ON THE INADEQUACY OF TYPE-3 AND TYPE-2 GRAMMARS FOR HUMAN LANGUAGES*

D. TERENCE LANGENDOEN

City University of New York Graduate Center

Chomsky (1956; 1957, chapter 3) and Bar-Hillel and Shamir (1961) present arguments that no type-3 (equivalently, one-sided linear, regular, or finite-state) grammar can generate all and only all of the sentences of human languages, such as English. Bar-Hillel and Shamir (1961) and Postal (1964) argue further that no type-2 (context-free phrase-structure) grammar can generate all and only all of the sentences of English and Mohawk, respectively. According to these arguments, the theory of type-3 grammar and the theory of type-2 grammar lack the weak generative capacity necessary for an adequate theory of human language.

Critics of these classical arguments, for example Daly (1974) and Levelt (1974), focus on two points. First they claim that the argument forms are of questionable validity. Second they challenge the major premiss on which these arguments are based, namely that at least some human languages contain infinitely many grammatical sentences that are nevertheless entirely unacceptable to those that know those languages. We take up each aspect of these criticisms in turn.

The invalidity (if that is what it is) of the classical arguments is easily corrected, as Levelt himself notes, by making use of the theorem that

*An earlier version of this paper appeared under a slightly different title, in *CUNY Forum* 1.1-12, 1976. I thank C. Kaniklidis and H. Savin for helpful suggestions.

if the intersection of a language H with a type-3 language R is a language of type-n, then H cannot be of type-m, where m is greater than n (Arbib 1969). If one takes examples like those used in the classical arguments, one finds that the results of intersecting H with R turn out to be languages of the following families.[1]

(1) Languages of the form $L = \{xu^n wv^n y : n \geqq 0 \ \& \ u \neq \phi \ \& \ v \neq \phi\}$, called n-dependency languages.

(2) Languages of the form $L = \{x\hat{x} : x \in F$, where F is a type-3 language & $\hat{x}$ is x backwards$\}$, called mirror-image languages.

(3) Languages of the form $L = \{xx : x \in F$, where F is a type-3 language$\}$, called copying languages.

The linguistic interest of n-dependency and mirror-image languages is that they are generally of type-2, while copying languages are of interest because they are generally of type-1 (context-sensitive phrase-structure). However, not all members of these various families are of the requisite type. Consider the following examples.

(4) $L_a = \{ab^{2m}a : m \geqq 0\}$.

Clearly, L_a is a type-3 language. Nevertheless, it is an n-dependency language, since we can set x = a, u = b, w = ϕ, v = b, y = a, and n = m/2. Hence, not every n-dependency language is a type-2 language.

(5) $L_b = \{(abba)^{2n+1} : n \geqq 0\}$.

Clearly also, L_b is a type-3 language. Nevertheless, it is a mirror-image language, since it is of the form $\{x\hat{x} : x \in F_b\}$, where $F_b = \{(abba)^n ab : n \geqq 0\}$. Hence, not every mirror-image language is a type-2 language.

(6) $L_c = \{a^n ba^n b : n \geqq 0\}$.

L_c is a type-2 language; in fact, an n-dependency type-2 language. Nevertheless, it is a copying language, where x = $a^n b$. Hence, not every copying language is a type-1 language.[2]

Therefore, in order for the classical arguments that human languages are not of type-3 or of type-2 to go through, the languages that result upon intersection of the given human languages with the requisite type-3 languages must not only be n-dependency, mirror-image, or copying languages but must also be languages of these families of the appropriate type. It would be useful, therefore, to know the necessary and sufficient conditions under which n-dependency and mirror-image languages are of type-2, and under which copying languages are of type-1. These conditions are set forth in the following theorem, the proof of which is given in the appendix.

(7) a. An n-dependency language $L = \{xu^n wv^n y : n \geqq 0;\ u \neq \phi;\ v \neq \phi\}$ is a type-2 language, unless $w = \phi$ and $u = v$.

b. A mirror-image language $L = \{x\hat{x} : x \in F\}$ is a type-2 language, unless there is a finite string r and finitely many finite strings q and s, such that $F = \{qr^n s : n \geqq 0;\ r = \hat{r};\ s = \phi$ or $s\hat{s} = r\}$.

c. A copying language $L = \{xx : x \in F\}$ is a type-1 language, unless there is a finite string r and finitely many finite strings q and s, such that $F = \{qr^n s ; n \geqq 0\}$.

