of the theory of performance, then we would have clear evidence of their grammatical character. Moreover, we should be able to decide whether those rules belong in the syntactic component of the grammar, or in a separate component of their own.[15]

[14] That is, the embedded clauses become sisters of the matrix clause. This occurs whenever the embedded clauses are the rightmost constituents of the matrix. When the embedded clauses are internal to the matrix clause, then the embedded clauses become daughters of the matrix clause. Since the nature of the adjunction can be determined from the configuration of the input, we shall say nothing further about it here.

[15] Obviously, the other components of the grammar—base, semantic, and phonological—may be ruled out a priori.

Let us therefore state formally the RR that relates (6a) and (6b); such a statement is given in (7).[16]

(7) *Embedded-Sentence Readjustment* (*ESR*)

$$\begin{array}{llll} X & - \ S & - \ Y & \\ 1 & 2 & 3 & \overset{\text{opt}}{\Longrightarrow} \\ 1 & \phi & 2+3 & \end{array}$$

Conditions: a. The A-over-A condition is inapplicable.

b. 2 is on a right branch; i.e. $1 = X'A_1 \cdots A_n$, where $A_1 \cdots A_n S$ is a B.

From an examination of the properties of ESR, we see immediately that if such a rule is in the grammar, it is not in the syntactic component, since it is not a phrase-structure rule, and since it violates at least two well-motivated principles governing the structure of syntactic transformations. First, since it must be applicable to occurrences of the category S embedded within S, the A-over-A condition (Chomsky 1964; 1973), which governs the mode of application of every syntactic transformation, must be suspended for ESR (see condition (a)).[17] Second, the structural description of ESR requires making explicit references to the internal constituent structure of its factors, contrary to otherwise well-motivated conditions on proper factorization (Chomsky 1961; Peters and Ritchie 1973). Even if such reference were to be allowed for syntactic transformations, however, condition (b) would still be in violation of the principle of minimal factors, according to which a nonvariable factor must either be directly affected by the transformation (deleted or adjoined to something else) or provide the immediate context for such an operation.[18] Thus RRs, or at least those like ESR, cannot be part of the syntactic component of a grammar.

From the formal properties of ESR, however, there are several excellent reasons for considering the rule to be part of grammar, rather than part of performance. First, like syntactic rules, ESR makes no reference to the particular relations that the constituents it applies to bear to one another: it does not matter whether the embedded sentences to which it applies are complement clauses or relative clauses; whether those clauses have complementizers or relative pronouns or not; or whether all of the clauses to which it applies in a given sentence are all of the same type, are similar in internal surface structure, or are dissimilar. If the rule were one of performance, one would expect that these particular properties of embedded clauses would play a

[16] The adjunction sign "+" is to be interpreted in the manner described in note 14.

[17] Note that this suspension of the A-over-A condition cannot be gotten around by having the rule apply either cyclically or anticyclically, since in either case wrong derived constituent structures would be obtained. If the rule can apply more than once to a given P-marker, it applies simultaneously.

[18] To my knowledge, the principle of minimal factors has yet to be discussed in detail in the literature, but it has been presented and motivated by Chomsky and others in public lectures.

role in the statement of the rule. Also, like syntactic transformations, ESR is capable of introducing structural ambiguity, hardly a property one would expect of rules of performance. To see this, consider again (6b). Not only can that structure be derived from (6a) by ESR, it can also be derived from (8) by ESR, in which the relative clauses are "stacked" modifiers of the head noun *cat*.[19]

(8) $[_{S}$this is $[_{\bar{N}}[_{\bar{N}}$the cat $[_{S}$that caught $[_{\bar{N}}$the rat$]_{\bar{N}}]_{S}]_{\bar{N}}$ $[_{S}$that stole the cheese$]_{S}]_{\bar{N}}]_{S}$

Second, the fact that the results of applying ESR to structures in which complementizers or relative pronouns are omitted (so that the clauses are no longer morphologically marked as being subordinate), or to structures in which the embedded clauses are of very different internal form are, to varying degrees, unnatural, is to be explained on the basis of the theory of performance as applied to those results, and that fact is not to be taken as limiting the applicability of the rule. On the other hand, the fact that ESR and other RRs like it appear to be "motivated" and perhaps even "explained" on the basis of the fact that human users of natural languages have very limited capacities for parsing multiply embedded structures, is no different in principle from the fact that many stylistic syntactic transformations also appear to be motivated on the basis of their ability to reduce CE structures to LE and RE ones (Yngve 1960; Chomsky 1961, 126; Miller and Chomsky 1963, 471). In each case, we are clearly dealing with rules of grammar whose general properties are ultimately constrained by the systems of language use (on this point, see also Bever and Langendoen 1971).[20]