The argument of Chomsky (1956, 1957) that English is not a type-3 language can now be stated as follows. Let H be English, and let R_1 be the type-3 language:

(8) $R_1 = \{(\text{if})^m$ it rains (then it pours)$^n : m, n \geqq 0\}$.

The intersection of H with R_1 is the n-dependency language L_1:

(9) $L_1 = \{(\text{if})^n$ it rains (then it pours)$^n : n \geqq 0\}$.

Since L_1 is an n-dependency language in which $w \neq \phi$ and $u \neq v$, it follows that L_1 is a type-2 language and hence that English cannot be a type-3 language.

Similarly, Bar-Hillel and Shamir's (1961) argument that English is not a type-3 language can be stated as follows. Again, let H be English, and let R_2 be the type-3 language:

(10) $R_2 = \{(\text{the woman, the men})^+ (\text{watches, study})^+\}$.

Intersecting H with R_2, one obtains the mirror-image language L_2:[3]

(11) $L_2 = \{x\hat{x}' : x \in F_2 = \{(\text{the woman, the men})^+\}$; x' the corresponding string with *watches* for *the woman* and *study* for *the men*$\}$.

Since L_2 is a mirror-image language in which $F_2 \neq \{qr^n s : n \geq 0;\ r = \hat{r};\ s = \phi$ or $s\hat{s} = r\}$, it follows that L_2 is a type-2 language and hence that English cannot be a type-3 language.

Bar-Hillel and Shamir's (1961) argument that English is not even a type-2 language can be stated as follows. Again, let H be English, and let R_3 be the type-3 language:[4]

(12) R_3 = {(the woman, the men)$^+$ and <the woman, the men> (smokes, drink)$^+$ and <smokes, drink> respectively}.

Intersecting R_3 with H, one obtains the copying language L_3:[5]

(13) L_3 = {xx' respectively : x ε F_3 = {(the woman, the men)$^+$ and <the woman, the men>}; x' the corresponding string with *smokes* for *the woman* and *drink* for *the men*}.

Since L_3 is a copying language in which $F_3 \neq \{qr^n s : n \geq 0\}$, it follows that L_3 is a type-1 language and hence that English cannot be a type-2 language.

Finally the argument of Postal (1964) that Mohawk is not a type-2 language can be stated as follows. Let H be Mohawk, and let R_4 be the type-3 language:

(14) R_4 = {a (e, f)$^+$ d b c (e, f)$^+$ d : a = the girl (in Mohawk); b = admires; c = this; d = house; e = the liking of; f = the praising of}.

The intersection of H with R_4 is the copying language L_4:

(15) L_4 = {a x d b c x d : x ε F_4 = {(e, f)$^+$}}.

Since L_4 is a copying language in which $F_4 \neq \{qr^n s : n \geq 0\}$, it follows that L_4 is a type-1 language and hence that Mohawk cannot be a type-2 language.[6]

Thus, there is a valid argument form for the classical arguments that human languages like English and Mohawk are neither type-3 nor type-2 languages. Let us therefore now turn to the challenge to the assumption on which those arguments rest, that at least some human languages contain infinitely many grammatical sentences that are nevertheless unacceptable to anyone that knows any of those languages.

That such an assumption is necessary to the arguments that the theories of type-3 and type-2 grammar are inadequate for human language can be seen upon examination of any of the examples used in those arguments. For example, consider the claim that English contains all of the sentences of L_1 = {(if)^n it rains (then it pours)^n : $n \geq 0$}, but none of the sentences of R_1 = {(if)^m it rains (then it pours)^n : $m, n \geq 0$}, in which $m \neq n$. While, indeed, all of the latter sentences are unacceptable to those that know English, so are all but finitely many of the former sentences. In fact, only two, or at most three, of the sentences of L_1, namely those for which n = 0, 1, and possibly 2, are readily accepted by those that know English. Thus, if English contains all of the sentences of L_1, then infinitely many of the grammatical sentences of English are unacceptable. The same is true for L_2, L_3, and L_4. Clearly therefore the proponents of the classical arguments must provide justification for the claim that the infinitely many unacceptable sentences of languages like L_1 through L_4 are grammatical.