Third, a striking reason for considering ESR to be a rule of grammar is the fact that the rule must be modified to express the idiosyncratic fact that certain phrases

[19] Despite the fact that the structure of (8) appears to be LE on $\bar{N}$, the recursive category is really S, which appears always on right branches in the stacked-relative-clause construction.

[20] In other words, rules of grammar are never specific responses to usage needs. Although it may be obvious that application of a specific rule, such as Extraposition or ESR, has desirable properties from the point of view of language users in a wide variety of cases, it also happens that these rules can have effects that are quite undesirable (with the result that certain sentences in which those rules apply are felt to be unacceptable). In the case of Extraposition, although it is true that its application in the derivation of a sentence like (i) results in a more comprehensible sentence than if it had not applied (as in (ii)), its application in (iii) results in a less comprehensible sentence than if it had not applied (as in (iv)).

(i) It's a pity that they don't want children.
(ii) That they don't want children is a pity.
(iii) It suggests that they consider parenthood a drag that they don't want children.
(iv) That they don't want children suggests that they consider parenthood a drag.

Concerning ESR, it may also be observed that when it is applied to embedded clauses, it converts them into coordinate-like adjuncts of the main clause; when applied to coordinate structures, it destroys the coordinate relationship! Thus, consider a sentence like (v).

(v) I believe that the cat chased the rat and that the rat stole the cheese.

This has the syntactic structure as indicated in (via); since the second complement clause is on a right branch, and since the A-over-A condition is suspended, ESR can extract the second conjunct and adjoin it as a sister to the main clause, resulting in (vib), assuming pied piping of *and*.

(vi) a. $[_{S}$I believe $[_{S}[_{S}$that the cat chased the rat$]_{S}$ and $[_{S}$that the rat stole the cheese$]_{S}]_{S}]_{S}$
b. $[_{S}[_{S}$I believe $[_{S}$that the cat chased the rat$]_{S}]_{S}$ and $[_{S}$that the rat stole the cheese$]_{S}$

containing the verb *be* may be "pied piped" (Ross 1967) along with an immediately following embedded clause. Thus, consider the sentences in (9).

(9) a. I believe that his objection is that the election procedures are too complicated.
b. I believe that his objection cannot be that the election procedures are too complicated.

In both of these examples, an intonation break may appear after the noun *objection*, indicating that *is* in (9a), and *cannot be* in (9b), may be pied piped together with the embedded clause *that the election procedures are too complicated*.[21] That the pied piping of the verb along with the clause is limited to phrases with the head *be* can be seen from an example like (10), in which case an intonation break cannot appear between the noun *objection* and the following verb.

(10) I believe that his objection remains that the election procedures are too complicated.

While it is not out of the question, it is hard to imagine why a performance theory should contain the possibility of moving elements under the conventions governing pied piping (though such a theory independently must contain rules for interpreting structures in which pied piping has taken place). For these two groups of reasons, then, we take it to have been established that ESR is a rule of English grammar, and that it belongs in a distinct readjustment-rule component of the grammar. We shall discuss its position in the grammar further below.

Before taking up the question of the existence of other RRs besides ESR in the RR-component of the grammar of English, let us examine more closely the applicability of ESR to a variety of structures other than the strictly RE type exemplified in (5). We have already considered one such variety, namely that illustrated in (8). Now consider (11), which has one degree of CE, although all of the embedded clauses are introduced on right branches. Its syntactically motivated constituent structure is that given in (12a); the result of applying ESR to that structure is given in (12b).

(11) The cat that caught the rat that stole the cheese was sick.