For convenience, let us call the sentences of languages like L_1 through L_4 *crucial sentences*. Let us also call the premiss that all of the crucial sentences of at least some human languages are grammatical the *crucial premis* If only finitely many crucial sentences are grammatical, then the crucial premiss is false, the classical arguments fail, and it follows that the theor of type-3 grammar is the optimal theory of human language.[7] If infinitely many crucial sentences are grammatical, but infinitely many others are not, then whether the classical arguments succeed depends on which crucial sentences are grammatical, and which are not.[8] Finally, if all but finitely many of the crucial sentences of a language are grammatical, then the classical arguments are successful. To simplify the following discussion, we assume that the only possible outcomes are either that all of the crucial sentences of some human languages are grammatical (i.e., that the crucial premiss is true), or that all but finitely many of them in all human languages are ungrammatical (i.e., that the crucial premiss is strictly false).

Three lines of argument have been developed to justify the crucial premiss. First, and perhaps best known, is an argument based on considerations of the simplicity of grammars. Second is an argument based on the observation that the number of acceptable crucial sentences increases, as one removes constraints on linguistic performance. Third is an argument based on the properties of the crucial sentences themselves. Let us consider each of these lines of argument in turn.

The simplicity argument is due to Chomsky.[9] He observes that in order to generate certain undisputedly grammatical sentences in certain human languages, certain rules of grammar appear to be justified. Those rules, if not modified so as to generate just the set of acceptable sentences, also generate certain unacceptable ones (and hence distinguish those unacceptable sentences from others, equally unacceptable; the former being designated grammatical and the latter ungrammatical). Since any modification of the rules so as to limit what they generate to just the set of acceptable sentences is ad hoc in the sense that the modification would serve no other purpose than to effect this limitation, and since such a modification would also complicate the statement of the rules of grammar, it is concluded that no such modification should be made. As a case in point, consider the rules of English grammar that are required to generate the acceptable sentences of L_1. Such rules achieve maximal simplicity and generality if any declarative English sentence is permitted to follow the word *if* and to precede the word *then*. But then, a grammar containing those rules also generates all of the unacceptable sentences of L_1, while failing to generate any of the sentences of R_1, in which $m \neq n$. Since any modification of the rules of English grammar that would serve to render ungrammatical the unacceptable sentences of L_1 would have no independent motivation, and would also complicate the statement of those rules, it may be concluded that all of the sentences of L_1 are grammatical, despite the unacceptability of all but two or three of those sentences. The force of this illustration, with appropriate changes, extends to all of the other examples used in the classical arguments.

However, the simplicity argument is easily rebutted. The fact that the 'simplest' formulation of the rules of grammar that generate the clearly acceptable sentences of a language also generate infinitely many unacceptable

sentences can just as well be taken to mean that the rules are incorrectly formulated, and not that those unacceptable sentences are grammatical. That those rules happen to be simpler than any alternative set of rules that also generates the acceptable sentences but none of the unacceptable sentences is irrelevant, since appeal to simplicity considerations is appropriate only if there is agreement about how to interpret the relevant data, and about what theory of grammar and set of notational conventions to use. In this case, there is no such agreement, since it has yet to be *decided* both how to interpret the acceptability data and what theory of grammar and set of notational conventions are appropriate. Hence any appeal to simplicity considerations in defense of the crucial premiss may be dismissed as irrelevant.

Consider next the argument based on the observation that the acceptability of crucial sentences increases as constraints are removed from linguistic performance. This argument is due to Miller and Chomsky (1963, p. 467), who point out that if a person is given time to reflect on the status of crucial sentences in a language, and if he is also given auxiliary unstructured computation space (for example, pencil and paper) to work out their properties, he finds more such sentences acceptable than he does under ordinary conditions of language use. Miller and Chomsky conjecture that as the availability of time and auxiliary computation space is increased without limit (i.e., as the conditions of idealized performance are approached), eventually all of the crucial sentences of any human language will become acceptable to any person who knows that language.

Certainly, if Miller and Chomsky's conjecture is correct, this line of defense of the crucial premiss is successful. However, if it is incorrect, and only finitely many crucial sentences become acceptable as the conditions of idealized performance are reached, the crucial premiss is not supported. There is, unfortunately, no experimental evidence that either unequivocally supports or refutes Miller and Chomsky's conjecture. If naïve subjects are given large amounts of time and unstructured auxiliary computation space to compute the grammaticality of crucial sentences, they do not generally do markedly better than they do under ordinary conditions.[10] However this result is inconclusive, since it could be maintained that people do not know how to integrate external unstructured computation space with their internal computation space, and that if they did, they would be able to determine the grammaticality of crucial sentences. Hence all that we can say at the moment about the second line of argument in defense of the crucial premiss is that it is inconclusive.