(12) a. $[_S[_{\bar{N}}$the cat $[_S$that caught $[_{\bar{N}}$the rat $[_S$that stole the cheese$]_S]_{\bar{N}}]_S]_{\bar{N}}$ $[_{\bar{V}}$was sick$]_{\bar{V}}]_S$
b. $[_S[_{\bar{N}}$the cat$]_{\bar{N}}[_S$that caught $[_{\bar{N}}$the rat$]_{\bar{N}}]_S[_S$that stole the cheese$]_S[_{\bar{V}}$was sick$]_{\bar{V}}]_S$

[21] When the *be*-phrase is pied piped, the resulting derived structure looks a bit strange, since a verb phrase is adjoined as a sister to a sentence. The node that is introduced by the adjunction operation is, of course, S, and not $\bar{V}$. If *is* is not pied piped in (9a), it still cannot be contracted with *objection*, exactly as predicted by ESR and the well-known restriction that the copula in English cannot be contracted with a preceding element in case the constituent that follows it is removed (King 1970; Baker 1971; Zwicky 1971). This observation provides independent evidence for the effect of ESR.

The intonation breaks predicted by (12b) seem entirely natural, so that we may conclude that ESR has properly applied in this case.

Now consider the application of ESR to the doubly CE sentence (13), in which, nevertheless, all of the embedded clauses are introduced on right branches. Its syntactically motivated constituent structure is (14a); the result of applying ESR to (14a) is (14b).

(13) The cat that the rat that stole the cheese was afraid of was sick.

(14) a. $[_S[_{\bar{N}}$the cat $[_S$that $[_{\bar{N}}$the rat $[_S$that stole the cheese$]_S]_{\bar{N}}$ $[_{\bar{V}}$was afraid of$]_{\bar{V}}]_S]_{\bar{N}}$ $[_{\bar{V}}$was sick$]_{\bar{V}}]_S$

b. $[_S[_{\bar{N}}$the cat$]_{\bar{N}}$ $[_S$that $[_{\bar{N}}$the rat$]_{\bar{N}}$ $[_S$that stole the cheese$]_S$ $[_{\bar{V}}$was afraid of$]_{\bar{V}}]_S$ $[_{\bar{V}}$was sick$]_{\bar{V}}]_S$

All that ESR did in this case was to flatten the structure somewhat; obviously, since the rule cannot affect the linearization of the string of formatives in any way, it cannot reduce the degree of CE. But, the intonation breaks that the application of the rule predicts (particularly those following *cat* and *rat*) accord perfectly with the phrasing of someone who has mastered the uttering of doubly CE structures. In other words, ESR predicts the intonation pattern of an ideal speaker, exactly as we would expect from the hypothesis that ESR is a rule of grammar.

From the discussion in section 6, we might expect that for every recursive category that appears on a right or left branch, there is a RR that raises that category to be coordinate with or immediately subordinate to the highest available occurrence of that category. This expectation is borne out. English does not happen to embed sentences on left branches; however, it does embed noun phrases on both right and left branches, and there is a RR for each of these types of embeddings. Consider first the multiple occurrence of noun phrases on right branches, as illustrated in (15). The syntactically motivated constituent structure for (15) is given in (16a); its readjusted structure, as indicated by its possible rephrasing, is given in (16b).[22]

(15) This is the friend of the daughter of the ambassador to West Germany.

(16) a. $[_S$this is $[_{\bar{N}}$the friend $[_{\bar{P}}$of $[_{\bar{N}}$the daughter $[_{\bar{P}}$of $[_{\bar{N}}$the ambassador $[_{\bar{P}}$to West Germany$]_{\bar{P}}]_{\bar{N}}]_{\bar{P}}]_{\bar{N}}]_{\bar{P}}]_S$

b. $[_S$this is $[_{\bar{N}}[_{\bar{N}}$the friend$]_{\bar{N}}$ $[_{\bar{P}}$of $[_{\bar{N}}$the daughter$]_{\bar{N}}]_{\bar{P}}$ $[_{\bar{P}}$of $[_{\bar{N}}$the ambassador$]_{\bar{N}}]_{\bar{P}}$ $[_{\bar{P}}$to $[_{\bar{N}}$West Germany$]_{\bar{N}}]_{\bar{P}}]_{\bar{N}}]_S$

The RR that relates (16a) and (16b), together with conditions that will be motivated below, is given in (17).[23]

[22] $\bar{P}$ may be read "preposition phrase". Since P has already been used as a category symbol for the possessive element *'s*, we use the symbol Pr for the category "preposition".

[23] From condition (a) it follows that A-over-A is not operative; hence violation of A-over-A is not mentioned specifically in the set of conditions in (17).

(17) *Embedded-Noun-Phrase Readjustment* ($E\bar{\bar{N}}R$)

$$\begin{array}{ccccc} W & - & X & - & \bar{\bar{N}} & - & Y & - & Z \end{array}$$

$$\begin{array}{cccccc} 1 & 2 & 3 & 4 & 5 & \xrightarrow{\text{opt}} \\ 1 & 2 & \phi & 3+4 & 5 & \end{array}$$

Conditions: a. $2^\frown 3^\frown 4$ is a $\bar{\bar{N}}$.

b. 3 is on a right branch.

c. If $2 = X'\ [_{\bar{P}}\text{Pr}$, then $[_{\bar{P}}\text{Pr}$ is pied piped, and if $4 =]_{\bar{P}}Y_1$ then $]_{\bar{P}}$ is pied piped.

d. There do not exist X_1, X_2, Y_1, Y_2 such that $2 = X_1[_S X_2$ and $4 = Y_1]_S Y_2$.

Condition (a) on $E\bar{\bar{N}}R$ is required to ensure that only embedded occurrences of $\bar{\bar{N}}$ are readjusted. Obviously, in a sentence like (18), the noun phrase *the movie* is not adjoined as a sister to the rest of the sentence, *I saw*.

(18) I saw the movie.

The need for conditions (b) and (c) is apparent. What condition (d) ensures is that no noun phrase will be readjusted out of a sentence that is itself part of a noun phrase. Without such a condition, the phrase *of the ambassador* in (19) would, incorrectly, be adjoined as a sister of the full noun phrase *the man who met the friend*, but with the condition, it can be adjoined to the smaller noun phrase *the friend*, in accordance with the facts.

(19) I know the man who met the friend of the ambassador.

Finally, consider the case of noun phrases embedded on left branches, as illustrated in (20). The syntactically motivated constituent structure of (20) is given in (21a). That structure may, however, be readjusted to yield (21b).

(20) My friend's oldest nephew's favorite teacher's strangest idea is that linguistics is profitable.

(21) a. $[_S[_{\bar{\bar{N}}}[_D[_{\bar{\bar{N}}}[_D[_{\bar{\bar{N}}}[_D[_{\bar{\bar{N}}}$my friend$]_{\bar{\bar{N}}}$'s$]_D$ oldest nephew$]_{\bar{\bar{N}}}$'s$]_D$ favorite teacher$]_{\bar{\bar{N}}}$'s$]_D$ strangest idea$]_{\bar{\bar{N}}}$ $[_{\bar{\bar{V}}}$is that linguistics is profitable$]_{\bar{\bar{V}}}]_S$

b. $[_S[_{\bar{\bar{N}}}[_D[_{\bar{\bar{N}}}$my friend$]_{\bar{\bar{N}}}$'s$]_D$ $[_D[_{\bar{\bar{N}}}$oldest nephew$]_{\bar{\bar{N}}}$'s$]_D$ $[_D[_{\bar{\bar{N}}}$favorite teacher$]_{\bar{\bar{N}}}$'s$]_D$ $[_{\bar{\bar{N}}}$strangest idea$]_{\bar{\bar{N}}}]_{\bar{\bar{N}}}$ $[_{\bar{\bar{V}}}$is that linguistics is profitable$]_{\bar{\bar{V}}}]_S$

The RR that relates (21a) and (21b) is given in (22).

(22) *Left-Embedded-Noun-Phrase Readjustment* ($LE\bar{\bar{N}}R$)

$$\begin{array}{ccccc} W & - & X & - & \bar{\bar{N}} & - & Y & - & Z \end{array}$$

$$\begin{array}{cccccc} 1 & 2 & 3 & 4 & 5 & \xrightarrow{\text{opt}} \\ 1 & 2+3 & \phi & 4 & 5 & \end{array}$$

Conditions: a. $2^\frown 3^\frown 4$ is a $\bar{\bar{N}}$.
b. 3 is on a left branch.
c. If $2 = X'\ [_D$ and if $4 = \text{'}s]_D\ Y'$, then $[_D$ and $\text{'}s]_D$ are pied piped.
d. There are no X_1, X_2, Y_1, Y_2 such that $2 = X_1\ [_S X_2$ and $4 = Y_1]_S\ Y_2$.