The third line of argument in support of the crucial premiss has not heretofore been fully presented, though traces of it can be found in Chomsky's writings.[11] It is based on the observation that all of the crucial sentences of a human language possess all of the linguistic properties of grammatical sentences, whereas the ungrammatical sentences possess none of them (except, perhaps, by analogy, or by the conventions of metaphoric or poetic use of language). Thus, it may be argued, the unacceptable crucial sentences of a language should be distinguished from ungrammatical sentences by being generated by the grammar of that language and by having their linguistic properties assigned to them by the interpretive conventions of that grammar. The most important of these properties to be considered in an evaluation of this line of argument is that of logical form (or, semantic interpretation, in the narrow sense adopted by Katz (1972)). Thus, consider the crucial sentences of English in L_1', obtained from those of L_1 by substitution of *rains* for *pours*.

(16) $L_1' = \{(\text{if})^n \text{ it rains (then it rains)}^n : n \geq 0\}$

From the logical form of the sentences of L_1', it can be determined that if n is even, the sentences of L_1' are synthetic, being true if the proposition expressed by *it rains* is true, and false if that proposition is false; and that if n is odd, the sentences of L_1' are analytic, being true no matter what the truth value of the proposition expressed by *it rains* is. On the other hand, none of the sentences of R_1' in which $m \neq n$ have logical forms in English, those sentences being ungrammatical:

(17) $R_1' = \{(\text{if})^m \text{ it rains (then it rains)}^n : m, n \geq 0\}$

Given that logical form is a property of English sentences by virtue of interpretive rules of the grammar of English, it follows that the grammar of English must generate all of the sentences of L_1', and by parity of reasoning, all of the other crucial sentences of English, and none of the unacceptable sentences that lack logical forms. To refute this line of argument, the critic of the classical arguments would have to show either that the unacceptable crucial sentences of a language, somehow, lack logical forms, or that logical form is not one of the properties assigned by the grammar of a language to the sentences of that language. Since the first of these possible replies is patently false, and since if the second is true it raises the (I believe) unanswerable question of what the mechanisms for assigning logical forms to the sentences of a language are a part of, if they are not a part of the grammar of that language, the third line of argument in defense of the crucial premiss appears successful. If it is, then the crucial premiss is true, and the classical arguments that the theories of type-3 and type-2 grammars are inadequate for human language are indeed valid.

APPENDIX: Proof of the theorem in (7).

a. If $w = \phi$, and $u = v$, then $L = \{xu^{2n}y : n \geq 0;\ u \neq \phi\}$, which is of type-3. Suppose L is of type-3. Then by the pumping lemma for type-3 languages (Bar-Hillel, Perles, and Shamir 1961; Arbib 1969), there is a positive integer p such that for all sentences $z \in L$ of length p or greater, there are strings u', w', v' such that $z = u'w'v' = xu^pwv^py$, and for all integers $k \geq 0$, $z_k = u'w'^kv' = xu^qwv^qy \in L$. Suppose $w \neq \phi$. Since $u, v \neq \phi$, w must be a substring of w'; i.e., that $w' = rws$. But then, in $z_2 \in L$, w occurs twice, contrary to assumption. Hence $w = \phi$. Suppose $u \neq v$. Then the string uv must be a substring of w'; i.e., that $w' = ruvs$. But then, in $z_2 \in L$, the strings u, v appear out of sequence. Hence $u = v$.

b. If $F = \{qr^ns : n \geq 0;\ r = \hat{r};\ s = \phi \text{ or } s\hat{s} = r\}$, then $L = \{qr^ns\widehat{qr^ns} : n \geq 0;\ r = \hat{r};\ s = \phi \text{ or } s\hat{s} = r\} = \{qr^ns\hat{s}\hat{r}^n\hat{q} : n \geq 0;\ r = \hat{r};\ s = \phi \text{ or } s\hat{s} = r\} = \{qr^m\hat{q} : m \geq 0\}$, which is a type-3 language (recall that q is any of a finite number of fixed finite strings).

Suppose that L is a type-3 language. Then by the pumping lemma for type-3 languages, there is a positive integer p, such that for all sentences $z \in L$ of length p or greater, there are strings u, w, v, such that $z = uwv = t\hat{t}$, and for all integers $k \geq 0$, $z_k = uw^kv = t_k\hat{t}_k \in L$.