The conditions on LE$\bar{\bar{N}}$R are motivated on the same grounds as the conditions on E$\bar{\bar{N}}$R.[24]

It is possible to construct an elaborate noun phrase containing both LE and RE noun phrases, resulting in CE of noun phrases within noun phrases. A case illustrating degree 2 of CE of noun phrase is illustrated in (23). The syntactically motivated constituent structure for (23) is given in (24a), and the result of applying E$\bar{\bar{N}}$R and LE$\bar{\bar{N}}$R to that structure is given in (24b), a derived structure that provides the basis for an ideal speaker's intonation pattern for (23).

(23) I borrowed the friend of the ambassador to West Germany's neighbor's sailboat.

(24) a. $[_S$I borrowed $[_{\bar{\bar{N}}}[_D[_{\bar{\bar{N}}}$the friend $[_{\bar{P}}$of $[_{\bar{\bar{N}}}[_D[_{\bar{\bar{N}}}$the ambassador $[_{\bar{P}}$to $[_{\bar{\bar{N}}}$West Germany$]_{\bar{\bar{N}}}]_{\bar{P}}]_{\bar{\bar{N}}}$'s$]_D$ neighbor$]_{\bar{\bar{N}}}]_{\bar{P}}]_{\bar{\bar{N}}}$'s$]_D$ sailboat$]_{\bar{\bar{N}}}]_S$

b. $[_S$I borrowed $[_{\bar{\bar{N}}}[_D[_{\bar{\bar{N}}}[_{\bar{\bar{N}}}$the friend$]_{\bar{\bar{N}}}$ $[_{\bar{P}}$of $[_{\bar{\bar{N}}}[_D[_{\bar{\bar{N}}}[_{\bar{\bar{N}}}$the ambassador$]_{\bar{\bar{N}}}$ $[_{\bar{P}}$to $[_{\bar{\bar{N}}}$West Germany$]_{\bar{\bar{N}}}]_{\bar{P}}]_{\bar{\bar{N}}}$'s$]_D$ $[_{\bar{\bar{N}}}$neighbor$]_{\bar{\bar{N}}}]_{\bar{\bar{N}}}]_{\bar{P}}]_{\bar{\bar{N}}}$'s$]_D$ $[_{\bar{\bar{N}}}$sailboat$]_{\bar{\bar{N}}}]_{\bar{\bar{N}}}]_S$

If we now compare the statements of the three RRs that we have proposed for English: ESR in (7), E$\bar{\bar{N}}$R in (17), and LE$\bar{\bar{N}}$R in (22), we see that they have much in common, suggesting that they can be collapsed into a single RR schema. To reconcile the differences among them, we note first that the various pied-piping conditions could be specified as separate conditions on the RR schema; second, that an analogue to condition (a) on E$\bar{\bar{N}}$R and LE$\bar{\bar{N}}$R can be imposed on ESR without altering the effect of that rule, so that the structural description of ESR can be written with the same five factors as E$\bar{\bar{N}}$R and LE$\bar{\bar{N}}$R; third, that whether adjunction is to the left or to the right depends only on whether the recursive category is on a left branch or on a right branch; and fourth, that condition (d) on E$\bar{\bar{N}}$R and LE$\bar{\bar{N}}$R, while not appropriate to ESR, can be specified in the schema so as not to affect the operation of ESR. Without specifying again the various pied-piping conditions, the RR schema for English may be stated as in (25).[25]

[24] Since pied piping of the determiner and possessive nodes is obligatory (one might be able to argue that the pied piping of the preposition and preposition phrase in E$\bar{\bar{N}}$R is optional), one might ask why it is not done directly in the structural change of the rule itself. The reason is that ultimately we shall be replacing these specific rules by a general rule schema (see (25)), in which the pied-piping conditions are listed separately.

[25] The RR schema (25) does not abbreviate all of the RRs in English, but only those involving the reduction of degree of embedding. The formation of words from syntactically separate items, such as a rule that attaches *'s* to the noun it immediately follows no matter what its scope is in syntactic surface structure, or even in the output of the RR schema, are also RRs that belong in the RR component.