Lemma. It is always possible, for all sentences $z \in L$ of length p or greater to pick w such that $u = \hat{v}$.

Proof of lemma. Suppose $u \neq \hat{v}$. Then either (i) $u = u_1 u_2$ and $u_1 = \widehat{u_2 v}$, or (ii) $v = v_1 v_2$ and $uv_1 = \hat{v}_2$. Consider case (i). We have: $z = uwv = u_1 u_2 wv = u_1 u_2 w_1 u_2 v = u_1 u_2 w_1 \hat{u}_1$. Pick new $w' = u_2 w_1$, and set $u_1 = u'$ and $\hat{u}_1 = v'$. Then $z = u'w'v'$, where $u' = \hat{v}'$, and for all $j \geq 0$, $z_j = uw^j v = u_1 u_2 (w_1 u_2)^j v = u_1 u_2 (w_1 u_2)^{j-1} w_1 u_2 v = u_1 (u_2 w_1)^j \hat{u}_1 = u'w'^j v'$, where $u' = \hat{v}'$. Case (ii) is handled similarly. This completes the proof of the lemma.

If $u = \hat{v}$, then for all $i \geq 0$, $z_{2i} = uw^{2i}v = uw^i w^i v$, where $t_{2i} = uw^i = \widehat{w^i v} = \hat{v}\hat{w}^i = u\hat{w}^i$, and hence $w = \hat{w}$.

Also, if $u = \hat{v}$, then for all $i \geq 0$, $z_{2i+1} = uw^{2i+1}v = uw^i w_1 w_2 w^i v$, where $w_1 w_2 = w$, and where $t_{2i+1} = uw^i w_1 = \widehat{w_2 w^i v}$. Then $uw^i w_1 = \hat{v}\hat{w}^i \hat{w}_2 = u\hat{w}^i \hat{w}_2$, and hence $w = \hat{w}$ and $w_1 = w_2$. Setting $u = q$; $r = w$; and $s = w_1$, the result follows for all sentences of L of length p or greater. Since only finitely many sentences of L are shorter than p, their first halves may all be represented by sentences of F, by means of appropriate choices of q and s. Hence the result follows for L as a whole.

c. If $F = \{qr^n s : n \geq 0\}$, then $L = \{qr^n sqr^n s : n \geq 0\}$, which is a type-2 n-dependency language.

Suppose now that L is a type-2 language. By the pumping lemma for type-2 languages, there is a positive integer p such that for any sentence $z \in L$ of length p or greater, $z = xuwvy = tt$, where w is nonnull (and where neither u nor v are null, if L is not a type-3 language), and for all $k \geq 0$, $z_k = xu^k wv^k y = t_k t_k \in L$. Suppose that neither u nor v are null. Then it must be the case that for all $k \geq 0$, $t_k = xu^k w_1 = w_2 v^k y$, where $w_1 w_2 = w$. Setting $x = w_2 = q$; $u = v = r$; and $w_1 = y = s$, the result follows for this case. If v is null (the case where u is null is handled similarly), then it must be the case that for all $k \geq 0$, $t_{2k} = xu^k = u^k wy$, and $t_{2k+1} = xu^k u_1 = u_2 u^k wy$, where $u_1 u_2 = u$. Setting $x = q$; $u = r$; and $u_1 = s$, the result follows for this case, and hence for all sentences of L of length p or greater. Since only a finite number of sentences of L are shorter than p, each of their first halves can also be represented by sentences of F, by appropriate choices of q and s. Hence the result follows for L as a whole.

We note without proof that a copying language is a type-3 language if and only if, furthermore, $sq = r$ or $sq = \phi$.

NOTES

1. For notation, see Chomsky (1963) or Hopcroft and Ullman (1969). In particular, the expression $(x,y,\ldots z)^+$ means any nonnull string made up of any combination of the substrings x, y, ... z in any order.

2. Language L_c is of interest, as Daly points out, because it is a counterexample to Postal's claim that any language of the form $L = \{xx : x \in F$, F an infinite type-3 language over a vocabulary of cardinality $\geq 2\}$ is a type-1 language.

3. The notion of a mirror-image language is easily generalized to include languages inwhich the symbols of the second (backwards) substring correspond word-by-word to the symbols in the first (forwards) substring. However all such languages are of type-2 if the vocabulary of the backwards string is distinct from that of the forwards string (this is a corollary to (7b)).