(25) *Readjustment-Rule Schema for English*

$$\begin{array}{ccccc} W & - X & - A & - Y & - Z \\ 1 & 2 & 3 & 4 & 5 \end{array} \overset{\text{opt}}{\Longrightarrow}$$

$$1\begin{Bmatrix} 2+3 \\ \\ 2 \end{Bmatrix}\phi\begin{Bmatrix} 4 \\ \\ 3+4 \end{Bmatrix}5 \qquad \begin{matrix}\text{(i)} \\ \\ \text{(ii)}\end{matrix}$$

Conditions: a. A is a recursive category.[26]

b. $2^\frown 3^\frown 4$ is an A.

c. (i) results if 3 is on a left branch;
(ii) results if 3 is on a right branch.

d. If $3 \neq S$, then there are no X_1, X_2, Y_1, Y_2 such that $2 = X_1\,[_S X_2$ and $4 = Y_1]_S\, Y_2$.

e. Various pied-piping conditions, such as condition (c) of (17) and (22).

Note that the only English-specific aspect of the schema (25) is contained in condition (e). The structural description and change, as well as conditions (a–d), are formulated so as to suitably represent the nature of constituent readjustment in any language, and hence constitute a hypothesis about universal grammar. The notion that a schema like (25) should be universal follows from four assumptions: (i) that recursive embedding is a property of languages generated by transformational grammars; (ii) that all acceptable strings (strings whose surface P-markers have less than some small degree of CE and that are otherwise acceptable) receive phonological interpretation in performance; (iii) that the input to the phonological component consists of well-formed P-markers; and (iv) that strings whose surface P-markers have some small degree of embedding cannot be parsed, and hence cannot receive phonological interpretation in performance. Assumptions (i)–(iii) are uncontroversial, and assumption (iv) follows from Theorem 1. Hence a readjustment-rule schema like (25) must be universal.

Moreover, the structural description and change, in addition to conditions (a–c) of (25), reasonably follow from these four assumptions, in that they minimally accomplish the desired effect of reducing the degree of embedding of acceptable strings to 1 (that of coordinate structures), and hence make them capable of being parsed. Condition (d) has the further desired effect of preventing nonsentential constituents from being removed from the sentences containing them, hence limiting the disruptive effect on intelligibility that application of RRs inevitably has. Thus the schema (25), less condition (e), is a very reasonable hypothesis about universal grammar.

[26] It may be correct to identify the categories that satisfy condition (a) as the cyclic categories S and $\overline{\text{N}}$. Categories that appear to be recursive in surface structure, but that are not in deep structure, such as $\overline{\text{V}}$, do not appear to undergo constituent readjustment.

As we have already argued, the formal properties of RRs are quite distinct from those of syntactic transformations. Thus, we are in a position to affirm not only that schema (25) without condition (e) is universal, but also that it belongs in a separate component of the grammar, one that relates syntax and phonology, thus confirming the conjecture by Chomsky (1973, 254) that rules that "never change the terminal string of phrase marker but only its structure . . . can be restricted to the readjustment rule component of the grammar, which relates syntax and phonology."[27,28]

References

Anderson, S. and P. Kiparsky, eds. (1973) *A Festschrift for Morris Halle*, Holt, Rinehart, and Winston, New York.

Baker, C. L. (1971) "Stress Level and Auxiliary Reduction in English," *Linguistic Inquiry* II, 167–182.

Bar-Hillel, Y., M. Perles, and E. Shamir (1961) "On Formal Properties of Simple Phrase Structure Grammars," reprinted in R. D. Luce, R. R. Bush, and E. Galanter, eds., (1965).

Bever, T. G. and D. T. Langendoen (1971) "A Dynamic Model of the Evolution of Language," *Linguistic Inquiry* II, 433–463.

Braffort, P. and D. Hirschberg, eds. (1963) *Computer Programming and Formal Systems*, North-Holland, Amsterdam.

Bresnan, J. (1971) "Sentence Stress and Syntactic Transformations," *Language* 47, 257–281.

Bresnan, J. (1972) "Stress and Syntax," *Language* 48, 326–342.

Chomsky, N. (1959a) "On Certain Formal Properties of Grammars," reprinted in R. D. Luce, R. R. Bush, and E. Galanter, eds. (1965), 125–155.

Chomsky, N. (1959b) "A Note on Phrase Structure Grammars," *Information and Control* 2, 393–395.

Chomsky, N. (1961) "On the Notion 'Rule of Grammar'," reprinted in J. Fodor and J. Katz, eds. (1964), 119–136.