4. The notation <x, y> indicates a string consisting either of the string x or of the string y; angle brackets here thus stand for what curly braces ordinarily stand for in the statement of the rules of generative grammar.

5. The notion of a copying language is easily generalized to include languages in which finitely many extra words, such as *respectively* in L_3, appear.

6. Rigorous arguments to the effect that type-2 grammars do not have sufficient weak generative capacity for human languages are rare. Besides the arguments of Bar-Hillel and Shamir and of Postal for English and Mohawk, discussed here, I know only of the argument of Huybregts (1976) for Dutch.

7. That is, it follows as far as weak generative capacity is concerned. Given that the sentences of a human language can be weakly generated by a type-3 grammar, it does not follow that they can be strongly generated by such a grammar. For example, it may be the case that a type-3 grammar will not be able to assign all and only all of the structural descriptions of a particular sentence that it generates to that sentence. As I have pointed out elsewhere (Langendoen 1975)), this situation arises for sentences whose structures manifest multiple right- or left-branching. However, since in fact the full phrase-markers for such sentences cannot be recovered by human beings under ordinary conditions of language use, it could be maintained that an optimal grammar should not assign full phrase-markers in those cases. If this conclusion is reached, then the theory of type-3 grammar would be optimal also on grounds of strong generative capacity.

8. For example, if in Mohawk only sentences of the language L_4' of those in L_4 (and similarly for all of the other crucial sentences of Mohawk) are grammatical, then Mohawk could be generated by a type-2 grammar:

(i) $L_4' = \{axdbcxd : x \in F_4' = \{(ef)^n : n > 0\}\}$

9. See, for an early version of that argument, Chomsky (1957, pp. 23-24).

10. We can assume that a linguist who claims that he would accept all of the crucial sentences of a language he knows under conditions of idealized performance has that impression as a result of having consciously formulated rules that generate those sentences in that language. Hence his testimony would be irrelevant, since the conditions under which he accepts those sentences would be richer than those of idealized performance.

11. As in his remark concerning the crucial sentences of English: "They can be understood, and we can even state quite simply the conditions under which they can be true." (Chomsky 1957, p. 23)

REFERENCES

Arbib, M. 1969. *Theories of Abstract Automata*. Englewood Cliffs, N.J.: Prentice-Hall.

Bar-Hillel, Y.; M. Perles, and E. Shamir. 1961. "On Formal Properties of Simple Phrase Structure Grammars". *Zeitschrift für Sprachwissenschaft und Kommunikationsforschung* 14. 143-72.

________ and E. Shamir. 1960. "Finite State Languages: Formal Representation and Adequacy Problems". *Bulletin of the Research Council of Israel* 8F.155-66.

Chomsky, N. 1956. "Three Models for the Description of Language". *IRE Transactions on Information Theory* IT-2.113-24.

________. 1957. *Syntactic Structures*. The Hague: Mouton.

________. 1963. "Formal Properties of Grammars". *Handbook of Mathematical Psychology*, vol. 2, ed. by R.D.Luce, R.R. Bush and E. Galanter, 323-418. New York: Wiley.

Daly, R.T. 1974. *Applications of the Mathematical Theory of Linguistics*. The Hague: Mouton.

Hopcroft, J.E. and J.D. Ullman. 1969. *Formal Languages and Their Relations to Automata*. Reading, Mass.: Addison-Wesley.

Huybregts, M.A.C. 1976. "Overlapping Dependencies in Dutch". *Utrecht Working Papers in Linguistics* 1.23-65.

Katz, J.J. 1972. *Semantic Theory*. New York: Harper and Row.

Langendoen, D.T. 1975. "Finite-State Parsing of Phrase-Structure Languages and the Status of Readjustment Rules in Grammar". *Linguistic Inquiry* 6.533-54.

Levelt, W.J.M. 1974. *Formal Grammars in Linguistics and Psycholinguistics*, vol. 2:*Applications in Linguistic Theory*. The Hague: Mouton.

Miller, G.A. and N. Chomsky. 1963. "Finitary Models of Language Users". *Handbook of Mathematical Psychology*, vol. 2, ed. by R.D. Luce, R.R. Bush and E. Galanter, 419-91. New York: Wiley.

Postal, P.M. 1964. "Limitations of Phrase Structure Grammars". *The Structure of Language: Readings in the Philosophy of Language*, ed. by J.A. Fodor and J.J. Katz, 137-51. Englewood Cliffs, N.J.: Prentice-Hall.