Chomsky, N. (1963) "Formal Properties of Grammars," in R. D. Luce, R. R. Bush, and E. Galanter, eds. (1963), 323–418.

Chomsky, N. (1964) *Current Issues in Linguistic Theory*, Mouton, The Hague.

Chomsky, N. (1965) *Aspects of the Theory of Syntax*, MIT Press, Cambridge, Massachusetts.

Chomsky, N. (1970) "Remarks on Nominalizations," in R. A. Jacobs and P. S. Rosenbaum, eds. (1970), 184–221.

Chomsky, N. (1973) "Conditions on Transformations," in S. Anderson and P. Kiparsky, eds. (1973), 232–286.

[27] Downing (1973, 119–120) argues that a rule he calls Complement Detachment is a readjustment rule that must precede at least one movement transformation, namely a rule he calls Matrix Embedding. He concludes that "there is [thus] no reason to suppose that such purely restructuring rules belong in a special 'readjustment' component between syntax and phonology." However, his argument for the existence of the rule of Matrix Embedding depends on his analysis of the derivation of parenthetical expressions, an analysis that is easily challenged. In particular, he gives no arguments against the analysis in which parenthetical expressions are base generated as such, and for which no rule of Matrix Embedding is required.

[28] The application of readjustment rules effectively destroys the stress contrasts that motivate Bresnan's (1971, 1972) analysis of stress assignment in the syntactic cycle. Apparently, stress must be reassigned from scratch (or nearly so) in the phonological cycle if readjustment rules have operated. More study of the interaction between pre- and post-readjustment stress assignment is obviously called for.

Chomsky, N. and M. Halle (1968) *The Sound Pattern of English*, Harper and Row, New York.

Chomsky, N. and M. P. Schützenberger (1963) "The Algebraic Theory of Context-Free Languages," P. Braffort and D. Hirschberg, eds. (1963), 118–161.

Downing, B. T. (1973) "Parenthesization Rules and Obligatory Phrasing," *Papers in Linguistics* 6, 108–128.

Fodor, J. and J. Katz, eds. (1964) *The Structure of Language: Readings in the Philosophy of Language*, Prentice-Hall, Englewood Cliffs, New Jersey.

Hopcroft, J. E. and J. D. Ullman (1969) *Formal Languages and their Relation to Automata*, Addison-Wesley, Reading, Massachusetts.

Jacobs, R. A. and P. S. Rosenbaum, eds. (1970) *Readings in English Transformational Grammar*, Ginn-Blaisdell, Waltham, Massachusetts.

King, H. V. (1970) "On Blocking the Rules for Contraction in English," *Linguistic Inquiry* 1, 134–136.

Luce, R. D., R. R. Bush, and E. Galanter, eds. (1963) *Handbook of Mathematical Psychology*, Volume II, John Wiley and Sons, New York.

Luce, R. D., R. R. Bush, and E. Galanter, eds. (1965) *Readings in Mathematical Psychology*, Volume II, John Wiley and Sons, New York.

Miller, G. A. and N. Chomsky (1963) "Finitary Models of Language Users," in R. D. Luce, R. R. Bush, and E. Galanter, eds. (1963), 419–492.

Peters, P. S. and R. W. Ritchie (1973) "On the Generative Power of Transformational Grammars," *Information Sciences* 6, 49–83.

Rabin, M. O. and D. Scott (1959) "Finite Automata and their Decision Problems," *IBM Journal of Research and Development* 3, 115–125.

Ross, J. R. (1967) *Constraints on Variables in Syntax*, unpublished Doctoral dissertation, MIT, Cambridge, Massachusetts.

Wanner, E. and M. Maratsos (1971) "On Understanding Relative Clauses," unpublished paper.

Woods, W. A. (1969) *Augmented Transition Networks for Natural Language Analysis*, Report CS–1, Computational Laboratory of Harvard University, Cambridge, Massachusetts.

Woods, W. A. (1970) "Transition Network Grammars for Natural Language Analysis," *Communications of the ACM* 13, 591–602.

Yngve, V. H. (1960) "A Model and an Hypothesis for Language Structure," *Proceedings of the American Philosophical Society* 104, 444–466.

Zwicky, A. M. (1971) "Auxiliary Reduction in English," *Linguistic Inquiry* II, 323–336.

Ph.D. Program in Linguistics
C.U.N.Y. Graduate Center
33 West 42 St.
New York, New York 10036